Quality Assessment of Environmental Simulation Scenarios

Scientific, Expertise, Interpretive and Participatory Methods based on Mathematical, Epistemological and Socio-Anthropological Approaches.

June 2012

Thèse présentée en vue de l'obtention du
grade de Docteur en Sciences agronomiques et ingénierie biologique par

Juan David PIÑEROS GARCET

Promoteur:	Pr. M. Vanclooster	(UCL)
Membres du Jury:	Dr. F. Mélard	(ULG)
	Pr. P. Baret	(UCL)
	Pr. B. Feltz	(UCL)
	Pr. P. Bogaert	(UCL)
	Pr. P. Defourny	(UCL)

Université catholique de Louvain (UCL):
Earth and Life Institute/ Environmental Sciences
Louvain-la-Neuve
Belgique

A mon épouse—
magnifique, courageuse, attentionnée.

A mes filles, à ma famille, à ma belle-famille et aux amis. . .

Me voici arrivé, enfin!, aux remerciements, après des années de travail puis d'écriture: mon épouse, mes filles, ma famille et mes amis m'ont ainsi connu (et supporté) écrivant enfermé dans une cabane ardennaise, dans un monastère, dormant dans un bureau et ailleurs encore, ou apparaissant cinq minutes aux fêtes familiales entre deux chapitres. J'entends encore ma fille Julia frapper doucement à la porte du bureau... papa...

Mes pensées et mes remerciements vont donc en premier lieu à ma femme sans qui cette thèse n'aurait jamais pu être terminée et qui m'a soutenu, chouchouté et qui a tenu la barque familiale les moments où j'ai dû m'absenter. Je pense aussi à Julia et Danaé qui pourront davantage profiter de leur père. Je pense à Alice qui a été bien patiente aussi et à qui je promets le retour des promenades au parc. Merci à ma famille, ma belle-famille, à mes amis aussi.

Ensuite, je tiens à remercier le Professeur Marnik Vanclooster, pour la confiance qu'il m'a accordée depuis longtemps et pour son ouverture d'esprit qui ont rendu possible le passage de plusieurs frontières disciplinaires et un retour critique sur nos propres approches, ce qui représente une prise de risque qui n'est pas courante.

L'enthousiasme de F. Mélard pour mon sujet et ses conseils de lecture m'ont été plus qu'utiles, merci!. Merci aussi à l'ensemble de mon jury de thèse pour l'exigence dont il a fait preuve et qui a contribué significativement à la qualité de ma démonstration. Je remercie également Luc Pussemier de m'avoir permis de réaliser une partie de ma recherche participative et sociologique dans un institut habitué aux sciences "dures". Merci à François Goor qui m'a remis en selle après avoir ramassé avec moi mon paquet de pages éparpillées au sol.

Merci à C. (qui est timide) pour la curiosité de toutes choses, au Professeur Laurent qui m'a donné il y a longtemps des axes de départ précieux, à l'ensemble des personnes avec qui j'ai collaboré dans plusieurs projets de recherche (les équipes de APECOP, PEPAM, PRIBEL, PRIBEL pour la région wallonne, HAIR) et aux organismes qui y ont contribué: UCL/GERU, CERVA, la Région Wallonne (DGRNE), la Fondation Roi Baudouin, la Commission Européenne, BELSPO. Je remercie également les trois dames de Talitha Koum et les moines de l'Abbaye de Maredsous pour leur accueil inconditionnel de ce thésard et son baluchon de chapitres en chantier, merci à Philippe V.B. pour son humour antithétique et finalement synthétique.

Abstract

Simulation scenarios, defined as parameters with a physical meaning used with scientific models, are often used to supplement the limits and avoid the costs of experimental approaches, and to forecast the environmental consequences of regulatory decisions. Their quality assessment is important to support decisions in the public domain and avoid choices with unwanted effects. Building on mathematical, socio-anthropological, philosophical, and scientific developments, this study defines a generic framework for scenarios quality assessment, develops specific methods and adequacy criteria for regulatory scenarios and tests their applicability, limits and interest.

Quality is defined as the adequacy for a given use. Representativeness, theoretical consistency, logical consistency, juridical conformity and political criteria are defined. Two routes for quality assessment are defined: an expertise route making interpretation choices transparent, and a participatory route opening interpretation choices to citizens. Those routes are tested, respectively, with case studies of diffuse pollution simulation scenarios for groundwater and bees exposure to pesticides. The framework allows to assess multiple quality components, taking into account the links between them (including the juridical, scientific and political components) while preserving their specificity. It highlights the limits of the scientific knowledge in a regulatory context, makes transparent the choices and uncertainties, and allows to assess the social acceptability of such uncertainties and of the use of the scenarios. Due to the lack of a binding political mandate in our case studies, the possibility of the incorporation of the framework in formal democratic processes is still to be demonstrated.

Epigraphs

Few terms are used in popular and scientific discourse more promiscuously than "model". A model is something to be admired or emulated, a pattern, a case point, a type, a prototype, a specimen, a mock-up, a mathematical description - almost anything from a naked blonde to a quadratic equation - and may bear to what it models almost any relation of symbolization. (Goodman, 1976-1968)[1]

[B]ecause validation involves the art of modelling and the philosophy of science, validation will remain controversial![2] (Kleijen, 1999)

[L]a preuve, au sens où elle est scientifiquement fiable [n'est pas] à confondre avec les modèles [...] qui dépendent, comme le langage ordinaire, des facteurs que l'on a choisi de mettre en scène et ne testent donc que les conséquences de la mise en scène [...]. Ainsi, dès que l'on quitte [...] la question de la preuve, on entre [...] dans le domaine des négociations, et donc dans l'agencement des rapports de force qui imposent le cas échéant la prise en compte de facteurs 'gênants', ceux, notamment, dont la mise en scène de laboratoire a réussi à faire l'économie [...]. (Stengers, 2000)

La démarche anthropologique prend comme objet d'investigation des unités sociales de faible ampleur à partir desquelles elle tente d'élaborer une analyse de portée plus générale, appréhendant d'un certain point de vue la totalité de la société où ces unités s'insèrent. (Augé, 1979)

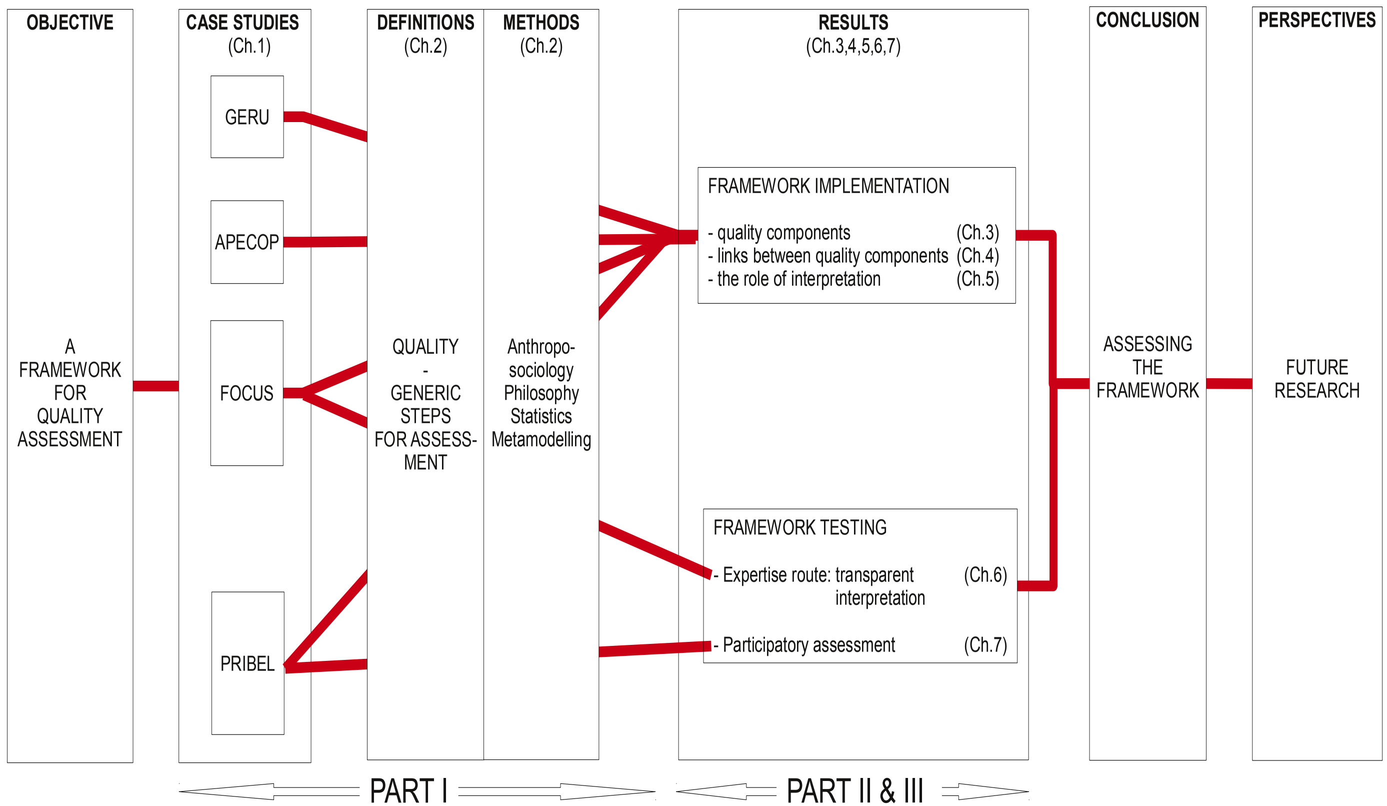
[1] For a re-published work, two years are mentionned: the first year refers to the consulted edition, and the second refers to the first edition.

[2] Bracketed text indicates a modification to the original text, a bracketed ellipsis "[...]" indicates deleted text, "[sic]" indicates that the preceding text is as in the original source.

Contents

Reading guide

Figure 1: *Reading guide for this thesis.* Using four test cases described in chapter 1, chapter 2 defines a framework to perform forecasting regulatory simulation scenarios quality assessment, and proposes a method to implement and test it. Chapter 3, 4, 5 presents the implementation, while the test of the framework is provided in chapter 6, 7. The framework is discussed and assessed in the conclusion. Finally perspectives and indications for future research are given.



Keywords

Simulation scenarios; scientific, logical, juridical and political quality assessment; environmental regulation; pesticides; groundwater; bees; representativeness; acceptability; expertise; citizens; participation; interpretation; transparency; anthropological laboratory studies; sociology of translation; actor-network theory; philosophy of representation; empirical adequacy; metamodels; neural networks; density mixture.

Foreword

After a few years working with models and scenarios, I came to think that modelers rely on some sort of relation between their scenarios or models and the reality, but without having an analysis of such a relation. This situation left me with an uncomfortable feeling, since the basis on which the modelling work rests – its ultimate reference – seemed to me to be unclear. I began, on the one hand, to listen to what other people involved in this kind of work were saying, and then I interviewed them formally. On the other hand, I began to develop an interest in developments in epistemology in order to build a foundation that seemed more solid with regard to my relation to reality through scenarios. Finally, the fact that I worked in applied projects for pollution regulation introduced the curiosity of understanding better the links between the modelling work, its results, and the social context in which it evolves. This is how the project of this PhD began.

Introduction

In environmental pollution science, and in particular in a regulatory context, a simulation model is generally considered as some kind of simplification of the reality (e.g. *''a simplified representation of a part of reality that contains mutually dependent elements''* (FOCUS, 1997)). According to Lamotte et al. (2009): *''the development of the model involves selection of the state variables representing the system, selection of the relations existing between these variables and of the parameters that characterise them and, finally, selection of the external factors affecting the functioning of the system [...]''.*

According to Grémy (2009), simulation is *''experimentation on a model. It is a scientific research procedure that consists in creating an artificial reproduction (model) of the phenomenon one wishes to study, in observing the behavior of this reproduction when making experimental variations of the actions that may be performed on the model, and from this inferring what would happen in reality under the influence of analogous actions.''* Yet another definition: *''Simulations characteristically are used in connection with dynamic models, i.e. models that involve time. The aim of a simulation is to solve the equations of motion of such a model, which is designed to represent the time-evolution of its target system. So one can say that a simulation imitates a (usually real) process by another process''* (Frigg and Hartmann, 2009). However, in many cases the word ''simulation'' is also used to refer to static models, without involving time (see e.g. Williams and Albertson (2005)). The word ''model'' has other meanings in science and logics that are reviewed e.g. in Frigg and Hartmann (2009), Mouloud et al. (2009)

and Nouvel (2002), but these meanings are not directly relevant for describing the diffuse pollution models we are dealing with.

Computer based simulation scenarios are defined here as series of values that are given as inputs to computer models (programs), which return outputs values called the simulation results. In that sense, a simulation is one complete execution of the computer program, including input and outputs. Scenarios are sometimes themselves constructed from simulation results. The scenarios we are working with are at the same time scientific objects (parameters with a physical meaning to be used with models) and regulatory objects (they were developed with the *purpose* to be *used* in a regulatory context). These two aspects are closely linked. Moreover, we are focusing on forecasting scenarios, meaning here that the attention when performing the simulations is on what can be known in advance, from present and past evidence and/or relevant theories, in order to support the regulatory decisions. Simulation scenarios are often used in regulation for supplementing the limits of the experimental approach: they are useful for spatial upscaling and extrapolation, and in many cases to forecast the likely environmental or health consequences prior to the actual production or marketing of the product being regulated.

Decisions based on scenarios that are not adequate for a regulatory use are likely to have unwanted consequences, such as the marketing of pesticides that have significant potential to leach to the groundwater. Quality assessment of regulatory scenarios is thus important because they support decisions in the public domain. The key question is how to define the meaning of the "quality of regulatory scenarios", and the methods for assessing it.

Actually, "scenarios" is a word with multiple meanings. Regarding the common language use, Makins (1991) defines it either as *a summary of the plot of a play, etc., including information about its characters, scenes,...* , or as *a predicted sequence of events*. We find sometimes definitions around the theme of *"circumstances under which something can append (e.g. traffic accidents' scenarios)"*. Sapio (1995) made a review of various meanings of the word "scenario", amongst which we find: *"a description of a certain aspect of the state of the world"*. Mahmoud et al. (2009) made a categorisation of scenarios for their use in natural resources management by scenario planning[3]: *''Scenarios are possible future states of the world that represent alternative plausible conditions under different assumptions''*.

Scenarios as considered in the present work —and as most often used in diffuse pollution modelling—are not made for scenario-based strategic planning. Rather, they are generally speaking to be located in the area of what is sometimes called traditional forecasting, as opposed to scenario planning:

> *[t]raditional forecasting techniques are founded on the hypothesis that the future world will be broadly similar to the present one, so they cannot anticipate primary changes in the environment [...]. Cause-effect relationships might permit rebuilding the past and extrapolate the future from the present, in a way we can call ''deterministic'', provided that we have adequate information about the current state, highly sophisticated models and unlimited power of calculation.* (Sapio, 1995)

While there is a continuum or combinations between scenarios used in traditional forecasting and scenarios for strategic planing, generally speaking the latter are more adapted to situations with high uncertainties, lack of knowledge, possibilities of paradigm shifts, and difficulties to formalize. In that sense, the scenarios we are dealing with in the present work corresponds to the information about the current state. Moreover, models are often used to make predictions when experimentation is not possible (the model serves as a substitute - simulation -for the functioning of a system), and the traditional forecasting scenarios often *represents* the characteristics of the system that one can vary depending on what one wishes to simulate. As an example of definition of a scenario used in the context of the regulation of pesticide diffuse pollution, we find:

> *[a] scenario [is] a unique combination of agronomic and environmental conditions that realistically* represents *significant areas within which conditions are relatively homogeneous with respect to modelling input parameters. These*

[3]Scenario planning, is a method for strategic planning under high future uncertainties using scenarios, by combining ''known facts about the future'' with plausible alternatives, to construct a set of alternative scenarios about the future.

> *conditions include climate, hydrogeology, surface water characteristics, soil and topography.* (FOCUS, 1997)

Such a definition poses scientific challenges. How can we define the representativeness of a scenario? Can it be measured or quantified? If it can not be measured, or confronted to some empirical evidence, can it be approximated or deduced through modelling? How to estimate the uncertainties of such a deduction?

In hydrological modelling, there exists an extensive literature about model validation (see e.g. Vanclooster et al. (2000)). However, regarding scenarios used in traditional forecasting for hydrology, we found no systematic account of scenarios validation, assessment or evaluation (as opposed to available formal frameworks for the case of scenario based strategic planning, e.g.Mahmoud et al. (2009)). Of course, scenarios or input parameters are considered in traditional hydrological forecasting under various aspects like the quality of the data from which they are derived, the number of samples needed, methods of measurement, parameter estimation methods and so on (e.g.Wosten et al. (2001)). Vanclooster (2002) and Vanclooster et al. (2003) presented some lines along which scenarios can be validated. Vanclooster et al. (2003) states that *"[q]uality assurance of exposure modelling implies the implementation of a good modelling practice, and the validation, documentation and maintenance of the modelling codes and scenarios."*

From the socio-anthropological side, many questions can also be raised about scenarios. Actually, when they are taken out of their scientific origin and introduced in risk assessment processes for environmental regulation (e.g. chemicals or pesticide), a purely scientific definition shows many limitations, and other points of view are to be taken into account to define them. How do scenarios adequate to a given situation? Which is the language chosen and how much is it efficient to the goal of the persons making the scenarios? What is the political function of a scenario? How are they integrated into the legislation? How are they interacting with the economical actors? What kind of negotiations are encountered when scenarios are established? How are they linked to the broad social context in which they are located?

Scenarios are also a philosophical challenge. While a representation-based definition matches the one given by scientists, it nevertheless presents its own problems. What is representation? What is reality?

What is a "good" representation? What are the relations between measurements, scenarios, and reality? Are scenarios representing the reality, or phenomena, or appearances, or arbitrary choices? Is representation a dual relation between a representor and a represented, or is-it a triangular relation between a subject, the representor and the represented? If a subject is involved, what is its role? Does the user introduce some kind of subjectivity or interpretation? How can scenarios represent something, if we consider the existence of scenarios that are not derived from measurements? Are they representing some kind of virtual objects, or concepts? Which is the relation between scenarios and theory? Are scenarios representing in the same way, if they are used for different purposes?

It is thus clear that a definition of scenarios from the modelling science becomes too narrow in order to take into account those multiple aspects of the scenarios. A wide approach is particularly important when scientists are confronted with demands for "valid" regulatory scenarios. Indeed, a request for validation combines the scientific, epistemological and anthropological challenges we just mentioned. It adds an additional challenge: is "validation" a good starting point? Is-it possible to define criteria to declare that a scenario is "valid"? On which basis? Is representativeness the only criteria as suggested by the scientific definition of a scenario?

The quality of simulations for predicted (forecasted) environmental concentration (PEC) of pesticide in regulatory assessments, will depend on both the quality of the PEC model and of the scenarios. We advocate that traditional forecasting methods like the ones used in the EU pesticide regulation can benefit from methods and concepts developed in data quality assessment, scenario-based strategic planning, philosophy of science and anthropology, because they allow to consider the quality of scenarios in a much broader meaning.

Considering such a large range of questions raised by the assessment of scenarios, it is clear that the development of a method for scenarios quality assessment must include the development of a solid and multidisciplinary theoretical basis.

The present dissertation has three objectives: (1) to define scenarios quality, and (2) to implement concepts, methods and means for a quality assessment of forecasting regulatory simulation scenarios, and (3) to test their applicability, limits, and interests.

Our objectives have thus normative and theoretical aspects on the one side, and pragmatical, methodological aspects on the other side.

The *scenarios' quality assessment* methods we develop must be useful as tools in an environmental regulation context. Thus, our concepts and criteria of scenario quality must be able to cover the following questions. What is a high quality scenario, from the regulatory point of view (thus at the cross-point between legislation and politics)? What is a high quality scenario from the scientific point of view?

Standard scenarios as the ones used for pesticides regulation, are complex combinations of different soil, climate and land use attributes and may therefore conform to the envisaged use in some aspects, but not in others. Scenario assessment may result in suggestions to improve the scenario, to reduce the uncertainty in the final (regulatory) simulations, and to take into account the uncertainty of the decisions based on them. Our aim is thus to develop a quality assessment framework that should evaluate the quality of a scenario in its multiple aspects in a transparent way, and give insights about (1) the influence of the scenarios quality on the use of the scenarios, (2) potential improvements to the scenarios, and (3) potential improvements or changes to the scenarios use.

In order to achieve those objectives, Part I presents the materials and methods and defines a framework for the assessment of regulatory simulation scenarios. Part II implements the framework, while Part III tests the framework with case studies. The framework is discussed and assessed in the conclusion. Finally perspectives and directions for future research are given.

The structure and contents of the chapters is as follows:

- Chapter 1 (*Case studies presentation: actors and scenarios*) presents the GERU laboratory, APECOP project, FOCUS scenarios, and PRIBEL scenarios case studies;
- Chapter 2 (*A Framework for quality assessment*) defines the concept of *quality* and presents *generic steps* to perform scenarios quality assessment, and proposes a method to implement and test those steps.
- Chapter 3 (*Defining the quality components to be assessed*) implements the framework by defining the *the scenarios quality components* to be assessed
- Chapter 4 (*Determining the links between quality components*) demonstrates that *links between the quality components* exists, due to the

necessity of interpreting the scenarios for assessing their quality. This is done through an anthropological observation.

- Chapter 5 (*The role of interpretation in quality assessment*) provides concepts and methods for taking into account the links between quality components into the quality assessment. Concepts from hermeneutics, philosophy, and sociology are used. Two quality assessment routes are identified: a quality assessment through the expert route, equipped with transparency of the interpretation, and a quality assessment through the participatory route, that opens interpretation of the scenarios to citizens.
- In Chapter 6 (*Making interpretation choices transparent through the scientific expertise route*), FOCUS scenarios go through the expert route, for assessing the representativeness aspect of their scientific quality component.
- In Chapter 7 (*Opening interpretation choices through the participatory route*), the PRIBEL scenarios go through the participatory route, for testing the possibility and cognitive added value of an assessment of their political quality component by citizens.

In order to locate the reader in our demonstration, two reading guides are provided at the beginning of each chapter: a flowchart locating its contribution amongst the other chapters, and a second one indicating the contribution of the chapter to the implementation or test of the quality assessment framework generic steps.

We point out to readers that, as with the problems we are dealing with, the theoretical disciplinary foundations used are multiple: this multidisciplinary work is located at the intersection of the natural sciences, anthropology and philosophy, with socio-political aspects. Thus, for helping the reader, some theoretical background is given in appendices (about rhetoric, pragmatics, sociology of translation and actor-network theory, symbolic objects formalism, Latour's power/collective/state concepts, and representation in philosophy), and reference is made to them in the corresponding chapters.

Part I

A Generic Framework for Quality Assessment: Definition, Materials, Methods

Summary of Part I

Part I presents, in two chapters, the case studies, definitions and methods for the implementation and test of a quality assessment framework for regulatory forecasting simulation scenarios:

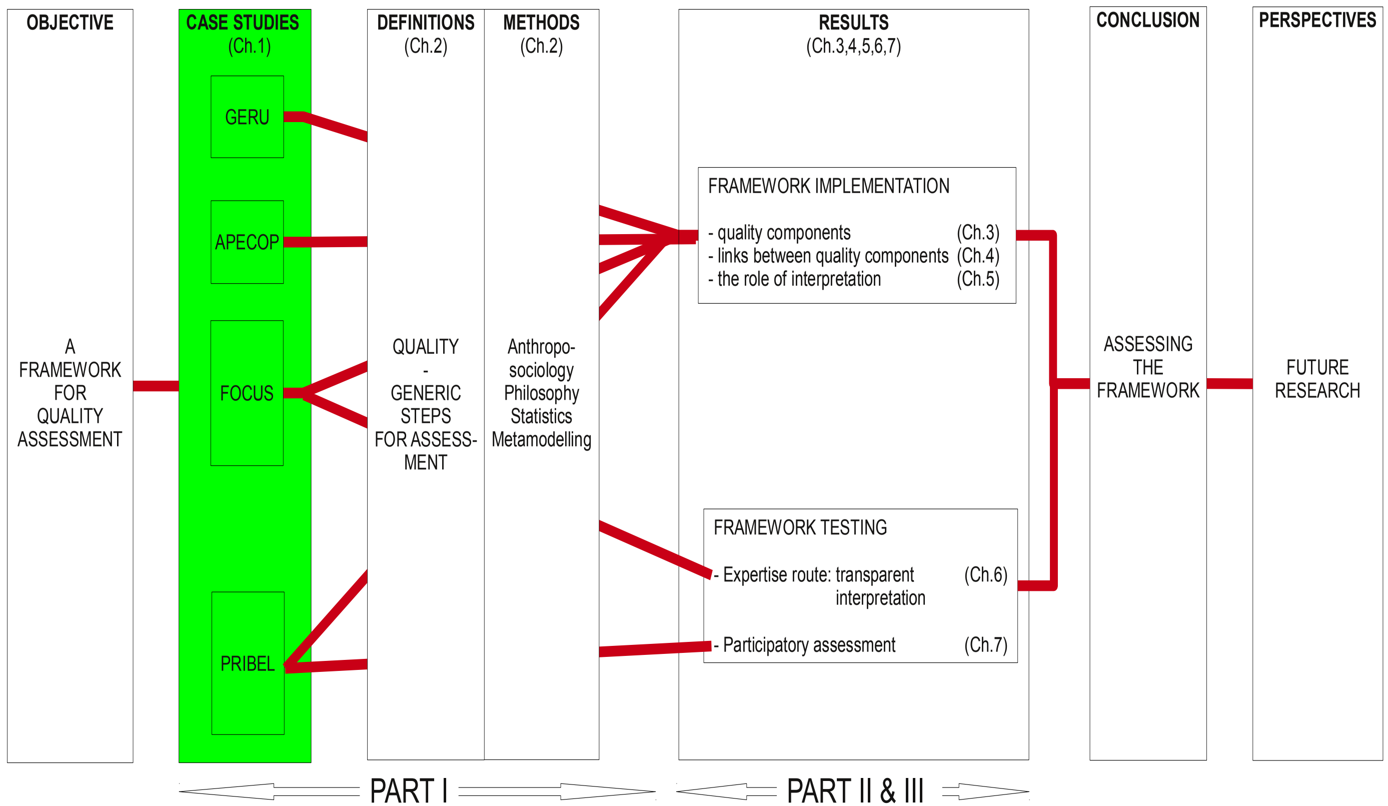
- chapter 1 presents the GERU laboratory, APECOP project, FOCUS scenarios, and PRIBEL scenarios case studies;
- chapter 2 defines the concept of *quality* and presents *generic steps* to perform scenarios quality assessment, and proposes a method to implement and test those steps.

Parts II and III make use of such a method for respectively implementing and testing the framework.

CHAPTER 1

Case studies presentation: actors and scenarios

Figure 1.1: *Reading guide of chapter 1.* This chapter presents four cases that will be used for implementing and testing the quality assessment framework generic steps. GERU is a water resources modelling academic laboratory, APECOP is a EU financed project aiming at the "validation" (quality assessment) of the FOCUS regulatory scenarios for groundwater exposure to pesticides, PRIBEL are regulatory scenarios for bees exposure to pesticides.



Summary

This chapter presents four case studies that are subsequently used in part II and III for implementing and testing the quality assessment framework:

- **Researchers in the GERU academic laboratory.** The field of expertise, history, and membership of the laboratory is presented here. The laboratory is studied in chapter 3 from the point of view of the concepts of its members about models and scenarios.
- **Researchers, regulators and industrials involved in various ways in the EU funded APECOP project**, aiming at the "validation" of the FOCUS scenarios. The negotiations in APECOP around what the FOCUS are supposed to represent is what is studied in chapter 4.
- **The FOCUS regulatory scenarios for groundwater exposure to pesticides** used in the 91/414/EEC Directive assessment of the risk of leaching at the EU level. The FOCUS parameters *use* determines what they represent. They are claimed to represent "realistic worst cases" in the directive 91/414/EEC, which is interpreted by FOCUS as an overall vulnerability to pesticide leaching approximating the 90th percentile of all possible situations, when used with the FOCUS models. An expert judgement leads to the choice of soil and climate parameters; the empirical basis for determining the parameters is rather weak. The FOCUS scenarios are used in chapter 6 for testing our concept of assessment of the representation.
- **The PRIBEL regulatory scenarios for bees exposure to pesticides**, used for aggregated risk calculations by the Walloon region and by the Belgian Federal pesticides use reduction plan). PRIBEL parameters are claimed to represent the reality in a different way than FOCUS, i.e. in the empirical sense, by being measured or survey-derived values (values of the amount of each pesticide used per year, pesticide half-live, distribution of crops per area in Belgium). This provides a strong base to the parameters choice, which is also a difference with the FOCUS scenarios. PRIBEL scenarios are used in chapter 7 for testing a participatory assessment method.

The descriptions of the scenarios are termed "narrations". The FOCUS and PRIBEL scenarios are constructed with inputs from industry, scientists and regulators, and there were various steering committees in

which they participate. They do not involve participation from NGOs or citizens in their construction. The way those case studies are used for implementing and testing the quality assessment framework is described in detail in chapter 2. The FOCUS and PRIBEL case studies are compared in table 1.1, which also indicates how they relate to the quality assessment cases we present in this dissertation.

Table 1.1: *FOCUS and PRIBEL case studies comparison* regarding their regulatory context, developers, development method, users, assessment projects, assessment results matrices (role in regulation, construction choices, their reasons, concerns, evaluation, improvement suggestions), simulation scale, representation mode (the way they represent as determined by their use, and the concept describing what they represent) and DPSIR descriptors of the environmental simulations made with them (driving force, pressure, state, impact, response).

	FOCUS groundwater scenarios	PRIBEL bees scenarios
Regulatory context	Council of the European communities Directive 91/414/EEC (placing plant protection products on the market)	Belgian pesticides reduction plan (Bel. Off. J.: M.B.11.03.2005); environmental reports (Walloon 27/05/2004 decree)
Developers	Regulators, industry and scientists (academia and public research centers) FOCUS (2002)	Data from regulators, industry, academia. Build by scientists (Piñeros Garcet et al. (2006b), academia and pub. res. centers)
Development method	Mainly values from expert judgement, no external participation besides developers.	Mainly measured or survey-derived values. Feasibility of co-construction with citizens explored in (Piñeros Garcet et al., 2006a) and in chapter 7
Users	Industry, EU and Member States Regulators	Belgian federal and regional regulators, sectoral thematic stakeholders groups of the reduction plan
Assessments presented in this dissertation	(1) APECOP project (Vanclooster et al., 2003), an assessment made by scientists (studied as an anthropological test case in chapter 4); (2) a second expert assessment of FOCUS, presented in chapter 6, with an improved model, applying the representativeness concept developed here, and aiming at transparency of the scenario interpretation choices.	(1) PEPAM project (Piñeros Garcet et al., 2006a): a participatory modeling exercise with citizens and stakeholders; (2) an analysis of PEPAM from the quality assessment perspective is presented in chapter 7, highlighting the opening of the interpretation choices to citizens and studying the way citizens assess scenarios.
Assessment results matrix	Role in regulation, construction choices and their reasons presented in this chapter and in chapter 4; concerns, evaluation, improvement suggestions presented in chapter 6.	Role in regulation, construction choices and their reasons presented in this chapter and in chapter 4; concerns, evaluation, improvement suggestions presented in chapter 7.
Simulation scale	EU, 20 years	Belgium, Walloon Region, one year
Representation mode	Functional representation: represent "realistic worst cases" of leaching, *when used* with FOCUS models.	Measurement representation: are used to represent *by being measured and survey-derived values* characterizing pesticides, their use, and land use.
Driving force	Agriculture	Agriculture
Pressure	Pesticides emission	Pesticides emission
State	90th percentiles of annual leaching concentration to groundwater	Bees exposure
Impact	Not modelled (regulation requires concentrations below $0.1\mu g/l$ for drinking water)	Bees acute risk
Response	Authorizing or not the pesticide on the EU market	Change pesticides use patterns and/or reducing quantities used, in Belgium

1.1 Introduction: scenarios narration

In the field known as *scenario planning*, scenarios are constituted by detailed *scenario's narratives* which are detailed storylines exploring plausible futures with a strong qualitative component (see e.g. Strengers et al. (2004) for the the Intergovernmental Panel on Climate Change (IPCC) emission scenarios, or L.Pintér et al. (2008) for the UNEP integrated environmental assessment method). It is a common characteristic of models and scenarios used in traditional forecasting like the ones for environmental fate of chemicals to be accompanied and documented by some sort of description, by a manual, by an explanation of what they are representing or how to use them. For the case of models, this is what Weisberg (2008) calls a *model description*. Such descriptions often lack what Piñeros Garcet et al. (2006a) are calling a *model equipment*, documenting in specialized tables how stakeholders and citizens reacts to choices in the model and the consequences of such choices.

We propose to call *scenario's narration*[1] the description of the whole context in which scenarios are used, the history of their genesis, the embedded choices and their consequences once they are used for modelling and regulatory purposes, the ex ante or ex post opinions of citizens and stakeholders. The term "narration" has the advantage of emphasizing the intervention of somebody narrating, and narration is here a construal, i.e. an interpretation in a particular way. Scenarios' narrations can be of a more or less generic nature without making an explicit reference to a specific model, or be designed to be used with various models as in the case of the FOCUS scenarios. Scenarios' narrations are thus partially independent from the models.

Part II will demonstrate that *scenarios' narrations* are important because they are the basis for scenarios interpretation, that, as we will show, is necessary for scenarios' quality assessment. Here below we provide models and scenarios' narrations, partially written by us, and partially as provided by the scenarios' designers themselves. Those narrations are however not complete since they lack e.g. citizens opinions: that aspect will be left for chapter 7 where a method for confronting citizens to simulation scenarios is presented.

[1]Narration is "a narrated account or story", as defined by Makins (1991).

1.2 The FOCUS-GW scenarios: pesticide leaching to groundwater

The FOCUS-GW scenarios are presented here in great detail because they are used as our main test-case for the scenarios' representativeness assessment concept we are proposing (chapter 6). The FOCUS groundwater (FOCUS-GW) scenarios' narration we present here relates to the period between 1991 and 2006. References to the regulatory texts are mainly to the versions of that period. Further developments of the regulatory context took place, e.g. new proposals for the way the FOCUS-GW scenarios are to be used, or the revision of the 91/414/EEC Directive (Anonymous, 1991) concerning the placing of plant protection products on the market. Some of those new developments are considered here, but a full description of them is not provided.

With the adoption of Council Directive 91/414/EEC of 15 July 1991, a harmonised legal framework was set up for the regulation of plant protection products in the European Community. A central decision-making regime for determining the acceptability of active substances contained in plant protection products was established on the basis of harmonised data requirements detailed in Annex II (active substance) and III (plant protection product) of the Directive[2].

Pesticide is a major chemical product in terms of market volumes. Development time (for industry) is long and costly, and industry wants to make profitable its huge investments before the patents falls in the public domain. At the same time, those products are under particular scrutiny from the public and the NGO's. Problems of pesticide pollution in groundwater are observed in aquifers, that have lead in some cases to restrictions to products that were already in the market. Agriculture is becoming increasingly dependent on pesticide. The 91/414/EEC Directive, and the FOCUS scenarios are then playing in a field with very important tensions. Industry is demanding for a regulation at the European level to allow them to make strategic planing in a science based stable and predictable regulatory environment.

The 91/414/EEC Directive specifies[3] that expert judgement will take into account predicted environmental concentrations in groundwater (PEC_{GW}) (i.e. simulation results) for deciding if lysimeter or field leach-

[2] Specific provisions exists for authorizations at the national level.

[3] Excerpt in Annex A.1.

ing studies are needed. It is common in pesticide registration procedures to proceed by stages, or tiers of questions. In first-tier assessments, models and scenarios are used to get a first indication of the leaching potential of a pesticide. PEC_{GW} is what calculated by the FOCUS-GW models and scenarios. Annex II refers to Annex III[4] where precisions are given about the way simulations (PEC_{GW}) should be done. It asks for "*suitable estimations (calculations)*" "*realistic parameters*", "*reliably validated with measurements*", "*relevant to the conditions in the area of use*", "*relevant agricultural, plant health, and environmental (including climatic) conditions*", and states that "*a realistic worst-case estimation must be made*". Annex IV[5] indicates how the authorization decision should be taken: "*No authorization shall be granted if the concentration of the active substance or of relevant metabolites, degradation or reaction products in groundwater, may be expected to exceed*" [a given concentration], "*as a result of use of the plant protection product under the proposed conditions of use*".

Besides those legal references, there is no official groundwater decision scheme for Annex I inclusion. Since the adoption of the 91/414/EEC Directive, guidance documents have been developed with the support of all Member States or by specific expert working groups, in order to further facilitate the harmonisation of evaluation procedures and decision making criteria, and to give guidance to industry on how to prepare "dossiers" and Member States on how to prepare draft evaluation reports ("monographs"). Examples of these guidelines are the ones prepared by the FOCUS [6] groups (FOrum for the Co-ordination of pesticide fate models and their USe) starting in the year 1993. FOCUS members and contrib-

[4]Excerpt in Annex A.2.

[5]Excerpt in Annex A.3.

[6]Presently, the role of FOCUS is being revised into the framework of the EFSA (European Food Safety Agency). In 2002, the registration of PPPs came under the responsibility of EFSA which setup the Panel on plant protection products and their residues (PPR) for giving scientific advice on PPPs registration, and the FOCUS group itself stopped to be supported by DG-SANCO in 2008. However, the FOCUS scenarios and models themselves are still used in registration, even if EFSA sometimes advices to modify part of its parameters (see e.g. EFSA (2007)). In order to take into account the work done by the APECOP project (Vanclooster et al., 2003), FOCUS is presently considering to modify two of the scenarios because of overestimated organic matter content in soils leading to underestimation of the leaching. Moreover, FOCUS is having discussions on the future use of its scenarios, and one possible recommendation is to supersede the lower tier assessments with higher tier assessments if the lower tiers shows evidence of risk.

utors are from industry, regulators, universities and research institutes. Industry took an important role in the FOCUS scenarios elaboration. In the framework of his work, FOCUS developed simulation scenarios for PEC_{GW} estimation. This is in fact a way to interpret and standardize the 91/414/EEC directive requirements regarding the way the expert judgement takes into account the simulation results, and it is also a standardization of the models themselves. FOCUS was the main forum where such a practice was discussed and formalized, regarding the use of models and scenarios. According to the FOCUS guidance documents[7], (FOCUS, 2002) "*The remit of FOCUS was to develop consensus amongst the Member States, the European Commission, and industry on the role of modelling in the EU review process of active substances*". "*Any PEC model calculation assumes a scenario which is therefore an important element of the guidance*". FOCUS (2002) states also that the scenarios "*increase the consistency of the regulatory evaluation process by minimising the subjective influence of the person who performs the PEC calculation*".

Only a limited number of ground-water scenarios were, for pragmatic reasons, identified in the FOCUS-GW procedure. Nine realistic worst case scenarios were developed by FOCUS, to be used with 4 chromatographic-flow "FOCUS models" (deterministic models PELMO, PEARL, PRZM, MACRO). The current practice is to demonstrate at least one safe use on a representative crop in a significant area of Europe: "*[i]f one or more of these relevant scenarios result in predicted groundwater concentrations less than 0.1 $\mu g/l$, then in principle the active substance could be included on Annex I*" FOCUS (2002)[8,9].

When a substance exceeds 0.1μg/L for all relevant scenarios, then Annex I inclusion would not be possible unless convincing higher tier assessments results demonstrate acceptable use. The higher tier assessment comprises both the use of specific scenarios/site specific data and data from lysimeter studies or field testing as model input.

FOCUS (2002) gives various indications, sometimes rather contradictory, about what the scenarios are supposed to represent. A full analysis of this will be presented in chapter 6, where we submit the FOCUS

[7]Excerpts in annex A.4.

[8]Excerpts in annex A.4

[9]For national assessments, if the compound cannot be used safely throughout the country, then the registration may be limited to the subset of conditions under which the compound can be used safely.

scenarios to a quality assessment of their representativeness, i.e. the quality of their representation. The determination of ''what the scenarios represent'' is a matter of interpretation, as we will see in the following chapters. According to our interpretation, FOCUS scenarios are claimed to represent major agricultural regions, span the range of temperature and rainfall occurring in EU arable agriculture, describe an overall vulnerability to leaching approximating the 90th percentile of all possible situations, and to avoid unrealistic combinations of climatic and soil properties; they do not ''mimic specific fields''[10].

For each scenario, soil, weather, and 25 crops were parametrised, and tables with the detailed parameters are given in the FOCUS report (FOCUS, 2002). Those detailed parameters values are what the scientists calls the scenarios[11], typically the scenario's description is not part of the scenario. Table 1.2 shows a summary of some of the parameters.

Scenarios	**Mean Annual Temp. (°C)**	**Annual Rainfall (mm)**	**Topsoil†**	**Organic Matter (%)**
Châteaudun	11.3	648 + I*	silty clay loam	2.4
Hamburg	9.0	786	sandy loam	2.6
Jokioinen	4.1	638	loamy sand	7.0
Kremsmünster	8.6	900	loam/silt loam	3.6
Okehampton	10.2	1038	loam	3.8
Piacenza	13.2	857 + I*	loam	1.7
Porto	14.8	1150	loam	6.6
Sevilla	17.9	493 + I*	silt loam	1.6
Thiva	16.2	500 + I*	loam	1.3

† = USDA classification (USDA, 1975; FAO, 1977)
I* = scenario also includes irrigation

Table 1.2: *Some important parameters for the FOCUS scenarios. The soil parameters (texture, organic matter) corresponds to surface horizons. As part of the scenario selection process, targets for annual rainfall and mean annual temperature where also selected. Four scenarios were identified as having irrigation normally applied to at least some crops in the corresponding area. Table adapted from FOCUS (2002).*

Vulnerability is mentioned in the description as part of what is represented by the scenarios. Since soil and climate were chosen as vulnera-

[10] See e.g. ''*[t]he scenarios should describe an overall vulnerability approximating the 90th percentile of all possible situations (this percentile is often referred to as a realistic worst case)*'' in FOCUS (2002). See more excerpts in annex A.4.

[11] FOCUS calls that the *scenarios specifications*.

bility indicators, this explains why they are important and were chosen to be mentioned in table 1.2 by the scenarios' designer.

1.3 The PRIBEL bees scenarios: acute toxicity of pesticide to honey bees

The PRIBEL scenarios, models and some of their developments are presented here because they are used in chapter 7 as a test-case for our participatory scenarios assessment and construction method. They also illustrate another regulatory use of scenarios, that will be used in chapter 3, together with the FOCUS case, for defining the quality components of regulatory scenarios.

In a communication from 2002, the Commission of the European Communities made a proposal for a thematic strategy concerning a sustainable use of pesticide (Commission of the European Communities, 2002). The problem was also addressed at the Belgian level through a plan for reduction of pesticide[12]. According to those plans, various measures for reducing the risk of pesticide use are studied. Both include an assessment of the impact of pesticide on human health and environment using indicators. It was proposed to make a risk assessment on a yearly basis.

The situation in 2006, in Belgium was that the plan proposes (amongst others) to reduce the risk associated with pesticide use, as well as the amounts used. It was expected that a decrease of pesticide use would decrease the risk on human health and on the environment (of course this assumption is valid if the reduction in doses is not counterbalanced by an increase in toxicity towards non-target organisms). To support implementation of the plan, the PRIBEL model (Pesticide Risk Indicator for Belgium) was designed[13,14,15]. PRIBEL is constituted of a limited number of indicators (seven) allowing the assessment of the risks asso-

[12] Anonymous. 28 mars 2003 - Loi modifiant la loi du 21 décembre 1998 relative aux normes de produits ayant pour but la promotion de modes de production et de consommation durables et la protection de l'environnement et de la santé. Belgian Official Journal: M.B. 29.04.2003, 2003.

[13] PRIBEL is described in a Belgian royal decree: *Anonymous. 22 fÉvrier 2005 - arrêté royal relatif au premier programme de réduction des pesticides à usage agricole et des biocides. Belgian Official Journal: M.B. 11.03.2005, 2005.*

[14] PRIBEL is also described in Vergucht et al. (2007).

[15] PRIBEL is a subset of the POCER-2 model, the latter being derived from the POCER-1 model. POCER-1 is described in Vercruysse and Steurbaut (2002) and POCER-2 in Maraite et al. (2004).

ciated with the toxicity of various molecules on different compartments (consumers, applicators workers on the fields, birds, bees, water organisms, earthworms, and groundwater) related to pesticide use. As it is the case for the models used by FOCUS, PRIBEL is typically a scientific instrument designed for specific purposes but put in relation with a political objective: to assess the objective of 25% risk reduction by 2010, mentioned in the Belgian agricultural pesticide reduction plan.

An aggregation method for the risks calculated by PRIBEL was further developed by Pissard et al. (2005), and by Piñeros Garcet et al. (2006b) in the framework of the Walloon environment Code (27/05/2004 decree) requiring environmental reports, open to citizens participation in Piñeros Garcet et al. (2006a), and implemented as a test case for risk aggregation in Piñeros Garcet et al. (2007). Aggregation is performed through: (1) the modelling of the statistical distribution of risks of individual risk events; and (2) the subsequent calculation of the p90[16] per agricultural region. An example of the outputs of the model is given in fig. 1.2.

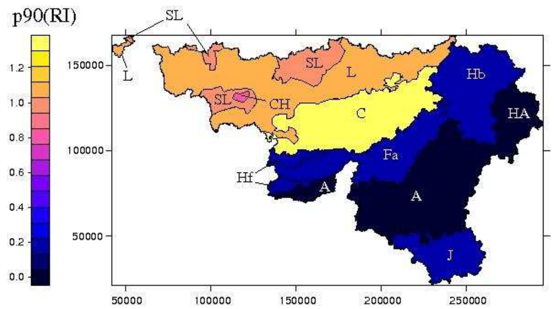


Figure 1.2: *90th percentile of acute risk for bees, per agricultural area in the Walloon region (Piñeros Garcet et al. (2006b), Ansay et al. (2007)).*
Agricultural regions: SL: Sablo-limoneuse, L: Limoneuse, J: Jurassique, Hb: Herbagère-Liège, Hf: Herbagère-Fagne, HA: Haute Ardenne, Fa: Famenne, C: Condroz, CH: Campine hennuyère, A: Ardene.

1.4 The GERU laboratory

The discourse, practices and projects of the members of the GERU laboratory, related to the creation and use of mathematical environmental scenarios and models, and particularly about their representation function, are used in chapter 3 as a contribution to the definition of the

[16]90th weighted percentile.

scientific component of the scenarios quality, through a anthropological structural symbolic systems analysis.

The studied laboratory is the former Agricultural Engineering unit (GERU) of the ''Université catholique de Louvain'', in Belgium as was existing in 2003[17]. According to the unit's web site (as it was in 2003), the unit provided courses for both undergraduate and graduate students in the field of machinery, construction, and water resources management. The main research activities focused on soil and water resources systems, such as: water-soil-plant relationships, erosion, surface and subsurface hydrology, irrigation, and monitoring.

There were around 25 members of the laboratory in 2003. As with many present-day university research laboratories, GERU was trying to maintain a balance between basic and applied research. The latter is sometimes viewed as a necessity from the financial point of view and sometimes as a legitimate university involvement in society. This is interesting because this explains why, according to many GERU members, regulatory scenarios were studied and developed by them. The laboratory had three academic staff members, three technical and administrative staff members, and around 20 researchers aged between 23 and 33 years. Researchers included those in receipt of grants, research assistants (i.e. financed by the university) and researchers on external projects (financed mainly by projects from public authorities such as the European Commission and the Walloon Region). This last category of researchers contains the majority of the researchers involved in applied research. The laboratory has gained substantial experience in the creation of models of the diffusion of pollutants in soil, especially with pesticide and nitrate, often in projects carried out jointly with the public authorities, so making the laboratory an appropriate object for our study of regulatory forecasting simulation scenarios.

1.5 The APECOP research project

The APECOP project is presented here because an anthropo-sociological study of it (chapter 4) is useful for showing that the quality components of a scenario are not necessarily independent. This contributes to the implementation of our quality assessment framework,

[17]Due to a reorganization of the University, research units do no longer exist at UCL since January 2010.

through a proposal in chapter 5 for taking those links into account.

APECOP (Vanclooster et al., 2003) is clearly defined in its program as a contribution to the FOCUS working group activities, and not as an ‘‘outsider’’ critique of the FOCUS scenarios. The aims of the project (APECOP, 1999) consisted in the ‘‘validation’’ of FOCUS-GW models and scenarios, as well as in proposing new and improved methods compatible with European legislation on pesticide, and to contribute to the improvement of the FOCUS-GW models and scenarios. APECOP was a project financed by the European Commission through research funds from DG Research (Fifth Framework Programme). The GERU laboratory was coordinator of the project.

The project had 10 partners in 7 member countries. It includes universities and public research centres. The beneficiaries of the project are, according to its coordinator (Vanclooster, 2002): *‘‘registration authorities at the European [...] and national scale (national ministries of agriculture and environment); developers and producers of pesticide; the agricultural sector and environmental, especially water, sector at the European and national scale’’.* The project is regarded as ‘‘sensitive’’ by the Commission and by industry as it touches on the procedures for regulatory approval, which means that substantial economic interests are involved.

The persons who form the basis of this project are a part of the group that co-signed the FOCUS guidance document on scenarios or are part of other FOCUS working groups. According to the coordinator, other persons were asked to participate to the project for a variety of reasons: in order to obtain European funding a group must contain members from several countries, while other persons facilitate access to databases or to experimental sites, needed to obtain the project objectives. The project, as proposed to DG Research for financing, is highly specific: it shows the weaknesses of the FOCUS-GW approach and proposes to document them and add improvements. It is clear that it is through their previous participation in FOCUS-GW that the project's scientists were able to acquire intimate knowledge of the problems and so to propose the project. It is worth noting that FOCUS scenarios are more easily used with models proposed by FOCUS-GW (the ‘‘FOCUS models’’) and that various persons involved in APECOP designed the FOCUS models as well as the FOCUS-GW scenarios.

CHAPTER 2

A Framework for quality assessment

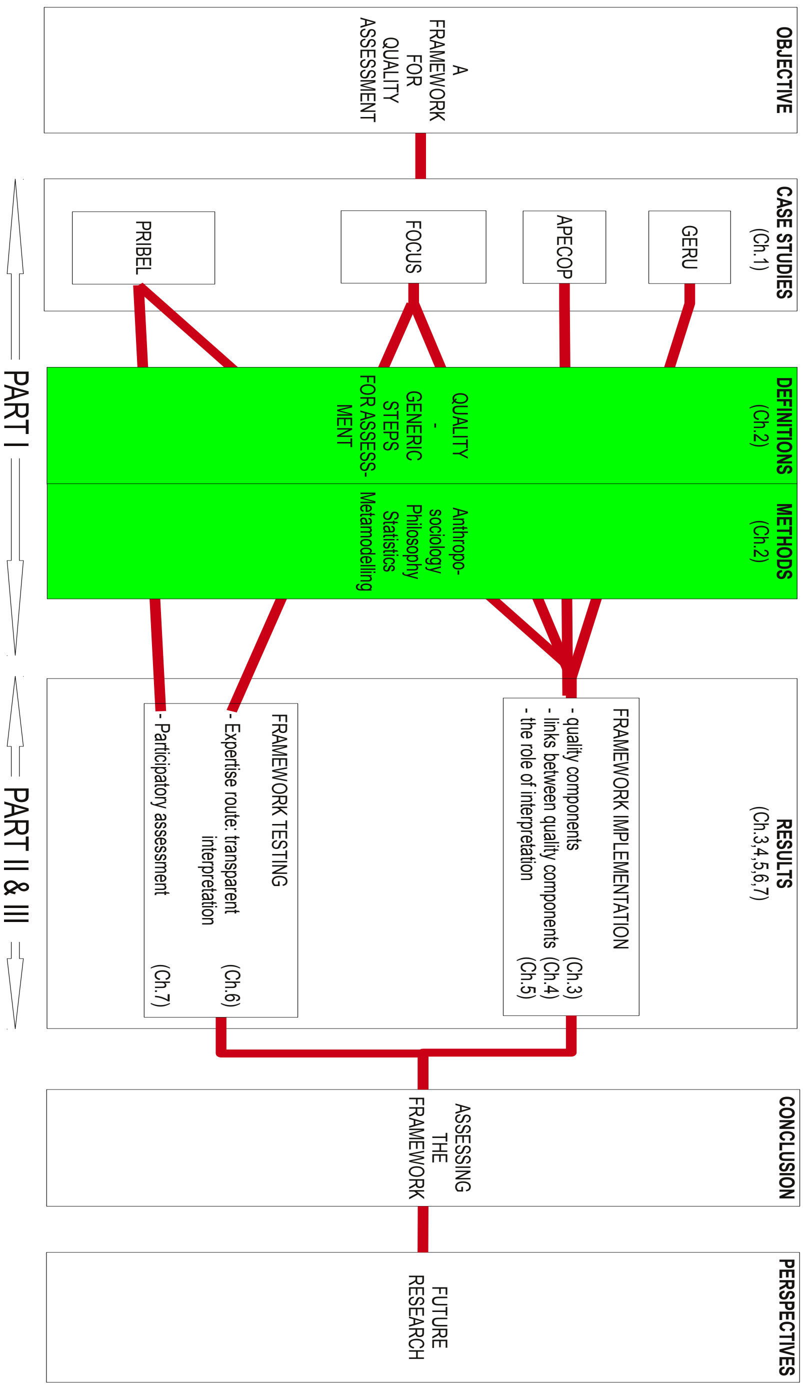


Figure 2.1: *Reading guide of chapter 2.* This chapter defines a generic framework for simulation scenarios quality assessment (i.e. a definition of the concept of quality and generic steps). Then a method is presented that combines inputs from anthropology, sociology, philosophy, juridical analysis, and science for implementing and testing the framework.

Summary

The objective of this chapter is to define a framework for the assessment of regulatory forecasting simulation scenarios, and to propose an approach for its implementation and test. In order to achieve this objective, this chapter:

- Establishes a definition of scenarios quality, in view of their assessment: *scenario quality is the adequacy of the characteristics of a scenario to given use(s) or purpose(s).*
- Defines a *quality component* as the set composed by a scenario characteristic, a scenario use, and an adequacy definition.
- Proposes generic steps for the quality assessment, that includes the definition of quality, of the assessment purposes, of the quality components, of the links between the quality components, the assessment in itself, as well as the assessment of the quality of the assessment, its use, and specifies the role of interpretation and narrations in quality assessment.
- Presents implementation objectives and methods for Part II.
- Proposes test objectives and methods for Part III.
- Indicates how the case studies previously presented in chapter 1 are used for implementing and testing the framework.

For that purpose, this chapter starts with a literature review about the word “validation” in philosophy and modelling (that shows that despite is wide use, is not appropriate as a concept for scenarios assessment). Then the concept of quality is examined from a sociological perspective (quality is seen as the result of a social process rather than an intrinsic quality of the scenario), and a literature review about quality assessment practices, data quality and quality in modelling is presented (that highlights the importance to define quality relatively to a given use). Then the proposed steps for a generic quality assessment framework are presented, along with the specific contribution of the following chapters to the implementation and test of each step.

2.1 Quality assessment review

2.1.1 "Validation" concepts in modelling and in philosophy of science

Models and modelling

In the modelling literature, the word "validated" is often used, and one might think it is a synonym for "good quality". Oreskes et al. (1994) discussed verification, validation and confirmation for deterministic models with examples in hydrology and geochemistry. In verification, the assumption is (according to Oreskes et al. (1994)) that models can be assimilated to logical propositions, and therefore their truth value can be known from the confrontation to observational data and a logical verification. This locates verification in the tradition of the logical positivism (e.g. Carnap (2002-1928)). Oreskes states that verification is not possible for scientific deterministic models, mainly because the logical fallacy called "affirming the consequent"[1]. If a model fails to reproduce observed data, then we know that the model is faulty in some way, but the reverse is never the case (e.g. errors in hypothesis can cancel each other out).

Regarding validation, Oreskes sees it as to be restricted to the fact that a model can or not contain errors of logic. So for Oreskes et al. (1994), the use of "validation" as meaning "consistent with observational data" is erroneous (they prefer to describe it as "consistency within a system or between systems"). The use of validation as meaning "accurate representation of reality" is an impossibility she says. We refer for this to van Fraassen's views about what a model can represent: according to him, empirical adequacy does no involve belief in the truth of a theory or model (see van Fraassen (2008), and section 5.1 and appendix H).

Monnoyer (2003) indicates that other views exist: that scientific progress is made through ruptures (Kuhn, 1996-1962) indicates that corroboration can not be based on the "normal" current science only, and some says that refutability is not essential to theories because they involve interpretation of facts (Feyerabend, 1993).

Oreskes et al. (1994) prefer the term *confirmation*, based on a

[1] Makins (1991): the fallacy of inferring the antecedent of a conditional sentence, given the truth of the conditional and its consequent.

hypothetico-deductive and probabilistic[2] view close to Popper , where models are seen as scientific hypothesis: *''confirming observations do not demonstrate the veracity of a model or hypothesis, they only support its probability''*. Moreover *''there will always be the possibility that more than one theory can explain the available observations''*. This is also an argument against verification. Oreskes et al. (1994) makes the observation that validation or verification vocabulary implies a binary response, while confirmation is a matter of degree: some data agree with predictions, other not. They suggest to use the term ''relative performance'' (relative to data, other models, expectations, etc) because this invites to a contextual definition.

Hendley (2003) advocates for seeing models as frameworks to help interpret or extend data, because assumptions are inherent to modelling.

Besides semantic and philosophical considerations, models ''validation'' is now a current practice in modelling science and in the vast majority of cases, the term ''validation'' is used. We argue that simply scientists and philosophers are not using ''validation'' with the same meaning and in consequence it is a matter of defining the terms we use (this is also supported by Rykiel (1996)). I shall replace the use of the ''validation'' term by ''quality assessment'' which is probably a good concept to describe what scientists are doing under the concept of ''validation'', as we shall demonstrate here below.

Model ''validation'' is being criticized for various scientific reasons (e.g. scale extrapolations involved once validated models are used, lack of measured data for extreme values, validation method not addressing the regulatory purposes, ''weight of evidence approaches'' could be more suitable, etc, see e.g. Hendley (2003)). On the other hand, Hendley (2003) emphasises that monitoring data are ''typically a snapshot of a particular set of circumstances'', and in consequence, both modelling and monitoring are needed. Rykiel (1996) made a chronological review of models validation in science, from which I give here some elements. The validation subject in ecological modelling appeared in discussions in the 1960s (see e.g. Levins (1966)).

[2]Popper uses the term ''logical probability'' (which is not equivalent to statistical inferences) and sees it as proportional to the possibility of falsification. Theories are for him either corroborated hypothesis or false.

Rykiel states:

> *Validation means that a model is acceptable for its intended use because it meets specified performance requirements. [...] Validation therefore is the process of showing that a model (1) accords with the facts (data) as we know them at the time; and/or, (2) with our judgment of what is true or accepted as true in ecology; and/or, (3) is justifiable and appropriate for our purposes.* (Rykiel, 1996).

This is based on a distinction between two general uses of models: to understand (scientific hypothesis testing) and to predict (engineering testing / operational validation). ‘‘Operational validity’’ is limited to given conditions and assumptions. ‘‘Qualification’’ is aimed at discovering the domain over which a ‘‘validated’’ model may properly be used. Levins (1966) states that the purpose of model testing is to increase credibility (and not prove truth).

Russell et al. (1994) considers that selecting appropriate models for regulatory purposes includes model applicability, ‘‘validation’’, capability, ease of use and documentation, and highlight problems commonly encountered in modelling including limited accuracy, lack of defined objectives and standard modelling practices, and misuse of models and results.

Scenarios

Regarding scenarios, in contrast with the large number of publications available on models ‘‘validation’’, few studies dealt explicitly with ‘‘scenarios’ validation’’. Hendley advocates that:

> *A fundamental point [...] is that almost all scenarios are ‘‘valid’’ as long as their underlying assumptions and implications are understood. While they may not closely reflect reality, they may be perfectly suitable for a specific regulatory purpose. [...] There is nothing ‘‘wrong’’ with [...] assumptions [...] as long as their impact on the assessment [is] documented.* (Hendley, 2003).

Hendley (2003) gives an example of a comparison between scenario’s assumptions and ‘‘reality’’, and emphasises that inputs/factors selection for being included in the scenario can have a high influence on model outputs.

However, if scenarios are considered has ‘‘input data’’, a more important corpus exists about ‘‘data validation’’ that is related to scenarios. Caswell (1988) states that neither models or data can be ‘‘validated’’. Rykiel states:

> *Validation of data certifies that the data meet a specified standard (quality assurance/quality control). We cannot assume that data accurately represent the real system and therefore constitute the best test of the model. The validity not only of the data but also of its interpretation must also be demonstrated.* (Rykiel, 1996).

In this citation, Rykiel (1996) is clearly limiting quality assessment to the measurement case. This is reemphasised in the same publication:

> *[...] the quality or property of being in accord with fact or reality [...] correspond to: [...] consistent with available data* [sic] (Rykiel, 1996)

van Fraassen (2008) also states that we do not have an access to phenomena, we rather have access to appearances, data models, and in that sense data (or data incorporated into simulation scenarios) can not be validated through some direct access to the phenomena. Oreskes et al. (1994) states that *‘‘what we call data are inference-laden signifiers of natural phenomena to which we have incomplete access’’*. This is applicable to scenarios build from measurements. Oreskes et al. (1994) gives an example from the atmospheric modelling: we do not have a direct measurement to the average temperature of the Earth. This can be translated to our pesticide leaching scenarios, by noting that e.g. the mean organic matter content attributed to a $5x5\ km^2$ pixel in a raster soil map can not be directly measured, it is only accessible through premises, hypothesis and inference.

Assimilation of scenarios to ‘‘measured input data’’ is too short as we will see in our section 5.1, because the existence of the so-called ‘‘virtual’’ but useful scenarios, like the FOCUS scenarios that are claimed not to represent any specific field but rather to represent when used with models. On the contrary, PRIBEL scenarios are claimed to represent by virtue of being based on surveys. This lead us in chapter 3 to distinguish between functional and measurement representation.

A review of literature about models "validation" can give us indications about what scientists have in mind when "validating". However, scenarios development is a central task in modelling, and many publications propose some criterion or method, generally ad-hoc, for scenarios development and describes in some way what "a good scenario" is.

We took a large corpus of publications (87) about modelling and performed a text analysis to selected context terms related to "validation". Some of the terms more or less related to "validation" we have found about modelling are[3]

> *verification, evaluation, corroboration, confirmation, trustworthiness, veracity, assessment, (cost of) errors, error (indicator/analysis/calculation), (acceptable/desired) accuracy, comparison, precision, bias, efficiency, reliability, testability/testing, plausible, acceptable, refutable, (empirical) adequacy, performance (criteria/measures), relative performance, proof, consistent with experience, convincing, degree of correspondence, (level of/building) confidence, user requirements, goodness, transparency, clear and agreed expositions, strengths, weaknesses, assumptions, sensitivities, clearly understood, regulatory validity of outputs, qualification, generalisation / generalisable, internal validity, blind validation, model expressing easy understandable assertions, easily interpretable model principles.*

Some of those terms have ambiguous meanings. As mentioned already verification and validation usage in philosophy is different from common use in modelling science. Another example is "scenario's evaluation" which is sometimes used with the meaning "to run a model with a scenario".

[3]We do not intend to present here a commented review of specific procedures for models "validation", since other made comprehensive surveys, see e.g. Kleijnen (1995), Kleijnen and Sargent (2000), Hassan (2004), Jones and Mangels (2002), Hendley (2003), as well as Law and Kelton (1991), Mayer and Butler (1993), and Power (1993).

Regarding scenarios, our literature review and text analysis about their development gives the following related concepts:

> *scenario's acceptance, discrimination between calibrated scenarios by ''best compromise between goodness of fit and parsimony'', scenarios: protective / conservative / standard / reasonable / (un)likely / possible /realistic / admissible / represent(ing/ative) / accessible / well defined / adequate, nature of the sources, complete/proper sets of data, evaluation of the quantity and quality of the input information, design of simulation experiments (number, sampling of scenarios), represent / representativeness, intended use, derived from accurate data, selection of inputs and modeler's influence, correlated inputs , relevance for assessment end-point, input parameters: model sensitivity to / uncertainties in, mismatches between model and input data.*

Such a diversity in the assessment of the scenarios is to be explained by a diversity of scenarios uses. Moreover, our philosophical review points to an assessment in terms of ''relative performance''. It becomes now clear that what we are searching for is a way of assessing the *quality* of simulation scenarios for a given use. This locates us in the field of quality assessment.

2.1.2 Quality according to the sociology of science

Sociology of science discusses extensively the relations between society and science, and ''validity'' is at the center-stage of the relations between knowledge and society. A branch of sociology of science called sociology of scientific knowledge, sees itself as the sociology of the scientific contents, theories, objects, rather than the sociology of the researchers. The translation and actor-networks theory (Callon (1986), Latour et al. (1992)) is a particular version of the sociology of knowledge. If we follow Latour and Callon's concepts, the quality of a scenario is not an intrinsic quality of it, it is rather the *result* of a controversy, of a social translation. We will see in chapter 4 that relations exists between a diversity of actors and the activity of scenarios' assessment. Our chapter 5 arrives at an integrated view between philosophical considerations, scientific methodological proposals and sociology of knowledge.

If social factors are involved in the way scenarios are assessed, it is better to render them as transparent as possible, rather than trying to avoid considering them or try to hide them. Our proposal for scenarios' quality assessment integrates those aspects.

2.1.3 Quality assessment, data quality and modelling quality

Reeves and Bednar (1994) presents a large review of definitions of "quality". They found a diversity of definitions along the following lines: conformance to specifications, conformance to requirements, fitness for use, loss avoidance, meeting/exceeding customers' expectations, excellence, value, and concluded that no universal definition of quality exists, rather various definitions are appropriate under different circumstances.

ISO defines quality as "the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs" (norm: ISO 8402: 1986, 3.1, "Vocabulary - quality").

EPA (2002) discusses model quality assurance procedures, and gives the following definitions:

- quality: the totality of features and characteristics of a product or service that bear on its ability to meet the stated or implied needs and expectations of the user;
- assessment: the evaluation process used to *measure* the performance or effectiveness of a system and its elements;
- performance evaluation: a type of audit in which the quantitative data generated in a measurement system are obtained independently and compared with routinely obtained data to evaluate the proficiency of an analyst or laboratory; and
- data quality assessment: a statistical and scientific evaluation of the data set to determine the validity and performance of the data collection design and statistical test, and to determine the adequacy of the data set for its intended use.

Pipino et al. (2002) discusses data quality assessment, and highlights that data quality is a multidimensional concept. They distinguish subjective data quality assessments "[that] reflect the needs and experiences of stakeholders" from objective assessments "[which] can be task-independent or task-dependent. Task-independent metrics reflect states of the data without the contextual knowledge of the application, and can be applied to any data set, regardless of the tasks at hand. Task dependent metrics, which include the organization's business rules, company and government regulations, and constraints provided by the database administrator, are developed in specific application contexts". Pipino et al. (2002) lists a variety of "quality dimensions" that can be assessed[4]. They indicate that: "Often, the most difficult task is precisely defining a dimension, or the aspect of a dimension that relates to the company's specific application. Formulating the metric is straightforward once this task is complete."

ISO, as well as EPA (2002) and Pipino et al. (2002) relates thus quality to needs, uses, tasks, and applications.

EPA (2002) is devoted to modelling in general, but gives indications about "inputs" and "parameters" which in fact are scenarios. They prescribe parametrization documentation and peer review, description of the intended use, performance or acceptance criteria, data quality indicators, relevance, appropriateness, and various other guidelines listed in appendix B. They list the things that should be done, but the particular way of defining acceptance/performance criteria are not detailed. This depends again on the use of the scenarios and model.

In the IPPC[5], identification of "good" scenarios (who meet user needs and requirements) for climate change modelling was discussed, and the characteristics were given:

> *Transparency in numbers and assumptions. Comparability among scenario results. Internal & temporally consistency within scenario families. Scenarios that have credibility (i.e. cover a range of plausible range of futures) and legitimacy (i.e. involve key stakeholders in their development). Policy relevant. Reproducible. High availability and dissemination. Easily communicated to the user-groups.*[sic] (IPCC Working Group III Technical Support Unit, 2005).

[4] Accessibility, appropriate amount, believability, completeness, concise representation, consistent representation, ease of manipulation, free-of-error, interpretability, objectivity, relevancy, reputation, security, timeliness, understandability, values-added.

[5] The IPPC is the Intergovernmental Panel on Climate Change.

Representativeness is of particular interest in our work. The EPA indicates:

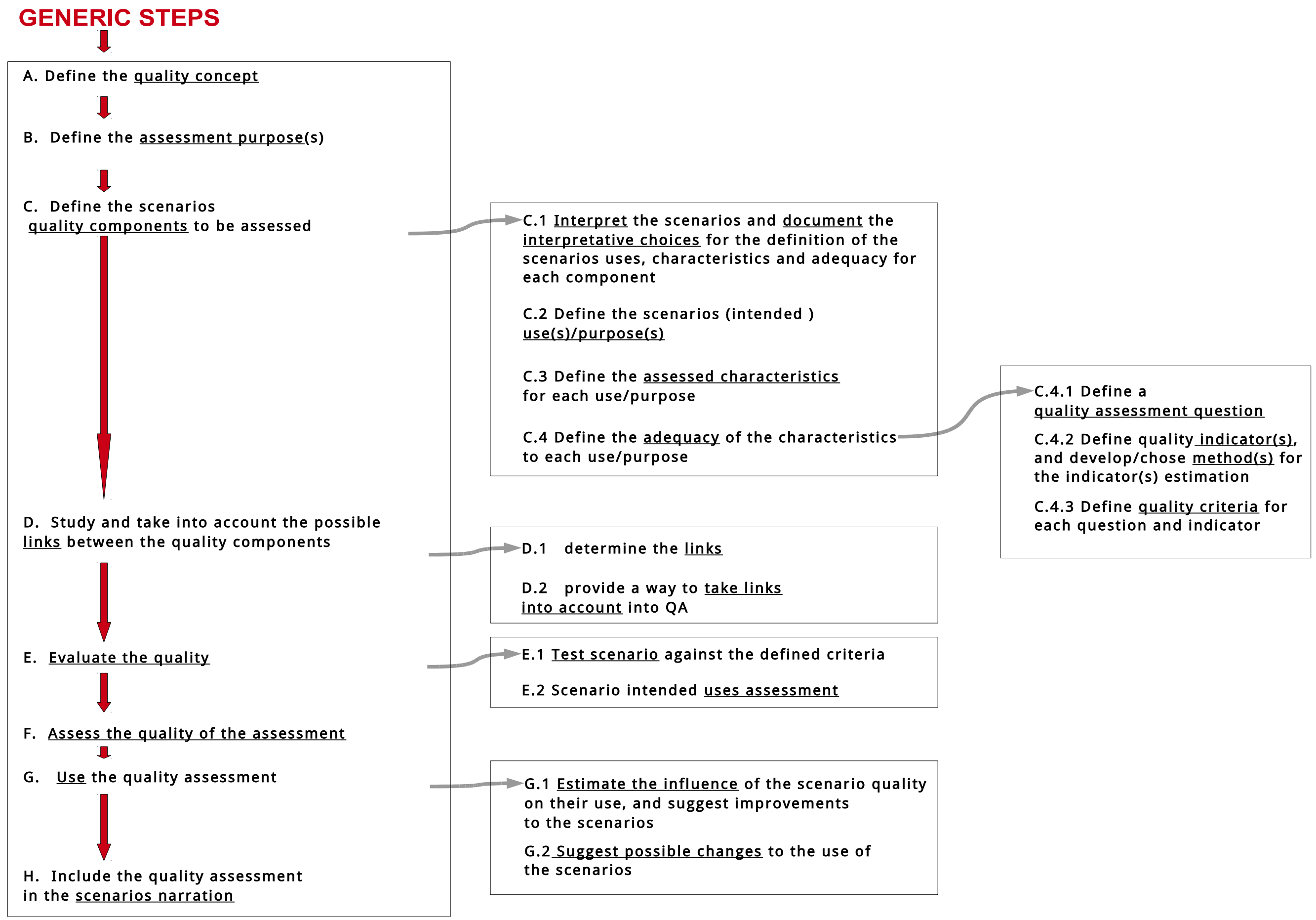
> *Representativeness: Were the data collected from a population sufficiently similar to the population of interest and the model-specified population boundaries? Were the sampling and analytical methods used to generate the collected data acceptable to this project? How will potentially confounding effects in the data (e.g., season, time of day, location, and scale incompatibilities) be addressed so that these effects do not unduly impact the model output?* (EPA, 2002).

This definition does not consider cases of ''virtual'' scenarios that however claims to be representative, as FOCUS scenarios. It relies on *similarity*, and we will see in our section 5.1 that representation does not always rely on similarity: it relies on it *if* someone uses similarity to represent. Moreover, this definition is hardly applicable to scenarios like the IPCC emission scenarios for climate change modelling because of its focus on data collected or measured. Indeed IPPC claims:

> *[IPCC emission] Scenarios are alternative images of how the future might unfold and are an appropriate tool with which to analyse how driving forces may influence future emission outcomes and to assess the associated uncertainties. [...] The possibility that any single emissions' path will occur as described in scenarios is highly uncertain.[...] A set of scenarios was developed to* <u>*represent*</u> *the range of driving forces and emissions in the scenario literature so as to reflect current understanding and knowledge about underlying uncertainties. [...] Any scenario necessarily includes subjective elements and is open to various interpretations. [...] Probabilities or likelihood are not assigned to individual SRES scenarios. None of the SRES scenarios represents an estimate of a central tendency for all driving forces or emissions, such as the mean or median, and none should be interpreted as such.* (IPCC Working Group III, 2000).

We will develop a representativeness concept that can take into account such cases. Moreover, if quality is defined relatively to a use, then regulatory scenarios quality assessment can be also performed for their representation function, political, economical and juridical functions, their internal logical consistency, their accordance to scientific theories, their social acceptability, their transparency, etc. All those components of the quality can in fact be assessed, as it will be shown in chapter 3.

Figure 2.2: Generic steps proposed for the assessment of regulatory forecasting simulation scenarios. The steps forms a sequence from A to H, and sub-steps are defined for some of them. A detailed explanation of each step is provided in the text section 2.2



2.2 Quality definition and generic steps for scenarios quality assessment

We argue that it is very important to highlight that quality is dependent on specific tasks or uses, and disagree about the usefulness of defining task-independent metrics, such as those listed by Pipino et al. (2002). In fact those "task-independent" metrics, while useful, are neverthless metrics used for implicit tasks: e.g. the consistency is only relevant when the data are used to be consistent. The ISO norm 8402:1986,3.1 (see section 2.1.3) use the terms "ability to satisfy stated or implied needs", which suggests the need for a *user*. This is in fact also the intepretation made by EPA (2002): they added "ability to meet the stated or implied needs and expectations of the *user*". In fact we interpret the fact that Reeves and Bednar (1994) did not found a universally valid definition to the fact that quality is always relative to a *use*. We will come back to this in chapter 3 when defining quality components, and in 5 where a detailed philosophical discussion will provide reasons for seeing the use as central in the definition of one of the main scientific components of the quality: the representativeness.

The first step of a quality assessment should then be an explicit definition of quality (step A). Our definition is that "quality is the adequacy of the characteristics of something to given use(s) or purpose(s)". This definition is valid both for the quality of a scenario, and for the quality of the assessment itself. For scenarios:

> *Scenario quality is the adequacy of the characteristics of a scenario to given use(s) or purpose(s).*

Taking that into account, the second step should be, for scenarios quality assessment: the definition of the use or purpose of the assessment, and the definition of the use or purpose of the scenarios. The purpose or use of the assessment (step B) may condition the scenarios uses that are assessed, and also the quality of the assessment itself.

The set constituted by {use/purpose,characteristic,adequacy definition} is called here a *quality component*. The definition of a quality component (step C) requires an interpretation (step C.1) of what is the expected use of the scenarios (step C.2). Such an interpretation should then be documented. The uses or purposes of the scenarios themselves are a guide to the selection of the characteristics to be assessed (step C.3), and for the adequacy definition (step C.4).

We make the proposal that the definition of the adequacy is facilitated by the definition of an *assessment question* (step C.4.1) that summarizes the relations that are being established between the use and the characteristic by the adequacy concept. This takes the form of a sentence like ‘‘How the characteristic X is, compared to Y, when the scenario is used for Z’’. An example is: ‘‘How much is a scenario parameter conform to the prescriptions of the relevant law, when use for regulatory purposes?’’.

Quality assessment questions should be translated into quality indicators (step C.4.2) related to the chosen scenarios characteristic, and quality criteria (step C.4.3) to decide if the value (quantitative or qualitative) of the indicator is above, equal or below a chosen threshold. That means also that methods for evaluating such an indicator must be selected or developed.

The next step (step D) is to investigate if there are links between quality dimensions (step D.1), and provide means to take that into account (step D.2). As an example, you may have various equally scientifically justifiable parameter choices, but the quality of the choice is determined by the fact that when the scenarios are used as regulatory decision-support systems, the relevant law may prescribe some parameters rather than others.

The next step (step E) proposed is to evaluate the scenario for a given use using the indicator and criteria (step E.1), and/or the assessment of the use of the scenario (step E.2) in a given context (social, regulatory, economical, etc). This distinction between steps E.1 and E.2 is important, because we will see in chapter 7 that the assessment that citizens makes about scenarios can encompass both aspects.

When doing a quality assessment, it is proposed to have a reflexive step, where the quality of the assessment itself is evaluated (step F). This is related, amongst others, to the purpose of the assessment.

A multitude of uses of the results of a quality assessment (step G) exists, but for simulation scenarios, we at least see two that are of prime importance: first to estimate the influence of the quality of the scenarios on the scenarios use (step G.1), and from there suggest improvements, and secondly, possibly suggest changes to the use of the scenarios (step G.2).

Finally, since quality assessment involves interpretation of the use, of what quality is, about the relevant characteristics for a given use, etc, we suggest that such interpretations, as well as the results of the assessment should be incorporated into the scenarios narrations (step H), for transparency and repeatability but also for helping new future interpretations, for democratic scrutiny, for mapping the various possible interpreations, for peer review, and for helping in defining the possible impact of the interpretation on regulatory decisions.

These steps are summarized in figure 2.2. What is the epistemological status of this generic framework? Following in this Bergson (1969), it is not a discovery of a pre-existing or existing fact, it is an *invention*: the invention of a method, and as such must be judged by its effects for specific tasks. Those generic steps should be implemented with the necessary tools for the specificity of regulatory forecasting simulation scenarios, and then tested for the benefits or drawbacks of their effects.

2.3 Approach for implementing and testing the framework

Having presented our case studies in chapter 1, and defined in this chapter a generic framework to perform forecasting regulatory simulation scenarios quality assessment, our objective is now to present an approach that implements and tests the framework and its steps, using our cases studies.

2.3.1 Overview of the approach

Step A of the framework (defining the quality concept), is already discussed in our previous section. Step B, the definition of the purpose of the assessment of scenarios, is discussed in our introduction, as well as in the preceding section, and will be also illustrated by our test cases: (1) the purpose of the assessment made by APECOP is discussed in chapter 4, along with the consequences of the chosen purpose for its results; (2) the reasons for doing an assessment of FOCUS and PRIBEL scenarios are presented respectively in chapter 6 and 7.

The implementation of steps C and D of the generic framework (definition of the quality components and their links) is presented in Part II (chapters 3,4, and 5) . Steps E,F,G and H (perform the assessment, assess the quality of the assessment, use the quality assessment and inclusion of the assessment into a narration) are part of the application of the framework with two case studies that is presented in Part III (chapters 6 and 7). The application of the framework to two test cases in Part III is a test of the framework, and gives test results for steps A,B,C, and D also. The term *test* is understood as an account of the applicability, limits and interests of the proposed framework, in view of the chosen goals for that framework: the ability of the framework to evaluate the quality of a scenario in its multiple aspects in a transparent way, and give insights about (1) the influence of the scenarios quality on the use of the scenarios, (2) potential improvements to the scenarios, and (3) potential improvements or changes to the scenarios use.

Table 2.1 gives an overview about the methods and case studies that each chapter uses to contribute to the generic steps implementation and test. A detailed account is given in next section. A reading guide is provided at the beginning of each of the coming chapters, in order to recall how they contribute to implement and test the generic quality assessment framework. Those reading guides are based on figure 2.2, and highlights in it the specific contribution of each chapter.

Table 2.1: Overview, chapter per chapter, of the methods, cases studies and contributions to the implementation and testing of the generic quality assessment framework steps.

Part/Chapter		Case study	Method	Step to which it contributes
Part I	A Generic Framework for Quality Assessment: Definition, Material, Methods	FOCUS, PRIBEL, GERU, APECOP	Study of cases, literature review, normative approach	A, B
Part II	Framework Implementation Using Conceptual developments, Normative developments, and Empirical Observations			
chapter 3	Defining the quality components to be assessed	GERU, FOCUS	Anthropology, mathematics, juridical analysis, political philosophy	C
chapter 4	Determining the links between quality components	APECOP	Anthropology	D.1
chapter 5	The role of interpretation in quality assessment	APECOP, FOCUS	Philosophy, sociology, quality assessment practices	D.2
Part III	Framework testing: two case studies			
chapter 6	Making interpretation choices transparent through the scientific expertise route	FOCUS	Interpretation, metamodelling, statistics	C.1, C.4, E.1, F, G.1, H
chapter 7	Opening interpretation choices through the participatory route	PRIBEL	Sociology, modelling, participatory practices	all except C.4, D.1, E.1

2.3.2 Specific Contribution of each chapter

Part II determines, for regulatory forecasting simulation scenarios, the relevant uses and defines the quality components, determines and takes into account the link between quality components. Methods and concepts from quality assessment practices, interpretation, anthropology, sociology, philosophy of science, political philosophy, juridical analysis, and mathematics are used.

Part II is composed by three chapters:

- Chapter 3 implements step C of the framework through: (1) the determination of relevant uses to be assessed for regulatory scenarios, and (2) the definition of the corresponding components (i.e. the set constituted by the use, the characteristic, and the adequacy definition), including quality assessment questions, quality indicators, and quality criteria.The defined quality components are illustrated with the FOCUS and PRIBEL cases. A particular attention is paid to the representativeness aspect of the scientific quality component (representativeness being the quality of the representation), starting from the GERU case and using mathematical approaches. The political component of the quality is also explored and illustrated using the FOCUS case.
- Chapter 4 contributes to step D.1 by investigating the links between the identified components through anthropological observations of the APECOP case. The existence and form of political choices into the scenarios is also explored.
- Chapter 5 implements step D.2 by providing concepts and methods for taking into account the links between quality components in quality assessment. That links are studied in detail and illustrated with the APECOP and FOCUS cases.

A proposal is made for considering *interpretation* as the link between components. The role of interpretation is thus considered in three ways in part II: in chapter 4, from the empirical point of view through anthropological observations, and in chapter 5, from a conceptual point of view through a philosophical approach and from a normative point of view for proposing a role for interpretation in quality assessment.

Part III presents two tests performed using the generic framework, the case studies, and implementation tools obtained in Part II. Methods from interpretation, modelling, metamodeling, statistics, sociology and participatory practices are used. Part III is composed of two chapters:

- Chapter 6 is a test of steps C.1,C.4,E.1,F,G.1,H through an assessment of the FOCUS scenarios for the representativeness aspect of the scientific quality component. The FOCUS scenarios representativeness assessment starts with a transparent documented account of the political interpretation choices made when the scenarios are interpreted by an expert, and then considers such choices as a boundary condition for the representativeness estimation from a scientific point of view. Emphasis is made on the test of steps C.1 and C.4: define the adequacy of the characteristics to each use/purpose, including quality assessment questions, quality indicators, and quality criteria definitions, and the documentation of the interpretative choices for the uses, characteristics and adequacy for each component. The interest of a documentation for making transparent political choices in the scenarios interpretation is discussed.
- In chapter 7, PRIBEL is taken as a test of a step further towards the full use of the framework. PRIBEL scenarios are used for testing a method of participatory assessment, that opens the interpretation political choices to discussion and modification, and tends towards a co-construction of the scenarios with lay-people, while keeping the specificity of the scientific work. This is a test of practically the whole framework including the quality definition, except steps C.4, D.1, E.1, through the development and test of methods for assessing the political choices in the scenarios, through the opening to discussion and documentation of the interpretative choices made by scientists when using scenarios in regulatory contexts, through the documentation of what lay-people think about such a participatory assessment, and through the use of the quality assessment for suggesting improvements to the scenarios and to the use of the scenarios.

2.4 Chapter conclusion

The objective of the assessment of scenarios is to evaluate the quality of a scenario in its multiple uses in a transparent way. A complete quality assessment is a multicriteria approach, and allows degrees of quality rather than an accept / reject approach. We conceive the quality assessment of scenarios as the evaluation of the adequacy of the characteristics of a scenario to given use(s) or purpose(s). A quality component is defined as a combination of a use, a scenario characteristic and an adequacy definition. A scenarios quality assessment is not possible without knowing the use context; even a scientific assessment presupposes a scientific use of the scenarios. Quality assessment presents a very applied, particular or local aspect, because it depends on the considered *use*. However, generic steps for a scenarios quality assessment were proposed, and should be adapted to each situation.

Although validation is often used in modelling science, using the term validation for scenarios or data is problematic for various reasons: amongst others, we do not have a direct access to the phenomena, but rather measurements results that are inference-laden and such inferences are in fine not verifiable. More over, validation is a concept that do not fit with the existence of a representation by ‘‘virtual’’ scenarios like the FOCUS ones that are claimed not to represent any specific field, but rather represent through their use for simulations with a model. Validation is a narrow concept when the various aspects of the scenarios are considered: if quality is defined relatively to a use, then regulatory scenarios quality assessment can be also performed for their representation function, political, juridical functions, their internal logical consistency, their accordance to scientific theories, their social acceptability, their transparency, etc. All those aspects can in fact be assessed, and none of the reviewed references proposes methods for the assessment of those components of the quality.

Methods are suggested to define relevant quality components for regulatory scenarios, investigate their links and take those links into account. The proposed steps as such are, epistemologically speaking, the proposal of a method, and not a discovery of a preexisting fact: they must be tested and evaluated for their effects. A test is here an account of the applicability, limits and interests of the proposed framework, in view of three chosen goals: the ability of the framework to give insights about the influence of the scenarios quality on the use of the scenarios, its potential to generate improvements to the scenarios, and its potential to generate suggestions for improving or change the scenarios use.

The methods suggested in this chapter are implemented in chapters 3, 4, 5, and tested in chapters 6 and 7.

Part II

Framework Implementation Using Conceptual developments, Normative developments, and Empirical Observations

Summary of Part II

Part I presented our case studies, the steps of a quality assessment framework and the method to use the cases for implementing and testing the framework. In part II, contrasting characteristics between our GERU, FOCUS, PRIBEL and APECOP study cases make them useful for the implementation of the framework when accompanied by conceptual developments:

- Chapter 3 implements step C of the framework (*define the scenarios quality components to be assessed*) through: (1) the determination of relevant uses to be assessed for regulatory scenarios, (2) defining the corresponding components (i.e. the set constituted by the use, the characteristic, and the adequacy definition), (3) the definition of quality assessment questions, quality indicators, and quality criteria for each component. This is done by the examination of the PRIBEL and FOCUS as examples of regulatory scenarios, as well as by an anthropological approach to the symbolic system of the GERU members.
- Chapter 4 contributes to step D.1 (*determine the links between the quality components*) through an anthropological observation of the APECOP project. Interpretation is playing a major role in that link.
- Chapter 5 implements step D.2 by providing concepts and methods for taking into account the links between quality components into the quality assessment. Concepts from hermeneutics, philosophy, and sociology are used, and interpretation acquires a central role in the framework by providing the link between the quality components. The linkage of political and scientific components for the case of the representativeness is developed in detail (representativeness being the quality of the representation). Elements for an assessment of the interpretations are given. Finally, chapter 5 identifies two complementary routes for scenarios quality assessment to be investigated in chapter 6 and 7: the transparency of interpretation in the expertise route, and the opening of the interpretation in the participatory route.

CHAPTER 3

Defining the quality components to be assessed

Figure 3.1: Reading guide of chapter 3. This chapter makes use of the GERU, APECOP and FOCUS case studies in order to contribute to implement the generic steps, through the definition of the relevant quality components (i.e. the scenarios uses, characteristics, and adequacy criteria). Next chapter (chapter 4) will study the links between those components.

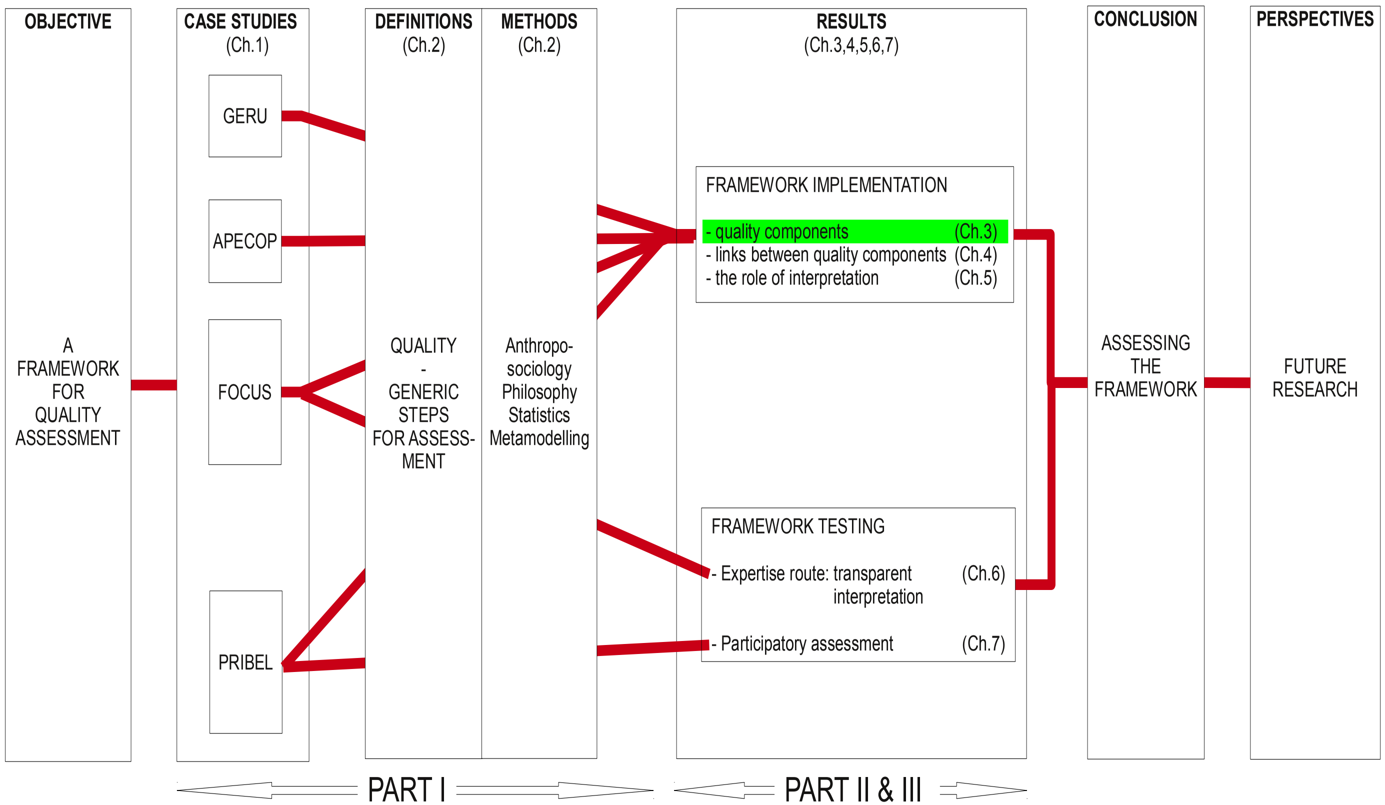
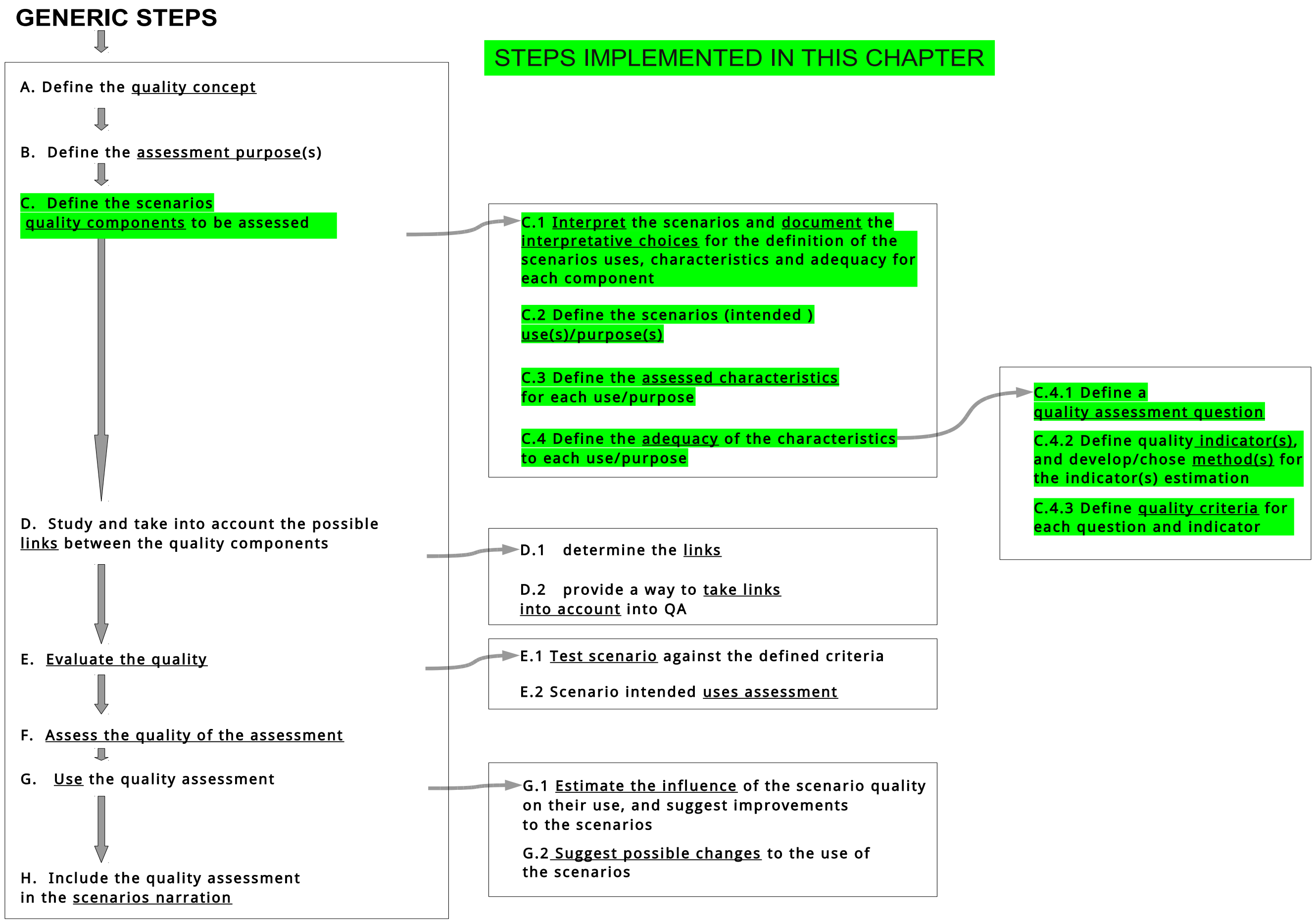


Figure 3.2: Reading guide: contribution of chapter 3 to the implementation of the quality assessment framework generic steps. This chapter implements step C and its sub-steps, in order to define the relevant quality components (i.e. the scenarios uses, characteristics, and adequacy criteria).



Summary

The objective of chapter 3 is to define quality components that are relevant for the use of forecasting simulation scenarios in a regulatory context. A quality component is the set composed by a use, a characteristic and an adequacy definition. This corresponds to the implementation of step C of our generic quality assessment framework (*define the quality components to be assessed*).

Chapter 3 starts with an anthropological and structural analysis of the symbolics system in which the GERU researchers conceives their scenarios. In that system, scenarios are mainly seen as having a use for representation, defined in terms of distance and simplification with respect to the reality.

A critical examination of that system, with philosophical developments, allows us to define the representativeness which is one of the scientific component of the scenarios quality. The theoretical consistency is also defined as a component of the scientific quality. We also take advantage of the scenarios narrations of the FOCUS and PRIBEL scenarios presented in Part I to identify additional scenarios uses (step C.2) that are relevant for a quality assessment in a regulatory context, besides representation: the political, juridical, and logical use. For each use a quality component is defined, and thus proposals for characteristics to be assessed and adequacy definitions are given (steps C.3 and C.4).

3.1 Introduction

3.1.1 Uses and purposes

Generally speaking, as defined in chapter 2, we conceive quality assessment as the evaluation of the *adequacy to a given use or purpose* (they are adequate when they give the expected results when used for a given purpose). In our framework, we consider three types of assessments: the assessment of the adequacy of a scenario to a given use (step E.1), the adequacy of the use itself according to some defined criteria (step E.2), and the assessment of the assessment (step F, as we are doing for the APECOP and FOCUS case). Use is related to actors, actants, purpose, intention, practise, habits, function, etc. In consequence, quality assessment always presents a very applied or local aspect, because it depends highly on the considered *use*. Moreover, scenarios used in diffuse pollution regulation are generally complex combinations of different soil, climate and land use attributes and may therefore conform to the envisaged use in some aspects, but not in others. The objective of step E.1 of the assessment of scenarios is therefore to evaluate the quality of a scenario in its multiple uses in a transparent way. This may result in suggestions to improve the scenario and to reduce the uncertainty in the modelling outputs and in the result of the decisions taken from there (step G.1). However, the assessment may also be done on the uses themselves (step E.2): that is what citizens tends to do, as it will be shown in chapter 7. In the latter case, improvements to the uses themselves may be suggested (step G.2). Considering the assessment itself, it is also a process that can have various uses or purposes (step B), and this can also lead to various criteria.

For the case of the PRIBEL, and FOCUS scenarios, both are scenarios designed to be used in a regulatory context (step C.2). This regulatory context links them to their juridical and political function, use or purpose. Various interpretations of their political or juridical purpose are possible[1], and so various criteria can be proposed. Both are meant to "represent something". Some of them are representing in the sense that they rely on selective resemblances with the things they are representing (PRIBEL). FOCUS scenarios are representing in the sense that they can be "used instead of something for a given purpose" (see section 5.1.3). Both are scenarios developed in a scientific language and their use supposes that they rely on some sort of internal coherence or logic, and are also consistent with scientific theories.

3.1.2 Defining the scenarios quality components to be assessed (steps C)

We believe that the scenarios qualities presented in table 3.1 can be relevant to a large set of scenarios used for regulatory purposes, and thus will develop criteria and methods for them in the rest of this chapter.

[1]E.g. Hendley (2003) thinks about regulatory models in terms of building confidence with the stakeholders.

Table 3.1: *Multicriteria scenarios' quality assessment.* The claimed, supposed or interpreted use of the scenarios determines the quality component being assessed, and thus the assessment question and the criterion to be used for the assessment.

Claimed or supposed use	Conceptual domain into which the assessed component is defined	Quality component being assessed	Assessment question (examples)	Quality Criterion (examples)
As part of a regulatory process	Juridical	Juridical conformity	How are the scenarios compatible with the relevant laws? E.g. 91/414/EEC Directive	Inclusion in the parameters of the maximum number and highest rates of application for which authorization is sought, as prescribed by the Directive
As part of a regulatory process	Political	Democratic nature	How do scenarios comply with some participation standard?	Involvement of citizens in scenarios construction
Used to represent ("instead of" something)	(use dependent)	Functional representativeness	How are the scenarios functioning, when used "instead of" something, in comparison to a reference function?	The scenarios lead to a 90th percentile of leaching when used with a FOCUS model.
Used to represent ("selective resemblances" with something)	Scientific	Measurement representativeness	How good is the matching between representor (scenarios' parameter values) and measurements made on the represented?	The parameters are in the range encountered in the relevant literature
Their use suppose that they comply with a given logic	Logic	Logical consistency	What are, if any, the internal contradictions in the scenarios' narration and parameters?	Absence of permanent crops in the parametrization, if narration claims that scenarios represent arable land
Used as scientific tools	Scientific	Theoretical consistency	What is the parameters meaning inside the accepted theories?	Presence of preferential flow parameters for cracking clay soils

Chapter 6 and 7 are devoted to the application of the proposed assessment criteria. Due to time constraints, consistency aspects of FOCUS scenarios could not be evaluated in detail in the present work. Most attention is paid to the evaluation of the ‘ functional representativeness’ of FOCUS scenarios, and to the political assessment of the PRIBEL scenarios.

3.2 Defining the scientific quality component

Our objective here is to define the scientific component of the scenarios quality. This is done by:

1. An analysis of how researchers conceives models and scenarios, that shows that representation is a main characteristic of the scenarios;
2. Formalizing the researchers' conceptions in a symbolic data analysis mathematical formalism;
3. Defining representativeness indicators and criteria; in a symbolic data analysis mathematical formalism;
4. Defining the theoretical component of the scenarios quality.

3.2.1 Researchers' conceptions about scenarios and models

Our objective here is to describe the way in which scientists from the GERU laboratory identify their relation to reality through their practice with models and scenarios. This is done through the draw up of tables of symbolic systems with the help of a structural analysis of content (‘‘Analyse structurale de contenus et modèles culturels’’ in Albarello et al. (1995)).

In order to obtain the necessary empirical material, a comprehensive observation followed by a comprehensive experimentation are performed (Juan, 1999): I noted phrases and fragments of conversations heard in the GERU laboratory, as well as at laboratory's project meetings when they related either to the relation to reality and/or the relations between science and politics. Then, non-directive interviews were held, although I did specify to the interviewees the overall theme of ‘‘mathematical models’’. The aim of the comprehensive observation stage is to obtain a certain number of key words used by laboratory members themselves, which were then used as the departure point for the comprehensive experimentation conversations.

The construction of symbolic systems based on empirical data is performed by using contextual analysis of the occurrence of selected key words, followed by a simple structural analysis of this context. The aim of the contextual analysis is to extract from the large amount of empirical material the part that is most interesting for the structural analysis.

According to Albarello et al. (1995) and Albarello (2002), the structural analysis of a text is made by (1) making an inventory of interrelated meaning units (2) searching the oppositions between meaning units (contrary, inverse, etc) (3) look for associated pairs of oppositions and group them (4) describe the structure of the groups which is a sketch of the cultural model that is examined.

Observations: about scenarios and reality

The "reality" is continuously present in the discourse about scenarios (and models). Extracted from our observations, here after some example sentences containing words derived from the Latin root "res":

- *realistic values*
- scenarios [. . .] are considered as not very *realistic*
- a comparison with the *reality*
- the purpose of the scenarios is to represent various *realities* or situations
- there are other indicators of the *reality*: intuition, experience (the two are linked), the physical laws
- the *reality* is the ensemble of all the particular cases
- the *unrealistic* organic matter content in the scenarios
- the groundwater level (in meters) is something *real*
- the combination of crop, soil, climate and agronomic conditions should be *realistic*
- at *realistic* depths of various meters
- *real* data versus optimized parameters

The context in which those words are formulated is used here as the basis for the structural analysis in order to reveal the symbolic system. In the following frame, we present some sentences with emphasised context words.

- *data* can be initial *parameters* used as *inputs* in your *model*
- a model is a *system* which sets *relations* between various *elements* and which needs *start parameters*, which can be *start data* [...] for me validation is ensuring that when taking *real data* (coming from *measurements*), the so called "*inputs*", you get *measurable data* as *outputs*.
- for *parameters* for which we did not had *measurements*, we fixed them to an *arbitrary* value. Others were chosen, *calibrated* manually, in order to have a model giving *realistic* values.
- more FOCUS scenarios must be *created* because they are considered as not being very *realistic*
- when you *compare* with the *reality*
- we can define the scenarios as different *parametrizations* of a model, with the purpose of being compared between them and/or to *represent* various *realities* or *situations*
- [a scenario is] a *representative* combination of crop, soil, climate and agronomic *parameters* to be used in *modelling*; *representative* means in this context that the selected scenarios should *represent physical sites known to exist*, i.e. the combination of crop, soil, climate and agronomic *conditions* should be *realistic*
- with the *models* there are many *a priori, and hypothesis*. There is no guarantee that the procedure is secure, in the absolute: we can forgive a *process* in the model.

Parameters and scenarios are closely related in our informants representations. The three following excerpts, amongst others, illustrates this:

- we could define scenarios as different *parametrisation* of a given model
- a scenario is [...] a combination of crop, soil, climate and agronomic *parameters* to be used in modelling
- permanent crops have also been *parametrised* in the scenarios

Those examples show that scenarios are also to be defined regarding the way they are *used* (this will become central when we will define how scenarios represent, in section 5.1).

We may also question what are the bounds, at the symbolic level, between parameters and scenarios, and their respective relations to the "reality" as conceived by our informants. Let us see the context in which the word "parameter" (and derived or related words like "parametrisation", "parametrise") is used, and examples of phrases about specific parameters (like "organic matter"):

- for PARAMETERS for which we did not have measurements, we fixed that to an arbitrary value
- other PARAMETERS were calibrated manually
- PARAMETERS of soil and climate
- our choices of ORGANIC MATTER content
- we did not choose well certain soil PARAMETERS
- if you had to calibrate your PARAMETERS, what shall you say to industry?
- it is a matter of precision and controlled PARAMETERS
- PARAMETRIZATION of a model
- crop, soil, climate and agronomic PARAMETERS
- PARAMETER estimation
- soil PARAMETRIZATION
- climate PARAMETRIZATION
- permanent crops have also been PARAMETRISED in the scenarios

That list shows clearly that parameters values are negotiated, chosen, simplified. The plausibility of parameters is discussed, distinctions are made between some parameters that must be measured and others that can be calibrated.

Interpretation: scenarios and representativeness

Scenarios, parameters, data, measurements and reality are closely linked and related by our informers.

A first interpretation can be made starting from the use of the word "parameter(s)": they are the parameters "of something": soil, climate, organic matter, agronomic parameters. Moreover, there is an action which corresponds to the parameters: "parametrization". "Parametrization" is the word which indicates the action of mapping the parameters and the "thing", the object which corresponds to them. But it indicates also the values of the parameters resulting from this mapping when they are provided to a model. One can also notice a certain number of operations which carry out the mapping between the parameters and the object:

to fix
to calibrate
to create
to parametrise
to select
choice
estimation
measurements

The relation between "reality" and the scenarios or parameters is described (by our informants) in two ways: it is a matter of distance and representation. What is at the center of the representation in our anthropological terrain can be described as follows[2]:

+ reality	− scenario
the object as a thing in itself	the represented object
physical sites known to exist	

Moreover, the object itself and its representation are separated by a distance that is described in various qualitative ways:

+ small distance	− big distance
representative/realistic/real	unrealistic
measurable	arbitrary
real data	estimated
real	hypothetical
high precision	low precision

[2]In the following tables, the + and - signs are arbitrary signs to represent opposed terms, as usual in a structural analysis.

To those qualitative descriptions corresponds a series of methods. The use of a given method is an indicator of the distance to the "real object":

if	if
measures experimental data optimization/calibration	arbitrary theoretical data manual calibration choice estimation
then + small distance	then - big distance

This structural analysis of a symbolic system indicates us how our informants conceive the scenarios in the framework of "representation", and it gives indications about what are good or bad representations (i.e. this is a concept of "representativeness" quality). Using this findings, our next sections will formalize the researchers representation concepts in mathematical terms and implement it in a representativeness mathematical concept that is operational for quality assessment.

3.2.2 Formalizing the researchers' conceptions in mathematical terms

Our proposal is that a scenario can be assimilated to a symbolic object, and in particular that scenarios are intensions[3] which represent by having an extension. This is what we call our "View 1" about representation. More complexity will be added to in chapter 5, which will present View 2 and View 3 that will take into account sociological and philosophical concepts.

[3] It is necessary to distinguish between intention, intension, and extension. Intention is referring to the purpose, aim or goal. Intension is in contrast encountered in logic and sets theory. In the sets theory, the intension is a definition of a set by mentioning a defining property, while the extension is the definition of a set by enumerating its members. See Weisstein (2008a), Weisstein (2008b), Makins (1991). Another term that must be distinguished in philosophy from intension is the "intentionality" that is discussed p. 115.

Our concept will be formalized and *implemented* mathematically in the terms, concepts and notations used in the set theory, and particularly the *symbolic objects framework* as presented by Bock and Diday (2000)[4,5].

According to Bock and Diday (2000), the symbolic data analysis (SDA) extends the classical methods of data analysis to *symbolic variables* that may have values such as subsets, intervals, frequency distributions; logical relationships may be defined between the variables (by rules or taxonomies). Classical data analysis is a special case of the SDA. Interesting also for us is the fact that SDA allows for measuring *similarity* between symbolic objects, and we will make use of this for defining a quality assessment method. Moreover, as seen in chapter 1, our scenarios often comprises a geographic component, and geographical symbolic objects can be defined and used to make aggregated regional comparisons.

Any SDA (Diday and Esposito, 2003) is based on four spaces: the space of individuals and classes of individuals (that we will both term here "objects"), the space of concepts, the space of descriptions that model the objects, the space of symbolic objects that model the concepts. The formalism used to describe symbolic objects is described in Diday and Esposito (2003), we present a brief summary of it in appendix I.

We will implement part of this formalism in our description of the way scenarios are representing. Bock and Diday (2000) assimilates the symbolic objects to the classical concept of *intension*, so we will use it because we believe it is more widely known and helps thus in communicating our proposal.

In our view, a scenario, when representing, is a symbolic object which represents *by having an extension*. This framework is presented in fig 3.3.

Members of an extension can be described by a set of p variables $\{Y_1...Y_p\}$ (possibly of different types). Each of those variables takes its values in a set ϕ_i of possible levels, alternatives, numbers, called a *domain* (Bock and Diday, 2000). Here we use the term *description domain* for the domain of all the variables used to describe the objects. We distinguish between two extensions: ESE and DSE. ESE, the enumerative

[4] It worth noting that the book of Bock and Diday (2000) and the article Diday and Esposito (2003) seems to use the symbolic objects mathematical framework within a realist metaphysical framework, to which we do not agree.

[5] The use of mathematics for expressing philosophical concepts can be found e.g. in van Fraassen's use of mathematical structures.

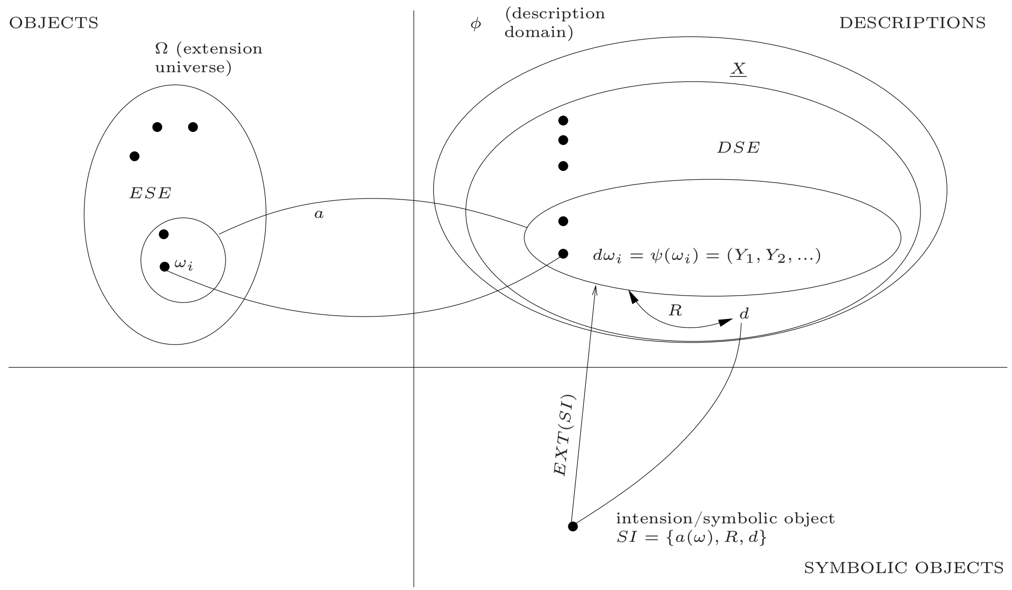


Figure 3.3: *Scenario's view 1. The scenario's intension, described as a symbolic object SI, represents by having an enumerative extension ESE. A scenario's intension is defined by a,R and d:* $SI = \{a(\omega_i), R, d\}$.
R: a mathematical relation (e.g. $\{<, >\}$); d: a description vector (e.g. $\{25°C, 100\ mm\ mean\ annual\ rainfall\}$); ω_i: an object: the symbol used to indexically locate (represent) a phenomenon in the logical space of the scenarios; Ω: extension universe: the set of all objects (e.g. a population, in contrast to a sample); $d\omega_i$: the description of ω_i using a vector $\psi(\omega_i)$ of symbolic variables Y_i; ϕ: the description domain: the space of all the possible levels, alternatives, numbers of the symbolic variables; $\underline{X}$: the symbolic data array: the database (or data model) available for the descriptions; ESE: enumerative scenario extension: is the set of the objects that are represented by SI and is termed ''enumerative'' to highlight that it enumerates the objects' symbols (the i^{th}) and does not refer directly to the phenomena; DSE: descriptive scenario's extension: the array of the description vectors of objects in ESE; a is a membership function that allows either to decide which objects ω are members of ESE or to estimate the degree of membership of a given ω_i to ESE; $EXT(SI)$: extension of SI into ϕ.

scenario's extension, is the set of the objects that are represented by a symbolic object SI. ESE is termed here "enumerative" to highlight that it enumerates the objects and does not refer directly to the phenomena which in fact are only accessible through indexical location[6] which is not a mathematical operation. DSE, the descriptive scenario's extension, is the array of the description vectors of objects in ESE. The symbolic object "scenario" (SI) is defined by a mathematical relation R, a description vector d, and a which is a membership function that allows to decide which objects ω are members of ESE (or their degree of membership)[7]

In such a framework, it worth noting that a scenario represents by having a ESE, while DSE is not strictly necessary for representing: something can be taken to represent something else, without the need for any detailed description of the represented objects. However, one of our goals is to introduce a concept of representation using scenarios that allows for some sort of empirical assessment of the scenarios' representativeness. As we will see in section 3.2.3, the descriptive scenario' extension DSE is the structure that will be matched to a *reference extension* in order to assess the scenario' representativeness.

3.2.3 Representativeness quality component

Scenarios' representativeness assessment and symbolic objects

Having in mind the conceptual framework of representation presented in section 3.2.2, it is possible to define an intension. From that intension, it is possible to find two corresponding extensions (scenario's extension (SE) and reference extension (RE)). A definition of the *representativeness* of a scenario can then be given, that allows its estimation: the representativeness is defined here as the *degree of matching* between the scenario's extension (SE) and a reference extension(RE), both extensions being described in the same common description domain. The degree of matching is estimated using a *matching function* ($MATCH$). This conceptual framework is presented in fig. 3.4.

[6]An indexical term is Makins (1991) in linguistics and logics a term whose reference depends on the context of utterance, such as I, here, now, you, tomorrow. In the present context, an indexical location is thus an operation made by a user of the representation (the "I") that establishes a link between the phenomena and the objects in the ESE.

[7]More formally in the formulation of Diday and Esposito (2003)a is a mapping defined by Ω in L depending on R and d, and L is either $0, 1$ or $[0, 1]$. See appendix I.

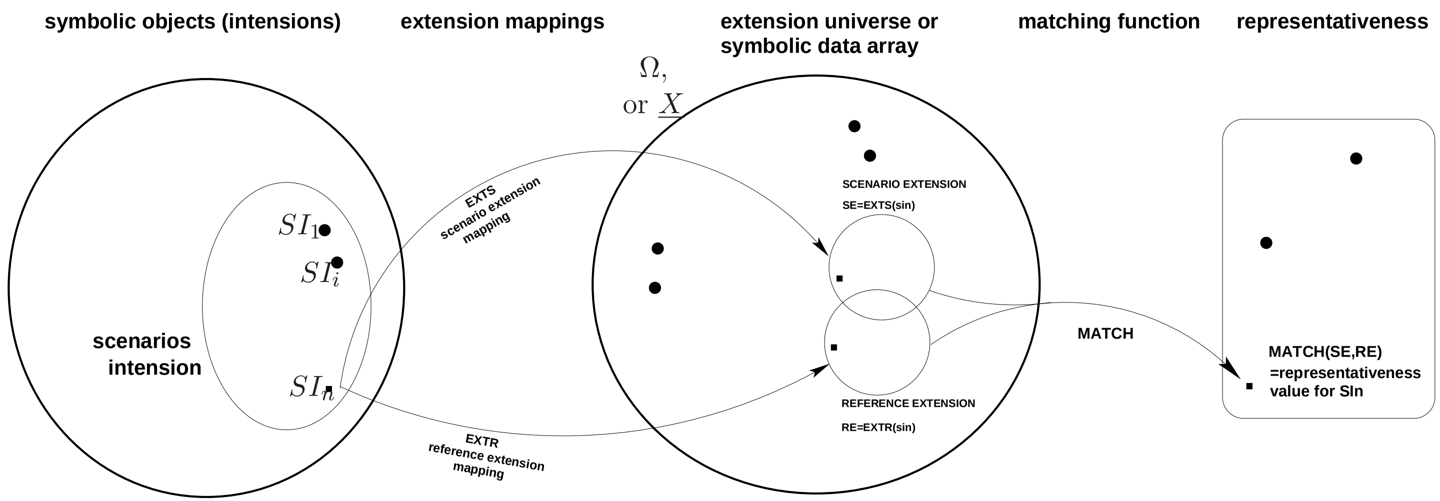


Figure 3.4: *Conceptual framework of the scenarios' representativeness estimation.*
From the intension of a scenario (SI), using an extension mapping ($EXTS$), it is possible to define its extension (SE). SE can be either ESE, the set of corresponding elements in the description domain having the properties that agrees the criteria defined in the intension of the scenario, or DSE, the descriptive scenario extension (the array of the description vectors of objects in ESE). Extensions can thus be defined on $\underline{X}$ or the extension universe Ω. SE is the extension that is claimed when the scenarios were constructed. From SI, it is also possible to define a reference extension (RE) using another extension mapping ($EXTR$) used as reference for quality assessment. The estimation of the representativeness is based on the comparison (matching) of a reference extension (RE) with the scenario's extension (SE), using a matching function $MATCH$.

The particular form of the matching function is dependent on the type of scenarios that are considered, the databases available, the models available, etc

The intension of a scenario is the result of a mathematical *interpretation* of the scenario's representing concept. More generally, interpretation is needed for defining the scenario's representing concept, intension, extension mapping, extension and description domain. Chapter 4 and 5 will show that the scenario concept is the result of an interpretation that in its turn is the conclusion of a social translation process. The intension is linked, through interpretation, to the scenarios' use, objectives and application context. This is particularly important in a regulatory context.

Having the scenario's intension, it is possible to define an *assessment question*, which is the translation of the intension into the interrogative mode (e.g. if a scenario is representing "worst cases of leaching" according to the scenario's interpretation, then an assessment question could be: "Are the scenarios representing worst-cases of leaching accurately"?).

> We propose to call *scenario's representativeness assessment* the action of answering to an assessment question by defining the reference intension, the extension function, a reference extension, a matching function, estimating the representativeness, choosing a representativeness threshold and finally performing a *representativeness test* which is the comparison between the result of the matching function against a *representativeness threshold* (an accepted representativeness level). The threshold depends on the purpose of the scenario and the particular context in which it is applied.

This representativeness assessment concept is applied in detail in chapter 6 to the FOCUS scenarios.

If we are looking to similarities between this matching concept and van Fraassen's empirical adequacy (van Fraassen, 2008)[8], DSE is playing the role of van Fraassen's model's empirical *mathematical substruc-*

[8] See appendix H.

ture[9], which is compared to (here: matched to) a *surface model*[10] (here a reference extension) for evaluating the empirical adequacy (the representativeness in our case).

Measurement representativeness

We use here the term "measurement representativeness" for parameter values coming from measurements. Instead of using the term "measurement representativeness", we could have used "plausibility", or "realism", but they have problematical philosophical meanings, while our choice is referring to an empiricist view[11] as presented by van Fraassen (2008) (as we will see section 5.1). We consider here the empirical adequacy as a quality assessment. Measurement representativeness assessment can be performed using the $MATCH$ function presented in fig. 3.4. In this case, the matching is between two surface or data models, as defined by van Fraassen (2008). This can also be interpreted in Latour's terminology as the successful alignment and the solidity of circulating references (chains of inscriptions) which have led to the values of the parameters integrated into the scenarios (this will be detailed in section 5.1, see also Latour (2005), Latour (2001)). This certainly also results from good laboratory practices, choice of the appropriate method, etc.

For the PRIBEL scenarios, there is a claim that their parameters values comes from measurements or/and surveys. So measurement representativeness applies to them.

For the case of the FOCUS ground-water scenarios, it is considered that the measurement representativeness of the scenarios is not an issue. Indeed, scenarios were defined by experts and we consider that it is very likely that the conditions of the standard scenario may occur somewhere in Europe. Moreover, FOCUS scenarios are said to be "virtual", i.e. there is no claim that they must have a measurement representativeness. What FOCUS scenarios represent is determined by their function (i.e.

[9] According to Minogue (1984), van Fraassen's structures are mathematical structures in general: *"any mathematical structure is open to use even if it lacks a precise meta-mathematical analysis"*. However, in the models theory (as part of mathematical logics), intuitively a structure is a set that one wishes to study, equipped with a collection of distinguished functions, relations, and elements (Marker, 2002).

[10] Which is derived from measurement through some sort of idealization or interpolation.

[11] See appendix H for philosophical background.

their use), and this is a matter of functional representativeness, as it will be shown here below.

Functional representativeness

In the case of functional representativeness, their *function* must be evaluated. Representation can be done by using something "instead of" something else for a given purpose: that conception will be detailed in chapter 5. At this stage we can however say that FOCUS scenarios, as we saw in chapter 1, are claimed to not represent specific fields by measurement representation. FOCUS scenarios are used instead of the real EU agriculture (fields) to find if safe uses of a given pesticide exist. In that sense, the scenarios, when used with a model, are used to represent worst cases of pesticide use (the 90th percentile). In consequence, FOCUS scenarios extension SE can be defined as the worst cases as calculated with the FOCUS models, and a reference extension RE can be defined as the set of the worst cases as calculated with a reference (more precise) method. RE and SE can be compared by a matching function like in fig. 3.4. We will apply this in detail in chapter 6 for the FOCUS scenarios case.

We suppose that the assessment of functional representativeness can be done in different ways in the case of other scenarios than FOCUS, and this will depend on the use that is made of such scenarios.

3.2.4 Theoretical consistency quality component

The scientific theoretical quality component is defined here as the consistency between the scenario parameters and the theory to which the used model belongs. It can also be understood as a critical assessment of the parameters choice, the theoretical implications of the chosen parametrization, the accepted simplifications, the unknown parameters that are calibrated, the relation between the scale at which the parameters are measured and the scale at which they are used in the model, etc.

3.3 Defining the logical component

For testing the logical consistency, it should be analysed if the scenario's narration[12] respects the rules of a reference logic. Semantic tech-

[12] Descriptions, history, etc, see chapter 1.

niques such as propositional calculus (Gries and Schneider, 1993) could be applied to analyse the logical consistency. It is possible to consider a scenario and its narration as a set of logical propositions, and look for consistency or truth values. Other logics can be considered as possible candidates to investigate their usefulness in internal consistency assessment (e.g. Zalta (1988), or Andrews (2009)).

To illustrate the issue of logical consistency, consider the following example. FOCUS scenarios have been developed as part of the implementation of the 91/414/EEC Directive at the EU level. However, the conditions of the standard scenarios inferred from western European data bases and experiences, will not be relevant for evaluating pesticide emissions in the tropical areas of the 'Département d'Outre-Mer' of France which also belongs to the EU. Other examples can be given: the scenarios are supposed to represent arable land, but grasslands are parameterised, most of them located in non-arable land; the same is true for olive trees.

This illustrates that FOCUS scenarios are not logically consistent in some aspects. These examples shows also that an interpretation is necessary for checking the logical consistency. We note however that logical inconsistency does not necessarily imply a bad representativeness.

3.4 Defining the juridical component

In this section, we will show that FOCUS has a legal role, in its function as the compiler of guidance documents establishing the way models and scenarios required by the 91/414/EEC Directive should be used. The assessment of scenarios produced by FOCUS-GW may therefore be considered from the legal perspective. This is not a unique situation: legislative documents increasingly make direct reference to the use of models and scenarios. In this sense, scenarios must conform to the "spirit" of the legislation that has warranted their existence. But how can one evaluate this conformity? In order to answer this question, we must start from what is the "truth" in legal matters.

3.4.1 The juridical truth

This section is in large part based on the book by Atias (1994). There are a number of contending theses in the area of legal epistemology. In them, we see the classic division between the nominalist, the Platonic, the Cartesian, the Kantian and phenomenological conceptions of what constitutes reality or truth. In parallel with these divisions, we may also take an interest in legal practices themselves.

On what can truth be founded, in matters of law? If we follow Kelsen in his notion of law:

> *[...] a norm [...] is valid if it conforms to the superior norm of the procedure according to which it has been laid down. Which means that the law must be formally compliant with the Constitution, the implementation decree must be formally compliant with the law, the provision must be formally compliant with the decree, the legal decision must respect procedure etc...* (H. Kelsen in Touret (2001))

Starting from this idea, it clearly appears that in order to be legally valid, the FOCUS scenarios must be compliant with the 91/414/EEC Directive. Now, this notion only postpones the problem: what is the fundamental norm on which all others are based? The Constitution is at the summit of the pyramid.

> *Since a norm can find its validity only in a higher norm, the question of the validity of the Constitution arises. So the fundamental norm is formulated thus: 'We must obey the commandments of the constituent'. It is a technical fact and not a moral and/or religious obligation. It is a functional necessity [...].* (H. Kelsen in Touret (2001))

But there is a second problem: even when the law goes into the finest detail, the law remains subject to interpretation. Atias writes that:

> *Power inevitably belongs to he who deduces the particular consequence from the general affirmation [...]. Everyone who contributes to the description, that is to say to the expression of law, owns a portion of it.* (Atias, 1994)

Following this, FOCUS guidance documents and scenarios have indeed a juridical function. The problem of the assessment of FOCUS scenarios in this context may be approached as an interpretation problem and I will deal with this aspect in the following section.

However, if we consider an epistemological point of view that goes beyond the “thinking subject/real object” distinction, and that incorporates the facts as fundamentally constructed, truth value can not be defined in terms of absolute objectivity unless, this is intersubjectivity. The utility of a hermeneutic analysis should be relocated within this context. In addition, according to Atias (1994), objectivity, in a certain idea of the world, might be the characteristic of what allows the enrolment of the largest number. He also says that we might consider that the criterion of the truth of law can only stem from a value judgement about the result obtained. Finally, we might say that the problem comes from a casuistical conception (Singleton, 2000) of law that seeks to understand every new situation as a particular example of a more fundamental essence, which would be the source of law. We should therefore take the position of considering only democratically guaranteed convictions as relatively absolute, and in this we would side with Atias (Atias, 1994). As a result, we should consider the FOCUS scenarios in terms of the use that has been made of them both by industry and by the administrative body in charge of examining pesticide dossiers, and we should submit this use to the guarantee of democratic conviction. We will come back to this point in section 3.5.

3.4.2 Juridical hermeneutics

If we wish to take into account the possibility of scenario’s assessment by using hermeneutics (despite all the reservations expressed in the preceding paragraph), what could be its foundations, method and relevance? Interpretation is always subject to discussion, Atias (1994) tells us. But is this not the role of interpretation? Not to contribute to defining a truth but to stimulate a debate and thereby contributing to the creation of a law based on democratically guaranteed convictions?

In this context, the word “assessment” would assume more the meaning of a continuous work towards a “negotiated interpretation”, the co-construction of an interpretation. In some way, assessment here would be saying: “Here is my interpretation of your interpretant. It is the product of a certain place and of a certain cultural logic. As such, it is not entirely singular; others could arrive at similar interpretations (i.e. I am probably representing a fraction of the population) and as such you, the legislator, could take them into consideration.”

However, the person performing such an assessment does not escape the role of expert, but in this case his legitimacy is not derived from his greater closeness to a supposed ‘‘objective reality’’ (‘‘I know what arable farming is and you don’t’’) but simply from time passed in reflection on the subject and on the plausibility that FOCUS members will give to his interpretants... If the texts produced by FOCUS are texts with legal value, the interpreter can take this into account in his interpretation of the interpretants produced by the largest possible segment of the population: the idea of agriculture produced by FOCUS versus that produced, for example, by a José Bové. It remains fundamentally true, however, following what Atias says (Atias, 1994), that legal power belongs to those who interpret... As such, my interpretation of the Directive itself (as performed in the APECOP project, see chapter 1 and Vanclooster et al. (2003)) had a very small yet a possible chance of modifying it, since it was located in a process of revision at the level of the Commission and of the Parliament. We will return to this point in the following paragraph.

3.5 Defining the political component

3.5.1 Political choices embedded in simulation scenarios

What do we mean by ‘‘political choices’’ in regulatory scenarios? Any general definition of the ‘‘political’’ will likely be controversial. Broadly speaking it concerns the domain of power, power allocation, resources allocation, the systems of authority, the state, and the public decisions that concerns citizens collectively, the whole social body, the collective. In that sense, a political choice is often the expression of a preference between various possible actions that may have a collective impact. Regulatory scenarios are used as decision-support tools generally for administrations in charge of the concerned domain, like the EU Commission and Member States administrations. The FOCUS simulations, based on FOCUS scenarios, are used as part of the arguments for decision-making. Ideally the criteria for decision making should be precise enough in the legal texts for scientists to translate them into a unique set of parameters in the scenarios. This is however not realistic. Alternative equally scientifically justifiable interpretation choices remains, within a certain range. This is clearly shown by our presentation of FOCUS scenarios in chapter 1: the 91/414/EEC directive states e.g. that a ‘‘*realistic worst-case estimation must be made*’’, and FOCUS translated that into a set of parameters and a 90th percentile vulnerability. Other equally scientifically justifiable choices could have been made for translating (interpreting) the directive into scientific parameters, and a space remains for interpretation of the 91/414/EEC directive. This is shown also in our chapter 4 were an account of negotiations of various interpretation options of the FOCUS scenarios is presented. That is the sense we give to ''political choices'' embedded in regulatory scenarios: *choices between various scientific options that are equally justifiable scientifically, and that may lead to different regulatory decisions.*

Then come various questions: Who should decide between those options? How to lead with such interpretation choices when performing scenarios assessments? What is a good political choice? How to assess the social acceptability or robustness of such choices? We present here below a vision in terms of political ecology (Latour, 1999) that will allow:

- the interpretation in terms of power of the role of FOCUS in its creation of scenarios;
- the proposal of an alternative idea of assessment that is not dependent on the philosophical realism and that does not short-circuit a certain vision of the democratic process.

3.5.2 Latour's power, collective and state concepts

For Latour (1999), our world (called ''the cosmos'') is constituted by ''institutionalized essences'' (both humans and non-humans), which enter into the world by a continuous process called ''the collective''. This process relies on four main steps: perplexity (sensibility to new proposals for essences), consultation (witnesses are searched for those essences), hierarchization (negotiation of the way the essence can enter into the cosmos), and institutionalisation (essences are accepted or rejected). These four functions are grouped in two *chambers*: the upper chamber (perplexity and consultation), and the lower chamber (hierarchisation and institutionalisation). The power of the upper chamber is said to be ''power of taking into account''; that of the lower chamber is said to be ''power of putting in order (arrangement)'', and there must be, for the good functioning of the system, a clear separation between the two powers. These concepts are described in detail in our appendix J.

3.5.3 Interpreting the role of experts in terms of power

If we attempt to view the work of FOCUS in Latour's terms, then the FOCUS-GW scenarios constituted a proposition that contained among other things the desire to "represent agriculture in Europe". The chosen representatives of this agriculture were mainly scientists (since the "discourse of agriculture" was translated into figures that constituted the scenarios). The jury was made up of politicians, industrialists and scientists (as shown by the composition of FOCUS steering committee). The work of hierarchization was short-circuited since "the scenarios have been developed scientifically, and therefore represent agriculture". They were therefore established, and the scenarios are in use at the present time.

In such a context, the vital functions of the collective as it currently works are short-circuited by the legislator and by scientific experts. Scientists do not only have a function of perplexity, which is normal according to Latour, but it happens that they present themselves not only as the only representatives of "objective scientific facts" but also as the only jury. Furthermore, in giving weight to the transcendence of "facts as they are", the work of hierarchization is disturbed since only scientific ("true") arguments are acceptable. Many propositions are therefore relegated or rejected as irrational or ideological. We thus end up at an unfortunate separation of power and at a concentration of power in the hands of scientific arguments. In addition, these arguments are put forward by industry, in the case of FOCUS, since industry contributed to choosing scientists as the only jury of the "scenarios" proposition.

What other juries and representatives might have been found additionally to those chosen by FOCUS? Scientists are not the only figures able to say, "what agriculture is". In fact, in the scenarios themselves there is a definition of the concept of agriculture that I might characterise as "existing practices but without guarantees of representativeness". A different witness might understand agriculture in a different way. The scenarios therefore conceal choices that are not without consequences under the innocence of objectivity. At the very least, this needs to be considered.

3.5.4 Proposals for a collective scenarios' quality assessment

Starting from the notion of the collective as defined by Latour, if FOCUS and the scientists plays both a function of consultation (e.g. as "representatives of groundwater") and institutionalisation (when having the last word about scenarios' "validity"), then we are in a situation where the separation of the powers of the two Latour's chambers is not guaranteed. FOCUS role should be limited to the consultation function. The word "validation" should also be redefined, due to the fact that in its present use in the sciences it is often used to mean "comparison in relation to data", with data conceived as representing the phenomena "as they are". I would suggest that an assessment is seen as the outcome of a slow process of creation of a common world. This would also be closer to what happens in practice when mathematical models are used. In fact, very often mathematical models consider complex phenomena for which there exists no known experimental protocol. Knowledge of the elements of a system is used as the basis to deduce the functioning of the whole, which rapidly and unproblematically ends in discussions in terms of "plausibility". I have been able to observe this on several occasions in various scientific projects, but only when the researchers are among themselves. Why then does "validation" take on the meaning of "tell the truth" when scientists are in interaction with, for example, politicians, as in FOCUS? Are scientists instrumentalised or do they actually know what power they have in using "truth" as a weapon? A little of both, I believe...

3.5.5 A type of political assessment: participatory modelling

In Latour's terms, it is proposed to introduce participation in the following steps:

- Regarding the perplexity function (i.e. at the step of the construction of the scenarios and before their assessment), scenarios can be open to a variety of actants. This is a way to improve the functioning of the collective, as opposed to a situation where only scientists and industry are closely involved in scenarios' construction.
- Consultation implies the search for trustworthy witnesses, who are the representatives of propositions. If scenarios involves choices, interpretation, use in a context, then regulators, scientists and industry are not the only possible witnesses (as in FOCUS). This opens the door of assessment e.g. to NGOs, but also to lay persons.

Chapter 7 presents the development of a scenario's participatory construction and quality assessment method, for pesticide risk assessment, as well as its feasibility test. Participatory construction means here that the construction of the scenario itself is made participatory, and not only its use or assessment. Although NGO's, politicians, regulators are involved in the method presented in chapter 7, the emphasis is put on *lay persons* participation. This clearly is a different approach than the classical stakeholders' involvement.

Hierarchisation and institutionalization are often devoted to regulators, administrations, politicians, but lay persons may have a role in those processes. However this aspect is not developed here: as it will be presented in chapter 7,we focus on the co-construction, consultation and cognitive aspects. The incorporation of our participatory proposal into formal democratic decision processes is not tested in this work.

3.6 Chapter conclusion

In order to implement step C, that requires the definition of the scenarios quality components to be assessed, representativeness, logical consistency, juridical conformity, and scientific theoretical consistency were presented, as well as an approach to a political quality. We propose that (1) for testing the logical consistency, it should be analysed if the scenario's narration respects the rules of a reference logic; (2) juridically speaking, scenarios must conform to the legislation that has warranted their existence; (3) the scientific theoretical component of the quality is defined as the consistency, in a broad meaning, between the scenario parameters and the theory to which the used model belongs, and in that sense the implications of the chosen parametrization should also be assessed (justification of the choice, simplifications, scale effects, etc); (4) for testing the representativeness, the degree of matching between a scenario extension and a reference extension is to be assessed; and (5) politically, a collective assessment must be allowed to discuss the choices involved therein and the way interpretation is made.

A symbolic systems analysis of the scientists' discourses was presented. It appears that the relation between the scenarios and the reality is at the center of scientists' conceptions about scenarios, and scenarios are continuously linked to parameters and measurements. It shows that the link to reality is conceived by our informants as a representation. The representation is seen as a relation between an object in itself (existing without the observer) and the object as represented, the distance between the two being expressed through a series of qualitative terms. Moreover, the method used to construct the scenarios is in itself an indicator of the distance. This discourse will be contrasted in chapter 4 with the practices of representation through scenarios, that shows that the referent of the representation is fluctuating according to the evolution of a controversy, and involves interpretation.

The representativeness is defined here as the degree of matching between the scenario's extension (SE) and a reference extension (RE), both extensions being described in the same common description domain. The degree of matching is estimated using a matching function $MATCH(SE, RE)$. The implementation of our generic steps for the representativeness assessment is as follows: answering to an assessment question (step C.1, C.4.1), by defining the reference intension, the extension function, a reference extension (step C.4.2), a matching function (step C.4.3), choosing a representativeness threshold (step C.4.3), and finally performing a representativeness test (step E) which is the comparison between the result of the matching function against a representativeness threshold (an accepted representativeness level). The threshold depends also on the purpose of the scenario and the particular context in which it is applied (step C.2). The particular form of the matching function is dependent on the type of scenarios that are considered, the databases available, the models available, etc.

Two cases of representativeness assessment must be distinguished: measurement representativeness, and functional representativeness. When the parameters of the scenario come from measurements, and that the use of the scenarios is to use those parameters to represent phenomena through those measurements, then the matching is between two surface or data models (in van Fraassen's sense). In the case of functional representativeness, the use of the scenarios must be evaluated to define the matching function, that is consequently highly dependent on the particular scenarios use. This is shown with the FOCUS-GW scenarios that are conceived to represent worst-cases of leaching when used with a FOCUS model. These scenarios do not represent through their parameters (that are explicitly claimed not to represent any specific field and are thus termed ‘‘virtual’’ by scientists) but through their function in modelling: a functional representativeness can thus be assessed by a comparison of the scenario extension (i.e. the description of the worst case according to the modelling results) to a reference extension describing the worst-cases via a more accurate method.

In the case of a functional representation, measurements may not play any role in the way scenarios represent. This will be exemplified in chapter 6.

Regarding the political component, starting from a view of the politics in terms of Latour's power, collective and state concepts, as well as an interpretation of the role that experts and their scenarios assessment should have in the collective, it is argued that scenarios quality assessment should be collective. Assessment is seen as the outcome of a slow process of creation of a common world. Representation using scenarios involves measurements, but also interpretation. A collective assessment must be allowed to discuss the choices involved therein and the way interpretation is made. This is all the more critical when mathematical models consider complex phenomena at scales for which there exists no known experimental protocol. As a particular form of collective assessment, lay persons should be allowed to participate into the perplexity and consultation phases of the collective functioning. This discussion in terms of political philosophy is the starting point of a cognitive test of a participatory assessment presented in chapter 7, that aims at knowing its potential for generating a range of interpretation choices of the scenarios, scientifically acceptable, but that have potentially different effects on the regulatory decisions. Those equally justifiable choices are termed here as political choices embedded in the scenarios.

Some of the criteria we developed for scenarios assessment are partially independent from the models with which they are used. However scenarios and models are linked, and thus the assessment of scenarios depends also on the models: this is e.g. the case for the FOCUS scenarios assessment, for which the scenarios extension depends on the use of the scenarios with a particular set of models. This is also the case if a juridical assessment of the FOCUS scenarios is to be made, and the 91/414/EEC directive mentions scenarios through their application with models. When assessing theoretically a scenario, it should be verified at least if the scenario provides the necessary parameters for the model. In a political assessment, the use and role of scenarios in a regulatory context can also be evaluated through the social consequences of the use of scenarios with a model. Despite this link between scenarios and models, there is also some independence between the two from the assessment point of view: a scenario may have a high quality and be used with a poor quality model, or a scenarios may be used with more than one model.

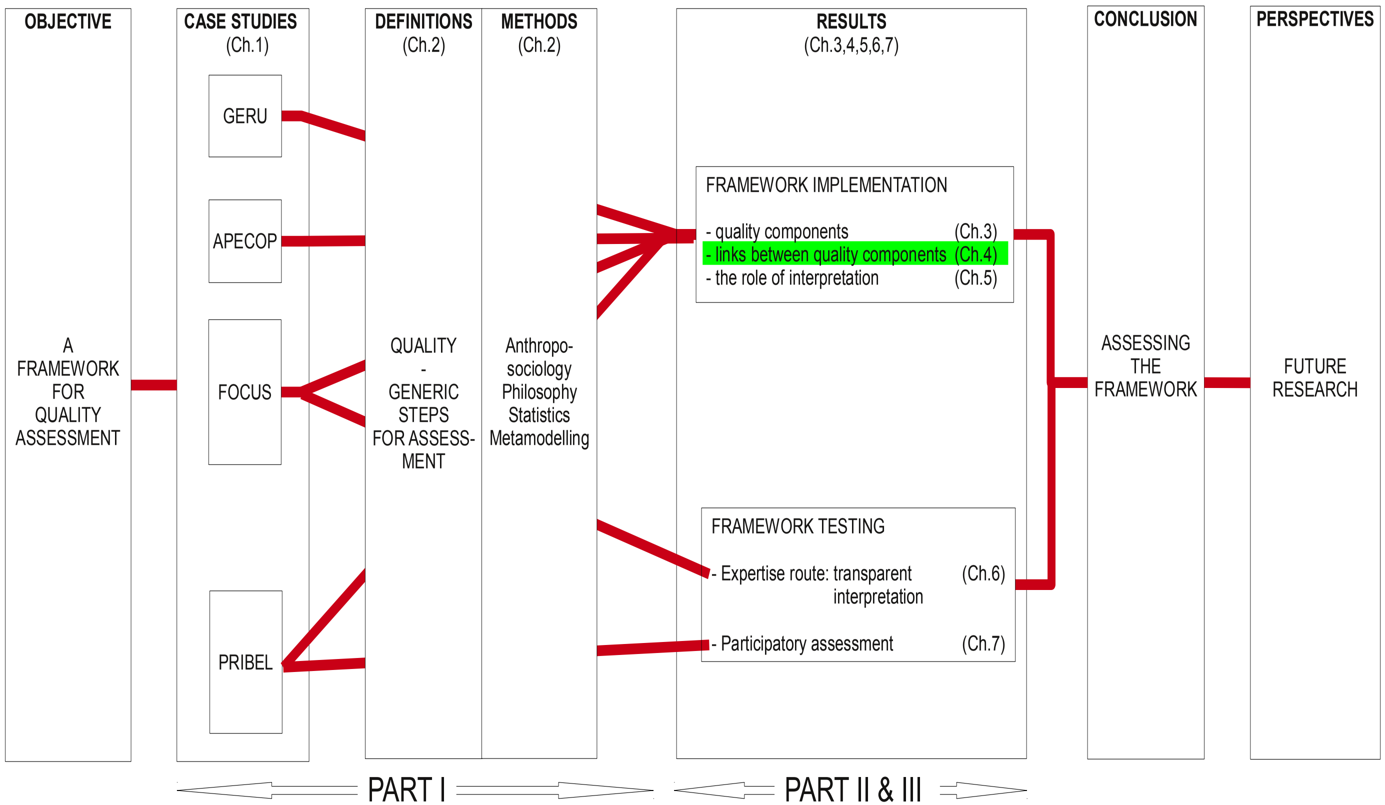
If the six criteria presented in table 3.1 are examined in the way how they can be implemented, it appears that interpretation is part of the process of quality assessment in each case. Juridical assessment implies laws' interpretation; political assessment implies e.g. interpretation of the political function of the scenarios according to some reference political conception of the society; representativeness involves scenarios' narration interpretation; logical consistency implies interpretation of scenarios' narration and parameters into some reference logic; theoretical consistency assessment implies interpretation of scenarios' parameters into some reference theory.

The two next chapters aims at defining the role of interpretation in quality assessment: chapter 4 will present an anthropological participant observation of interpretation processes in quality assessment, and chapter 5 will explore that question philosophically, and propose concrete ways to take into account interpretation in quality assessments.

CHAPTER 4

Determining the links between quality components

Figure 4.1: Reading guide of chapter 4. This chapter analyses, through an anthropological observation of the APECOP project case, the links between the quality components that where defined in chapter 3. The results are used in chapter 5 that presents ways to take the links into account in quality assessments.



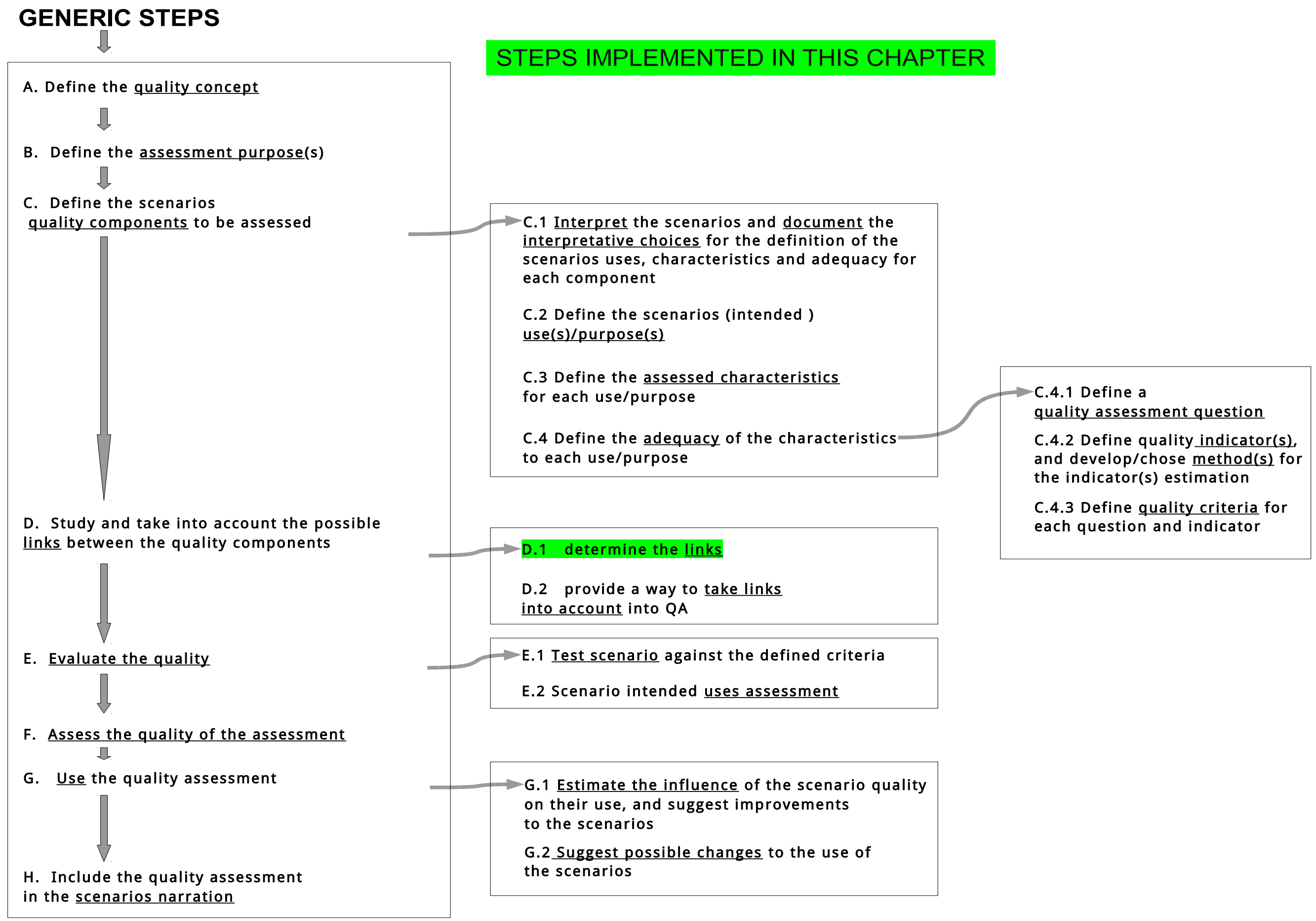


Figure 4.2: Reading guide: contribution of chapter 4 to the implementation of the quality assessment framework steps. This chapter implements step D.1 by analysing the links between the quality components, through an anthropological observation of the APECOP project case.

Summary

The scientific conceptions about representation by scenarios in terms of distance to and simplification of the reality (as describe in chapter 3) are quite different from the practices of representation of the scientists when confronted to a regulatory context. The links between the social actors negotiations and the outcomes of APECOP, studied from a socio-anthropological point of view, shows that some of the quality components defined in chapter 3 are not independent: different scientific quality levels are obtained for the same FOCUS scenarios, depending on initial interpretation choices that are equally scientifically justifiable. The choice between various interpretations may then result in different proposals for changing the scenarios, with, as a consequence, different regulatory decisions. The choice between interpretation options is thus a political component of the quality, and exemplify a link between the political and the scientific components of quality. Similar links between other components are displayed in our observations, but less attention is paid here to them. The existence of links between components can however not be generalized to all components: there may be independence between quality components as exemplified by e.g. the logical inconsistency that does not necessarily imply a bad functional representativeness. In consequence, an assessment of the different quality components must provide a way to take those links into account, and not simply take for granted that they are independent as one could think starting e.g. from conceptions about representation in terms of distance to the reality.

4.1 Introduction

Scenarios are often considered to be based on arbitrary hypotheses and so to be, by definition, impossible to "validate". Another tendency seen in the scientific use of the word "scenario" is to consider scenarios as a more or less accurate representation of a physical reality. Generally when scenarios are considered, they are given only a brief definition of their representing concept. Here is an example of such a brief scenario's representing concept:

> *A representative combination of crop, soil, climate and agronomic parameters to be used in modelling ; representative means in this context that the selected scenarios should represent physical sites known to exist, i.e. the combination of*

> *crop, soil, climate and agronomic conditions should be realistic* (FOCUS, 1995)

In this context, the only possible ‘‘validation’’ is of a scientific nature[1]. Scenarios become enclosed in the objectivity/subjectivity debate and in a certain conception of representation, before the criteria and methods for their ‘‘validation’’ have even been discussed. Once this brief sketch of the scenario's representing concept has been supplied, scenarios practice shows that generally one moves immediately to the list of parameters without any presentation of the method that has allowed the creation of such a list, and without mentioning the criterion that would allow a ‘‘validation’’ of their “representativeness”. At the most, it is noted that the values embedded in the scenarios are the result of databases, without adding much further detail.

The question of the relation between scenarios and reality might seem to be of purely epistemological interest if models were not used increasingly frequently to support decisions of a political nature. As we will attempt to show, both in terms of anthropology (in the present chapter) and in terms of epistemology (in section 5.1) we reach the conclusion that scenarios, in both their operational aspect and their assessment, must also be approached from the point of view of their social dynamic, without losing sight of the political dimension. The classic position of the expert is often described as being that of the person who brings certainties, on which politics makes its decisions. The legitimacy of models and scenarios might then be closely linked to their relation to reality. If models and scenarios contain elements determined by a paradigm, and not by experience then they open the way to drifting if the paradigm itself has not been made clear and subjected to political debate. This is even more important in the environmental field, where modelled systems are complex and take place on a scale that prevents their assessment by experimental methods. This is the case with many models used in the environment, such as the models for the prediction of pesticide leaching that we will examine in particular. But it is also the case, for example, with very long-term climate change prediction models whose predictions for the future are beyond the scope of experimentation. Hence, I. Stengers writes: ‘‘*proof, in the sense in which it is scientifically credible (not to be con-*

[1] In this chapter we will present anthropological observations, and our informants used the word ‘‘validation’’. We argue that assessment is more appropriate, so we let ‘‘validation’’ between brackets when it is used by them.

fused with models and statistical arguments, which depend, like ordinary language, on factors one has chosen to present and which therefore only tests the consequences of such a choice) is the business of the laboratory [...]." "[...] once we leave questions that we will term as strictly scientific, because [...] the kind of experimental methods relating to them all have as their aim the matter of proof, we enter [...] the domain of negotiations, and therefore of the arrangement of power relations" (Stengers, 2000). Does this mean that models should be condemned as frauds? Yet models have acquired political legitimacy (the Rio Protocol, for example) and legal legitimacy (e.g. the intensive use of mathematical models to allow or reject pesticides on the European market). In both these examples, models involve powerful economic interests. The debate often becomes confined to a science/politics dichotomy that operates in parallel to other dichotomies: objective/subjective, nature/culture. In parallel to the BSE[2] crisis and the dioxin affair, and not unconnected to them, the precautionary principle has come to the fore at the political level. This principle itself does not escape from a variety of interpretations in the literature (see Zaccai and Missa (2000)): possible intrusion of the irrational into political decisions, enforced democratic control over decision-making, or a demonstration of the failure of science. In this chapter we will attempt to show, by means of a number of examples, how an advance in this debate can be made if it is accepted that objectivity does exist, but not in an absolute way: a scenario may be objective but within a given paradigm; it implies a certain number of presuppositions, which are necessary to specify and examine. It is also necessary to show the political and economic aspects that help form the paradigm used.

The main aim of this chapter is, starting from a series of observations, to obtain a coherent anthropological interpretation of scenarios. The basis for our observations is the APECOP project, whose goal was to "validate" the FOCUS groundwater (FOCUS-GW) scenarios and models (see FOCUS scenarios in section 1.2 and the APECOP project in section 1.5). The project itself will be presented in detail in this chapter.

The theoretical tools and concepts we will use are taken from the fields of sociology and from anthropology of science. We will make a fairly brief presentation of those tools and concepts whenever relevant to the case of concern in this chapter. For the reader not familiar with pragmatics and rhetoric, a summary of those fields of the language study is presented in

[2]Bovine spongiform encephalopathy.

appendix C. A summary of actor-network theory and social translation is presented in appendix D. We will however locate our approach in the wide context of sociology of science in section 5.1 , when considering a proposal for the way ‘‘how scenarios represent’’.

We make use of the actor-network theory and of the sociology of translation. In that context, the translation postulates sometimes new actants and tries to bring them into existence. Studying an actant, then, is a matter of following its construction. When we examine the socio-technical graphs that track innovations, presented by Latour et al. (1992), we note that they are based on the classification of actants into two groups: programmes and anti-programmes. Here we see the equivalent of holders of the antagonistic proposition and holders of the received proposition, respectively, in rhetoric. There is, however, a point at which these two types of analysis diverge strongly: the question of interests. In fact, for the sociology of translation, interests do not pre-exist interaction; they are the outcome of interaction. For Woolgar (1981), interests are not accessible to observation: sociologists deduce them from their supposed effects and therefore this is a vicious circle in which effects are explained by effects. To circumvent this problem, Callon (1986) puts forward the notion of problematisation, which is not the description of the interest that the sociologist or anthropologist supposes the actant to have but is rather a description of the way in which the actant presents the interests of the other actants. The view put forward by Woolgar (1981) also contradicts the concept of the field proposed by Bourdieu (2001). For Bourdieu, what is at stake in the scientific field is the capital specific to this field: scientific authority. This field is structured by the distribution of this capital between scientists and the issue or interest is very clearly present: increasing one's capital. However, we will attempt to follow the logic of the sociology of translation to its end. We will also refrain in our pragmatic-rhetorical analysis from prejudging the interests of one or other; we will approach this issue solely through the examination of what informants have obtained at the end of the process of translation and how they problematise their interlocutors at the beginning. In addition, a phrase such as:

> *I want FOCUS to be THE standard for at least 5 more years.*[3]

[3] personnal communication from the EU Commission (DG-SANCO) coordinator of FOCUS

can not be classified as part of the rhetorical argument if the rhetorical question is: ‘‘Are the scenarios valid?’’. But it would be possible to take it into account via the actor-network theory, where standardisation would constitute one of the actants. Pragmatic-rhetorical analysis is generally complementary to analysis in terms of actants.

In the present work, we shall use the tools developed by the actor-network theory mainly for the practical way they allow to explore scientific controversies by highlighting the complexity of their context and the variety of actors (actants) that are involved. Moreover, that theory allows to present such context in a synthetic way. However, this does not imply that we adhere entirely to the philosophy of science that e.g. Latour is proposing. This deserves some clarifications, in light of recent controversies (the so called ‘‘science wars’’). Part of the polemic is about the ‘‘scientific truth’’ and the existence (or not!) and the extent of social determination of scientific truth. According to Vinck (1995), the controversy between rationalists and relativists can be traced up to the ancient Greece. For rationalists, the power of a proof comes from reasoning and its relation to experience. For relativists, there are no absolute rationality criteria. For them, consensus comes when an argument is conform to a local proof criteria. A series of intermediary positions exists between those positions. Socio-epistemologists (according to Vinck (1995), Fourez (2001)) advocates that scientific content is determined by a mixture between intra and extra scientific factors. In that sense, we are close to socio-epistemologists. Our philosophical position will be made explicit in section 5.1: it is close to van Fraassen's structural empiricism, but leaving space for social influence in scientific content in the sense that extra-scientific choices are used as initial conditions for the scientific work in quality assessment.

4.2 Terrain presentation

The regulatory and economical context of APECOP was presented in chapter 1, and additional details can be found in appendix E. In the present chapter, we are mainly concerned by the work done by APECOP around ‘‘scenarios’ validation’’. In order to understand the discussions (that we are going to analyse in detail), it is necessary to introduce some of the concepts and scientific tools used in the project.

Despite that concepts evolved during the project, and despite divergence of opinions, the participants had at the beginning some common understanding about models and scenarios. Models were seen as some sort of simplified representation of the reality, by means of equations solved by computer programs. Strictly speaking, what they call "models" were the computer programs. The kind of models used with the FOCUS scenarios are called "numerical" or "deterministic" models (as opposed to statistical models). Scenarios are inputs provided to the models (see section 1.2 for a presentation of the FOCUS-GW scenarios) and rapidly after the project kick-off, they were seen as "representing" something, but *what* and *how* they represent was not clearly defined at the beginning of the project and such uncertainty was at the origin of long discussions that we are going to analyse in the present chapter. Those discussions came into the floor when APECOP started to work on one of its objectives: scenarios' "validation". Scenarios' "validation" was a complicated task because there was no ready-to-use method at the beginning of the project, and because of the uncertainties about how and what the scenarios represent.

In order to "validate" the scenarios, a series of tools were considered necessary, shortly after the beginning of the project: (1) pan-European soil, climate and land-use geographical databases were considered necessary in order to assess if the FOCUS scenarios were in some way representative for Europe; (2) the FOCUS models in order to obtain modelling results from the FOCUS scenarios; (3) a statistical or analytical metamodel[4] in order to extrapolate the results of the FOCUS models to Europe, using the pan-European databases.

As we will see here below, during the project, the method that was chosen for scenarios' "validation" consisted (in short) to use the results of the extrapolation of FOCUS models across Europe to obtain a picture of vulnerability of soils and climate to pesticide leaching. From that picture, it was expected that it is possible to assess if FOCUS scenarios were (or not) representative for "realistic worst cases".

However, having defined that method, APECOP had to face a series of challenges that generated the discussions that we will examine here. Some challenges can be formulated as follows: What are exactly representing the scenarios? (i.e. how to define the scenario's representing

[4]A metamodel is some sort of "model of a model", i.e. a metamodel mimics the behaviour of a full model by extrapolations of model outputs, but having the advantage of being faster and having less input parameters so allowing to use them on big databases with few parameters like the pan-European ones.

concept precisely?) If they represent agriculture, how can we define agriculture, and how to translate that concept in quantitative terms using databases? The definition of the concept of agriculture lead to a series of sub-questions, regarding e.g. the definition of arable lands, the definition of the exact frontiers of the agricultural area represented by the scenarios, etc. Other challenges were: Is it possible to construct a metamodel? Is a metamodel really needed? How can vulnerability be defined? How can a ‘‘validation’’ criteria be defined? Must the result of the validation analysis yield a binary answer, or some more nuanced answer? Is representativeness the only ‘‘validation’’ criterion for scenarios? What is the role of ‘‘validation’’? It it the role of ‘‘validation’’ to improve the scenarios? What if they are ‘‘invalidated’’?.

A question was central during the discussions about FOCUS scenarios: *the definition and mapping of the geographical extension of the scenarios*. This question arose from the fact (see chapter 1) that FOCUS-GW scenarios are claimed to represent ‘‘major agricultural areas’’. However, the scenarios’ narration does not provide any practical definition of such areas. After many discussions an agreement was found, and the result was called ‘‘the FOCUS areas’’.

Another question must be mentioned for a good understanding of the debate: FOCUS scenarios are claimed (FOCUS, 2000) to represent (amongst others) the *90^{th} percentile of vulnerability*. In that sentence, the *90^{th} percentile* (i.e. p_{90}) is a probabilistic concept describing the distribution of a random variable (lets say X). If $x = p_{90}(X)$ then the probability that a random variable X will be less than x is at most 90/100. The 90th percentile is often referred to as a *worst case*.

Vulnerability is a problematic concept, since it supposes that it can be evaluated from the internal characteristics of the system (soil, climate, crops) without having the characteristics of the pesticide. This is solved in FOCUS by stating that: *‘‘the scenarios should describe an overall vulnerability approximating the 90th percentile of all possible situations’’*[5].

[5]However, this can have various interpretations, amongst which two are particularly problematic: (1) it means that for any situation, i.e. for all known and future pesticides, this sentence is true, individually for each pesticide; and (2) it means that the sentence is true when all pesticides are taken together, for all known and future pesticides. Sentence (1) is problematic since it is expected that it will be false for at least one pesticide. Sentence (2) is problematic since it can not be tested because we do not have the population of all future pesticides.

4.3 Method: comprehensive observation and experimentation, rhetoric and actor-networks

The position of the observer-anthropologist in relation to his field is a determining one. It is necessary to try to objectify this position in order to grasp its implications and to make clear what observation is possible and what interaction is involved by virtue of the observer's position in the field. I worked in the APECOP project. In the course of my work, I was asked to contribute to the "validation" of scenarios from the FOCUS-GW group and to this end to develop a scenarios' "validation" methodology. How does one "validate" scenarios without having conceptualised what they are? How does one take into account the politically sensitive aspect of these scenarios? Indeed, economic interests are at stake and they are clearly taken into account both by FOCUS and by the European Commission. Through my laboratory work, I was confronted with the problem of giving an operational definition of what scenarios are.

As we will see, the anthropological intention will lead us to highlight the tense and difficult negotiations between actors that defined the scenarios. In which case, how can I say what scenarios are, if this question has meaning only through negotiation and not as the concern of an isolated individual? This is where the explanation of my position as participating observer becomes most useful: I am one of the parties who negotiated the definition of what scenarios are. My position, like that of the others, contributed to "making the scenarios what they are."

In following the terminology used by Juan (1999) to describe the main methodological types in the social sciences[6], we may say that the present work took place using two distinct methods: comprehensive experimentation and comprehensive observation. Our purpose when using these methods is to identify the way in which researchers modelling practice is used when it interacts with the political field through regulation.

We noted phrases and fragments of conversations heard at APECOP project meetings. These extracts from daily life in the project were gathered without the need for any special interaction on my part. The same technique was used at lectures, congresses and scientific meetings I was able to take part in by virtue of my work.

[6] Comprehensive/explanatory observation, comprehensive/explanatory experimentation.

Finally, I undertook research into written documents, both on paper and in electronic format. To this end, both emails and minutes of meetings were used, as well as documents available, which helped to clarify the political and economic context of the laboratory and the APECOP project we studied.

In addition, various assessment concepts were developed by myself in parallel with the APECOP project (taking into account the sociological analyses presented here) and were then presented to APECOP members. According to Juan (1999), this procedure is a special form of comprehensive experimentation in which the "intervention" element is substantial. It is similar to the techniques of action research or to the sociological intervention of A.Touraine (Touraine, 1978), in which the researcher takes on the goal of group transformation. This poses a number of methodological problems that need to be mentioned; for example, the fact that if change takes place, peer control becomes difficult since the object of study is itself transformed. This problem can be overcome in part by parallel studies of other groups. However, as we will see, my proposals were only partly accepted by the APECOP group, and then only once they had been translated into scientific terms. We may therefore conclude that the intervention has generated reactions, which is its aim, but without profoundly transforming the viewpoints of the actors, which allows us to affirm that the issue of reproducibility is a relatively minor one in this case.

We will present a discussion that took place in the APECOP project. A first and partial interpretation of this material, an interpretation that has an essentially descriptive purpose, is supplied by rhetoric and pragmatics (see appendix C for background). Our basis will be the method described by Doury (1997). This method will allow us to understand some of the issues and to see the kind of argumentation used as well as some of the arguments put forward. This will show how "what does the scenarios represent" is negotiated. This will then serve as the basis for a more global interpretation using actor-network theory (see appendix D for background).

4.4 Results and discussion

During the APECOP project, two questions arose that took an important place in the discussions: "Are the scenarios valid?", and from there "What does the scenarios represent?". In such a context, the relations between the "reality" and the scenarios were sometimes problematic, polemic and subject to interpretation, as it can be seen from the following sentences:

- *There is no consensus on the interpretation of the FOCUS scenarios. This was also noted at the EU modelling workshop in Lisbon*[7]*. Persons have different views (even between former members of the FOCUS Working Group). The definition of FOCUS areas—as used in the APECOP project—is not subscribed by several people; i.e. they deny that the FOCUS scenarios are representative of the realistic worst case within the areas defined by the climatic limits and claim the term is misleading.*(APECOP, 2002)
- *The scenarios were created by FOCUS and we can not take the liberty of interpreting them, we can just use them as they are.* (APECOP project coordinator opinion, April 2000).

The FOCUS scenarios were created with relatively few discussions about their ontology (the object itself they are representing). It seems that the question became far more important when the question of their "validity"'[8] came up with the APECOP project. APECOP leads to the opening of a controversy about what the scenarios are, how can they be interpreted, must they be interpreted, what do they represent and so on.

From a very extensive material that is presented in appendix F and G, we discuss here what is related to the following thematics:

- "Why do we have to validate the scenarios"?
- "Are the scenarios valid?"
- "What does the scenario represent?"
- "How do we validate a scenario"?

The results are analysed in terms of an evolving problematisation[9], following in this Callon (1986) and approaches of actor-network theory.

[7]The Lisbon workshop is a scientific conference about modelling and scenarios. Industry and universities representatives, amongst others, gather there.

[8]Quotes are used in some words in this chapter to signify that we mean the word as used by our informants, as observed in conversations, interactions and so on.

[9]One of the processes of translation, see appendix D

An approach of this questions in terms of interest of the actors involved is possible[10].

4.4.1 Quality components involved

We had at our disposal the APECOP project documents, emails exchanged between project members, and minutes of meetings of the FOCUS Steering Committee. Examination of the chronology and the manner in which the work performed in the APECOP project became known to various actors is of interest for more than one reason. The chronology is as follows:

- September 1999, project paper submitted to the commission;
- April 2000, beginning of APECOP project, financed by DG- RTD[11];
- May 2001, DG- SANCO[12] becomes aware of the project;
- May 2001, the FOCUS Steering Committee, which is part of DG-SANCO but which includes actors from industry as well as representatives of the Commission, becomes aware of the project;
- December 2001, the FOCUS Steering Committee asks for a meeting to be held between APECOP and 5 external experts (from industry, research centres and universities);
- July 2002, meeting between members from the APECOP consortium and industry.

In its formulation, as it appears in the project's paper submitted to the Commission (APECOP, 1999), APECOP proposes to improve the standard models and scenarios used by the Commission through a process of "validation". Scientists, administrative bodies for the regulation of product marketing approval (particularly DG-VI, the future DG-SANCO),

[10]Indeed, those questions, as well as the APECOP project, are located in a broader context: the pesticide regulation and its economical implications for industry. However, we are reluctant to an approach in terms of interests because interests can hardly be observed. Despite our reluctance to perform an analysis in terms of interests, some may want to have hypothesis to explain why the regulators and industry were worried about APECOP, why some "validation" criteria were not accepted or caused many discussions and why the priority was given to representativeness rather than to other criteria in APECOP. Elements that can be used for that are presented in appendix E. It must be said that direct interactions with industry were not the major part of our participant observations. In consequence, appendix E helps also to have a broader panorama, by collecting opinions from a large set of written sources.

[11]European Commission, Directorate General Research.

[12]DG health and consumers.

the pesticide industry and politicians are designated as “end users” or beneficiaries of the products from the APECOP project and therefore as having an interest in its existence. In the terms used by the sociology of translation, the APECOP project was presented to actants as the OPP (obligatory passage point) for each to overcome their specific obstacle in achieving their goals.

This project could not be the object of problematisation without intimate knowledge of the work of FOCUS. Hence, several former FOCUS members have participated in preparation of the project. Furthermore, the FOCUS report (FOCUS, 2000), published by FOCUS-GW to describe its scenarios and how they are constructed, already contains a chapter offering ways of later reducing a collection of uncertainties “in relation to the use of scenarios”. The continuation of FOCUS-GW had therefore already been conceptualised by a number of FOCUS researchers, who subsequently came together again in APECOP: *“we should do more research to reduce uncertainties.”*.

DG-SANCO’s reaction in May 2001 is one of concern regarding the APECOP project:

> *You know, if it comes to FOCUS Steering Committee, I tend to behave like a lion’s mother in the sense that I am very alert about anything which might invalidate the current system*[13,14].

The initiative for the APECOP project took place outside the control of DG-SANCO, and could question the work carried out by FOCUS. DG-SANCO made explicit why scenario’s “invalidation” might be problematic:

> *I want FOCUS to be THE standard for at least 5 more years until the review programme for existing substances is finished. We must not move the target any more.*[15]

It is interesting to note that, despite the substantial number of people who had been or were still members of FOCUS at the time when the

[13]Personnal communication from the EU Commission (DG-SANCO) coordinator of FOCUS.

[14] In that sentence, the “current system” refers mainly to the scenarios.

[15]Personnal communication from the DG-SANCO coordinator of FOCUS.

APECOP project was launched, DG-SANCO – which coordinated FOCUS – was not aware that this new project was beginning. This is even more astonishing given that APECOP's direct study objects were related to the FOCUS products.

As explained in appendix E.1, when the APECOP project started, the Commission experimented delays in its evaluation of pesticide available on the market. Furthermore, many actors already started using the FOCUS scenarios. A possible suggestion to change the approach would take several years to become operational and hamper as such the further implementation of the Directive. This pragmatic aspect of the FOCUS scenarios appears equally clearly in the minutes of meetings of the FOCUS Steering Committee:

> *Don't get lost in the scientific detail, focus on aspects for which concrete proposals can be developed*[16]*.*

For FOCUS and the Commission, it is essential to manage the pesticide registration process in an effective way, since acceptance by industry undoubtedly depends on effectiveness. This hypothesis is supported by, for example, the statements by Racke (2002): each year of delay in regulatory approval is a loss in terms of sales, but also a loss of a year of product life before the patent enters the public domain.

After becoming aware of the project, the FOCUS Steering Committee seemed more reassured:

> *The results of the APECOP project have the potential to improve the FOCUS-GW scenarios in the future, provided that due process is followed.*

The criticism was therefore accepted, if it followed the institutional path of the FOCUS Steering Committee and submitted to a kind of approval in the context of FOCUS and before any possible scientific symposia:

> - December 2001: *The [FOCUS Steering] Committee asked [...] to organise a meeting at an appropriate time between [APECOP] and about 5 external experts (e.g. 3 from industry*

[16] DG-SANCO, in the minutes of the 11th Meeting of the FOCUS Steering Committee, Brussels, December 2002.

> *and 2 from research institutes or universities) to discuss the results and underlying assumptions. It seems crucial to me that we have this meeting before results from the FOCUS scenario validation are presented at workshops or symposia*[17].
> - March 2002: *We need a scientific discussion of the concept of validation before submitting articles to scientific journals. We will therefore meet with industry in July*[18].
> - April 2002: *There was concern that there was rigorous checking that was being done [in APECOP] was consistent with the FOCUS groundwater approach*[19].

The reaction from industry was also one of concern followed by relief, and additionally, industry gave a glimpse of the possibility of financing a project after the end of APECOP, as shown in the minutes of July 2002 meeting:

> *This informal exchange of information has taken away many worries of industry. [...] [Industry] involvement could be both financial and active*[20].

Criticism of the scenarios by APECOP was accepted as it could be accepted without calling into question the existing system. In addition, the line taken by the APECOP project was to say that they wanted to contribute to improve the system. According to the project coordinator:

> *APECOP's work is not to demolish FOCUS's work, but to contribute to improving things*[21].

The chronology we just outlined is indicating that tensions exists between various objectives (i.e. various problematisations) for scenarios: they must be at the same time: useable in a regulatory context, acceptable by industry, institutionalised, and scientifically valid. In fact, in the terms of our chapter 3, those are quality dimensions of the scenarios.

[17] Personal communication from a member of FOCUS-SC and APECOP, December 2001.

[18] APECOP coordinator, personal communication, March 2002.

[19] Minutes of the 10th Meeting of the FOCUS Steering Committee, Brussels, April 2002.

[20] Minutes of the meeting of APECOP with industry, July 2002.

[21] APECOP coordinator, personal communication, November 2002.

Moreover, DG SANCO is here using arguments that neither are juridical or scientific, but rather political. The involved quality dimensions are thus the juridical, political, and scientific.

Institutionalised here means created by committees set up by the Commission, and accepted by the parties who make up these committees (industry, researchers, EU Commission and Member States in the case of the FOCUS Steering Committee). DG-SANCO's reaction illustrates those tensions, and this chronology demonstrates also the tensions to which APECOP is submitted.

4.4.2 Scientists performing interpretations.

One of the purposes of the APECOP project was to estimate the degree of "validity" of FOCUS scenarios. Many of the debates and polemics were thus derived from the question "Are the FOCUS scenarios valid?". A rhetorical analysis of our material, is presented in appendix F. The first step into a rhetoric analysis of a controversy is the identification of two antagonist propositions (the doxa and the antagonist propositions, see appendix C), as well as the rhetorical question[22].The rhetorical question is "Are the FOCUS scenarios valid?".

Starting from that question, two propositions can be distinguished from the recorded participant observation material:

- Proposition 1 (doxa). The FOCUS scenarios are valid, but they have some defects that we must correct; our classical modelling method is the good one to find the defects and improve the scenarios.
- Proposition 2 (antagonist). The scenarios may have more defects than those you expect; other methods are needed to find those defects.

Starting from that distinction, appendix F presents a chronology of the variety of arguments used by our informants.

A first observation from that table is that there where tense discussions around the question "What does the FOCUS scenarios represent'?".

[22]Often, a rhetorical question is defined having the meaning of *erotema*: "to affirm or deny a point strongly by asking it as a question". For example, "Why are you so beautiful?" is likely to be a statement regarding one's opinion of the person addressed rather than a genuine request to know. See http://rhetoric.byu.edu. However, we are following here a different definition: the one used by Doury (1997), see appendix C.

The ‘‘validity’’ of the scenarios is conceived in terms of the quality of the representation. But what do they represent? It became increasingly evident that from various interpretations of the FOCUS narrations and the legal texts resulted various answers.

The role of interpretation in those discussions was also the subject of a polemic. At the beginning of the project, the translation[23] lead to a balance between the actants that was in favour of the idea that APECOP do not ‘‘has the right’’ to interpret the FOCUS scenarios: a ‘‘validation’’ was thought to be possible without interpretation. However the polemic evolved, then there was no consensus any more about what the scenarios represents, and a new translation step lead to the idea that the interpretation of ‘‘what the scenarios represent’’ should be as close as possible to the intention of the authors of the scenarios. APECOP thus resolved to consult those authors, but the result was deceptive, because the consulted authors never reached a consensus about ‘‘what they intended to represent with the scenarios’’. The next translation step lead to the idea that the interpretation of ‘‘what the scenarios represent’’ (i.e. the formulation of their representing concept) was to be made by APECOP itself, documented, and based on existing official texts only (scenarios narrations, legal texts), and not on the basis of the intentions of the scenarios authors.

4.4.3 Interpretation choices

In appendix F, a series of arguments are related to the scenarios, and particularly to what they represent. If we take the same doxa and antagonist propositions, but keep (from our observations in appendix F) only arguments related to ‘‘what the scenarios represent’’, we obtain the table that is in appendix G, showing a fluctuation of the definition of the scenario’s representing concept. We can summarize that table as follows: there where successive choices made between equally scientifically justifiable answers to ‘‘What do the scenarios represent?’’. Those choices have an implication on the representativeness estimation, and thus potentially on the regulatory decisions. Those choices are thus political in essence. Some of the more important choices where:

– It was implicit that a worst case means a 90th percentile. Scientifically speaking, other values could have been chosen, like e.g. a 95th

[23] In the terms of the sociology of the translation, see appendix D.

percentile. That will have implications on the number of pesticides authorized.
- Does the scenarios represent the whole EU or including the tropical areas, or only the European part, geographically speaking? The later alternative was chosen, despite that the legal text (directive 91/414/EEC) requires the whole EU. Per se, this ‘‘invalidates’’ the scenarios, juridically speaking.
- Does the scenarios represent climatic areas or agricultural areas? Different databases are to be used for the two options. Lively discussions in the project lead to the former option, for reasons like ‘‘it was never meant to represent agriculture’’, and thus ‘‘if we chose that, the scenarios will be automatically invalidated, even without doing the calculus’’.
- Having opted for climatic areas, various maps where available, with potential implications on the representativeness estimation. Institutional availability of the databases guided the choice.

The controversy lead to a number of concepts defining the scenarios. Those concepts, once agreed within APECOP, were used for the rest of the project. Moreover the agreed concepts were used as the basis for performing the calculations.The set formed by the agreed concepts is in fact a new definition of the FOCUS scenarios’ representing concept, that resulted from the controversy summarized above, i.e. from power relations and argumentation between various actors. The compilation of the agreed concepts, i.e. the new scenario’s representing concept, is summarised as follows:

> *Scenarios have different aspects, and thus are subject to different criteria, some are of academic interest only, others are more politically (pragmatically) useful. Scenarios are a subject for interpretation, but only the influence of interpretations on registration is interesting (i.e. pragmatically useful). The texts that describes the scenarios must be used for interpretation. Scenarios are subject to political/‘‘diplomatic’’ constraints. Scenarios include pragmatic choices. Scenarios represent vulnerability inside areas defined by arable lands and FOCUS climatic categories (no anthropogenic factor defines the areas). Scenarios have degrees of ‘‘validity’’. Scenarios are also subject to legal and logical constraints, in addition to scientific aspects. Scenarios are standards. Sce-*

> *narios are supposed to represent 90th percentile vulnerability (and that must be checked through "validation"). Scenarios have incoherency in some aspects.*

It worth noting that such a concept was never written as such, this list is made collecting sentences from various periods of the project and from various documents.

Such a definition of the scenario's representing concept reflects the concepts that were stabilized into the APECOP project, but not necessarily for other actants. Our concept of "stabilization" here is in-line with the characteristics of an agreement as defined by Michael Lynch (cited by Vinck (1995)): it is immanent to the situation, it is local, it is not dependent on having the individuals agreeing (or not) "in their hearths", an agreement is a part of a collective action. The agreement is produced when there are no more signs of disagreement, and the linguistic interaction ceases or changes its subject. In consequence, the definition given here above does not imply that everybody (or the author of the present work) agrees with it; simply it is the definition as showed by the state of the discussion into APECOP in 2003 (i.e. at its end), taking into account the power relations[24]. The fact that such an agreement was partly not written gives it less power to have consequences. It remains true that part of the agreement, when not written and not used inside APECOP, can be interpreted as a rhetorical strategy to avoid consequences of a stronger agreement. However, indications exists that at least part of what was stabilized into APECOP continued after it: e.g. in 2008, a new FOCUS groundwater group took the position that FOCUS scenarios represent vulnerability inside climatic areas (see section E.1). Moreover, the existence of the present work, partially done after the APECOP project finished, is a form of stabilization in the long term accepted by the University.

It clearly appears to us that the observed fluctuation of the definition of scenario's representing concept cannot be explained by a dynamic presumed to be internal to a closed "scientific field", to use Bourdieu's terminology (Bourdieu, 2001). In fact, arguments such as the need for standardisation of scenarios and the absence of operational regulatory

[24] Such power relations could be the subject of an analysis itself, starting from the kind of arguments used (pragmatics analysis of language), e.g. the use of authority arguments.

alternatives weighed heavily (cf. appendix F) in determining what scenarios represent. We might even add: the wish to maintain political leeway, the wish to "help rather than demolish", a wish for continuity in relation to the scenario's development work in which several members of APECOP had been involved. It even appeared that prospects of publication ("validation by agricultural areas is not scientifically interesting") contributed to determining what the scenarios represents or not (in the end, they did not represent agricultural areas).

4.4.4 The tools to "validate"

The question about the tools to "validate" turned out to be a choice between using a deterministic model, an analytical metamodel or a statistical metamodel. In fine, there remained two options: between the deterministic model and the statistical metamodel. Piñeros Garcet (2003) presents the translations that conducted to such an alternative. There where scientific advantages and inconvenients with the two solutions: the metamodel allows to take into account the intracell variability in soil maps, the deterministic model do not, but the metamodel introduces an additional fitting error. Also in favour of the metamodel was its direct use of soil maps without introducing an additional error coming from the linkage required by the deterministic model, between the soil map and estimated soil profiles.

The polemic resulted in the consensus that both methods should be used and compared. Interestingly enough, the two methods where supported by two different laboratories, who had thus interests in developing their respective methods.

The burden of the proof was not equal between the two models, in part because the metamodel was new, but maybe also because the metamodel results lead to the conclusion that the FOCUS scenarios representativeness was rather poor, while the numerical model results where more gentle with FOCUS scenarios. The choice between the two models may have implications on a revision of the scenarios, and thus on the regulatory decisions. Such a choice had thus political implications.

4.4.5 Interpretation in quality assessments, and the links between quality components

At this stage of the demonstration, we can conclude that interpretation is a constituent of quality assessment. Moreover, interpretation is involving extra-scientific arguments to chose between equally scientifically justifiable interpretation options. Those extra-scientific arguments are related in our case study mainly to the political and juridical fields. This is due to the fact that the scenarios are used also in political and juridical contexts.

If the quality is assessed, the result is linked to choices pertaining to the political and juridical quality components. If other juridical or political choices are made, another representation estimation is obtained. This fact is not due to the particular case of FOCUS, with a functional representativeness: choices are to be made even in the case of a measurement representativeness.

There are thus links between quality components. This is shown in fig. 4.3 and fig. 4.4. Figure 4.3 shows how a series of choices conducted to two different assessment results in APECOP. The choices are the result of the links between the scientific, juridical and political quality components. Figure 4.4 presents a thought experiment for investigating if other choices are possible and their consequences. Figure 4.4 shows that either we could arrive to a third representativeness estimation (termed X in the figure) or even to the fact that the representativeness can not be estimated. The later alternative is for the case that the only argument considered is experimental: there is no way to upscale it up to a percentile at the continental scale.

Our anthropological observations show thus that a scientific quality assessment of scenarios requires a series of interpretation choices as an initial condition. Those interpretations are either juridical, politic or scientific in nature. In consequence, the links between quality components must be taken into account when assessing scenarios.

Figure 4.3: Interpretation choices and links between quality components in APECOP. Horizontally in the figure are the successive interpretation choices that where made in APECOP, to arrive to two alternative representativeness estimations (called “validation” by APECOP). Vertically, in red, are the links between quality components that explains those choices.

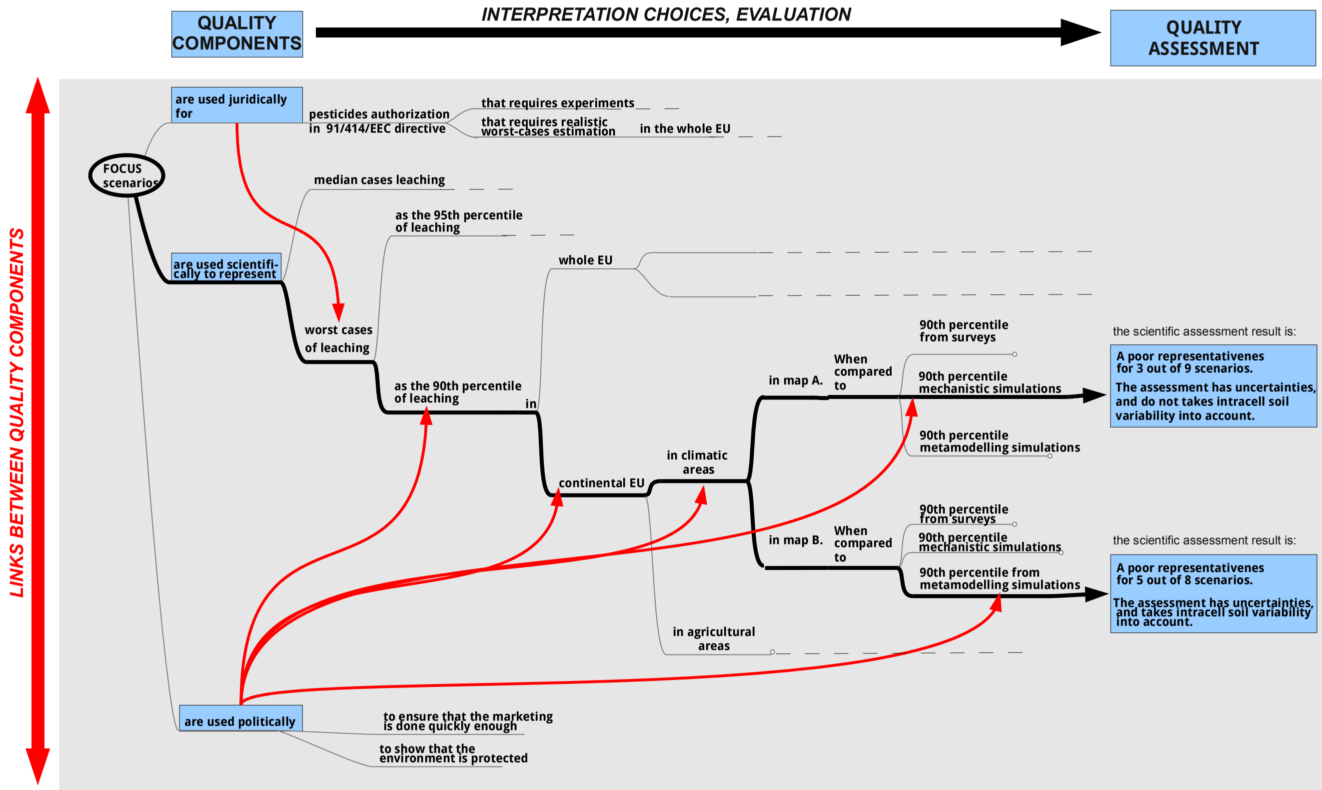
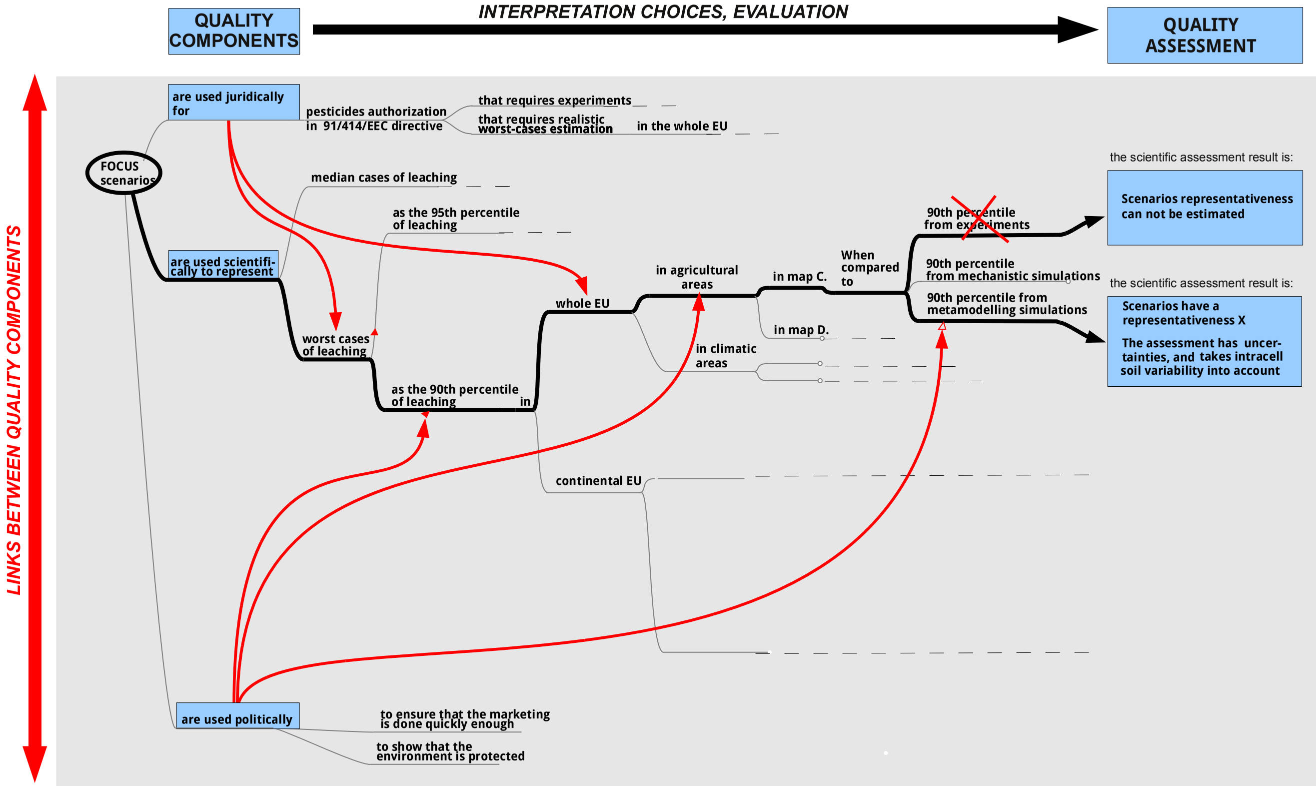


Figure 4.4: Thought experiment: what other choices could have been made in APECOP. Other equally scientifically justifiable choices could have been made, that leads to two different choices than those that occurred in APECOP. Horizontally in the figure are the successive interpretation choices that we make, to arrive to two alternative representativeness estimations. Vertically, in red, are the links between quality components that explains those choices. In one of the alternatives, the assessment is not possible.



4.4.6 Interpreted frontier objects

The identity of something appears to be intimately linked to the use one makes of it. To use Latour's words, the action name of an actant is a result of the tests it undergoes: it is by seeking to make it undergo a process of "validation" that a scenario is defined. In our case, the action name of the scenarios is "what they represent", i.e. their representing concept. But when it is a matter of conceptual practice (such as the development of scenarios), what can be the test? We might follow (Pickering, 1992), who sees conceptual practice as a production of associations that allow the elements to hold together. This production is seen by Pickering (1992) as dynamic (and dialectic, it seems to us): first, a movement of active choices and then forced movements resulting in a surrender before the constraints imposed by the system already accepted. This is also found in our observations on the subject of the development of the scenario's representing concept: each problematisation of what scenarios represents leads the actants to a new system of logical constraints.

However, we observed that actants from the worlds of politics or economics intervened very closely in the tests that scenario's underwent: the scenarios did not undergo a conceptual problematisation alone. In an interpretation in terms of actants, a technique produces heterogeneous combinations of statements, knowledge, techniques, administrative procedures, databases, etc. This is what Latour and Callon call a socio-technical network. When one of these networks is mobilised and acts as an actant, Callon calls it the actor-network. When this network has been sufficiently robust, it has mobilised in order to produce a kind of scenario's "validation" accepted and desired by all actants. In this interpretation, scientific expertise consists not in speaking the "true word" but rather in establishing relations and constructing actants. What scenarios represent and the ways of "validating" them, is constructed by the *articulation* (Latour, 2001) between different programmes or problematisations of all the actants.

We have been able to observe a fluctuation in the definition of FOCUS scenarios representing concept (as presented in appendix G), i.e. a fluctuation in the interpretation of what the scenarios represents. On the one hand those fluctuations correspond to successive stages of translation. If new actors enter the network (for example, a "validation" technique), what the scenarios represents changes under the effect of negotiations and struggles. But, on the other hand, the scenarios have appeared as profoundly polysemic: this is one of the typical characteristics of "border objects"[25]: they make it possible for heterogeneous actants and disjointed networks to co-exist and become linked. *It is therefore due to the existence of a border object such as scenarios that industry, administration and scientists were able to come together.* The fact that scenarios are interpreted border objects is what makes that their different quality components are linked.

A border object exists only as a function of networks coming into contact; hence the fact that the identity of such an actant (here the representation concept) largely depends on these networks. From this, we can deduce that the introduction of a new problematisation of the scenarios has induced by a multicriteria quality assessment as the one we advocate, has little chance of success if it is not carried out with the support of a network which would force/negotiate access to the border object and thereby enter into contact with the other networks already organised around this object. The possibility of its success depended on this search for allies.

4.5 Chapter conclusion

This chapter presented an anthropological approach to the study of simulation scenarios, a domain that has not been extensively studied so far, and completes the study of the representation in modelling using symbolic systems presented in section 3.2.1. The anthropological terrain was the persons and institutions involved in the APECOP project about models and scenarios "validation", presented in chapter 1.

A controversy took place in order to determine what the scenarios are supposed to represent, and those negotiations involved interpretation of the scenarios narrations and parameters. Negotiations also took place around the methods and criteria for determining what a "valid" representation is. The practice of scenarios' "validation" involves thus the definition of what a scenario should represent (that we term here the "representing concept" of a scenario), through a process of polemic negotiations that leads to a scenario's interpretation. Different persons have a different idea about what scenarios are representing, "the true thing that is represented" is searched using various categories of arguments including extra-scientific considerations. Some of our scientists believe that scenarios are entirely arbitrary.

The role of the legislative, political and economic actants on the social translation of the scenarios' interpretation was highlighted. The social processes observed during the interpretation of the scenarios were described here in terms of the rhetoric and translation. That analysis leads to the conclusion that the question *"are the scenarios valid?"* is not independent from three others: *"what are the scenarios representing?"*, *"what methods can be used to validate the scenarios?"*, and *"what are the criteria for a valid scenario?"*. The existence of that link opens the scenarios' "validation" to the influence of a variety of actants like the European Commission, the persons who constructed the scenarios, EU Directives, political considerations, industry, competing models, maps of tropical areas of the EU, stability pragmatism and standardisation needs for the pesticide registration, available databases, strategic considerations regarding scientific publications, and so on.

A series of successive representing concepts are observed through the controversy, until a stabilisation occurs. Extra-scientific arguments are used in those negotiations, and arguments from the political and economical fields were taken into account to arrive at the final scenario representing concept that was used to test the "validity" of the scenarios. Through our translation analysis, it appears that scenarios are actants with a very fluctuating identity, the identity being their representing concept definition. The representing concept results from the translation. This leads us to describe scenarios in risk assessment as a border object. The scenario representing concept viewed as a border object, is able to link quality components through the process of interpretation that results from multiple translations. They are polysemous in order to play their role: to be a communication point between networks using different languages.

This view is in line with the views of N.Bouleau:

> *[...] it is not evident that the qualities necessary for a description of reality in an attempt at rationality, objectivity and universality, are the same as those we might expect from a language of communication between actors, whether partners or adversaries, in the complexity of the social game. It seems that such language involves intervention from fields of knowledge of a different nature, contaminated because they are useful and perhaps also useful because they are contaminated. [. . .] permanent presence of negotiation, [. . .] the in-*

> *volvement of different actors through interdisciplinary practice. [...] the language vehicle of these fields of knowledge [...] proves to be semi-artificial: a mixture of vernacular language and of a key idea - that of the model [...] which itself uses the sciences [...]* (In Nouvel (2002))

Our description of scenarios as a border object is also in line with Howes (2001a) who sees the decision support systems as "a consortium of individuals using technology as a context within which they seek solutions".

Our anthropological participant observation shows thus a contradiction between a concept of scenarios seen as "representation of", and "distance to", a thing in itself, versus the scenarios practice which shows that the scenarios representing concept and thus the referent of the representation is fluctuating according to the evolution of a controversy through social process. Such a contradiction must be solved if a criterion for representativeness quality assessment is to be proposed.

We would like to say however that our anthropological conclusions in relation to scenarios cannot be applied indiscriminately to every type of scenario in every situation. Hence, scenarios in the context of more basic research rather than a regulatory context may involve other actants.

In striking the balance of the analyses we have performed, it seems that there might yet be much to extract from the ethnographic material gathered. Thus, it would be possible to undertake a deeper rhetorical analysis in order to show the influence of enunciation situations on the type of argument used (for example: in which situations do we encounter arguments from authority?). It would be possible to give a better description of the conversational territories, or to deal in more detail with the kind of legitimacy that speakers participating in the argument claim for themselves. These descriptions would allow us to build more bridges between the sociology of translation and rhetoric, which we believe complement each other, and from there to acquire more elements about the links between components.

We might also make more use of the tools developed in order to track innovations: for example, using the many numerical indices that exist, such as (Latour et al. (1991), translated in Latour et al. (2007)) the indices of negotiation, of yield, of solidity. This would allow us to understand better why, for example, certain scenarios validation criteria received more attention or investment than others, or why interpretation and translation lead to a given scenario concept and not to other equally justifiable interpretations.

It could be said that there is a symmetry problem in the work presented in this chapter, because industry and regulators are only seen here through what they said in specific contexts (applied research projects, regulatory and scientific conferences, etc), and were not confronted to the conclusions of our analysis. However, once this question arose, we gave the opportunity to industry to answer to direct questions after having presented the objectives of the present work, but the results were deceptive (we got answers but not to our questions). We believe that the present work shows that a collection of a large amount of sources combined with a detailed anthropological analysis of a particular project gives more useful material. Along the same lines, we could have made a return to the field by presenting a sketch of our work to the personnel of the GERU laboratory that is presented in chapter 3. The interaction thus stimulated would certainly not have lacked interest and could have then been integrated into our analysis. However, this return to the anthropological field is in part done here through the fact that the head of the GERU laboratory is closely involved in the present work.

It must be noted that other theoretical approaches to our ethnographic material are possible, like an analysis that sees participants in the APECOP and FOCUS groups as bearers of interests. Some elements towards such an analysis are presented in appendix E. An analysis in terms of fields Bourdieu (2001) could also be of theoretical interest. This would involve, for example, the problem of defining communication between fields, which Bourdieu himself sought to consider. While those analysis may give additional insights about the relations between science and society, the results obtained here are sufficient to demonstrate that (1) a concept of representation in terms of ‘‘distance to a thing in itself’’ can not provide a basis for a criteria of representativeness quality, (2) because interpretation is involved in the assessment and thus requires an interpreter, i.e. the intervention of a subject, and (3) that interpretation choices are the way through which quality components are linked.

CHAPTER 5

The role of interpretation in quality assessment

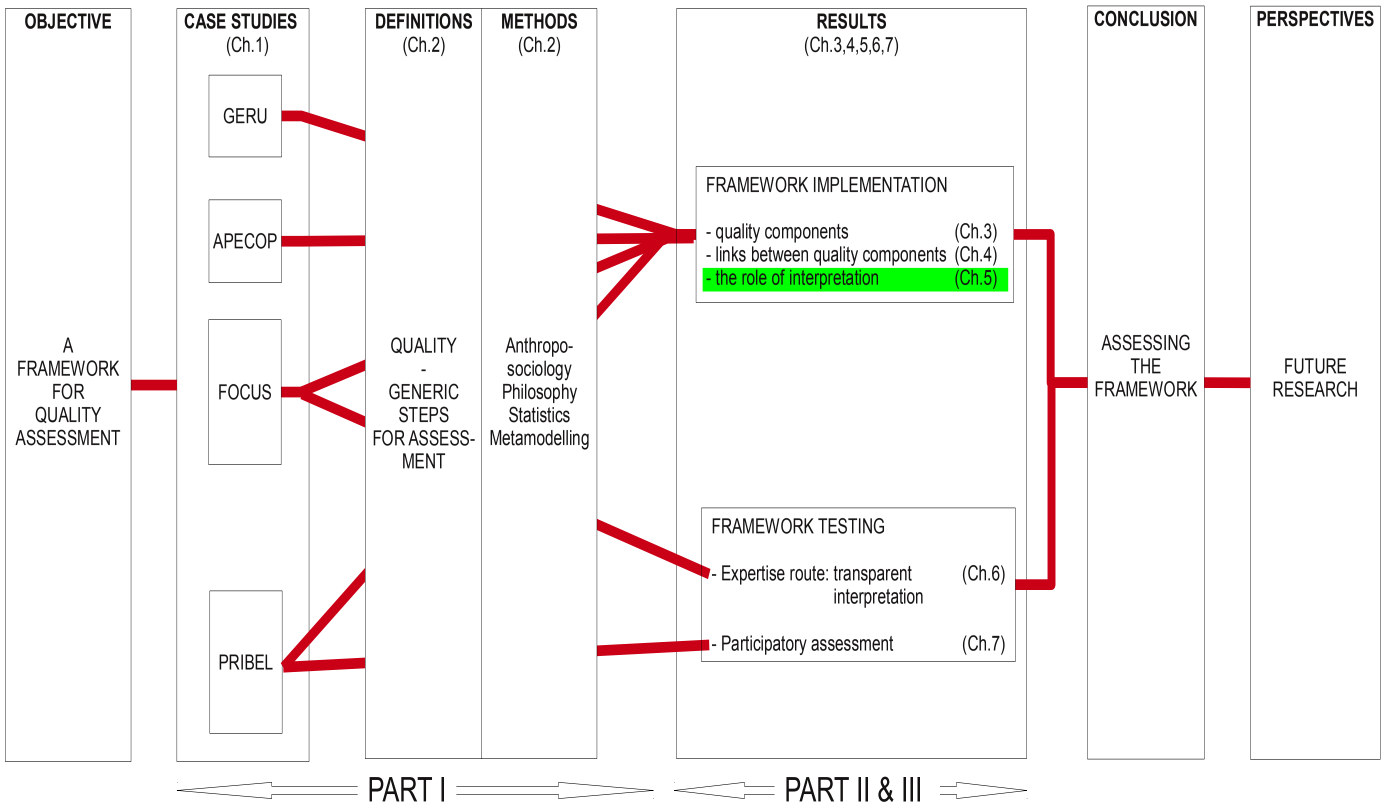
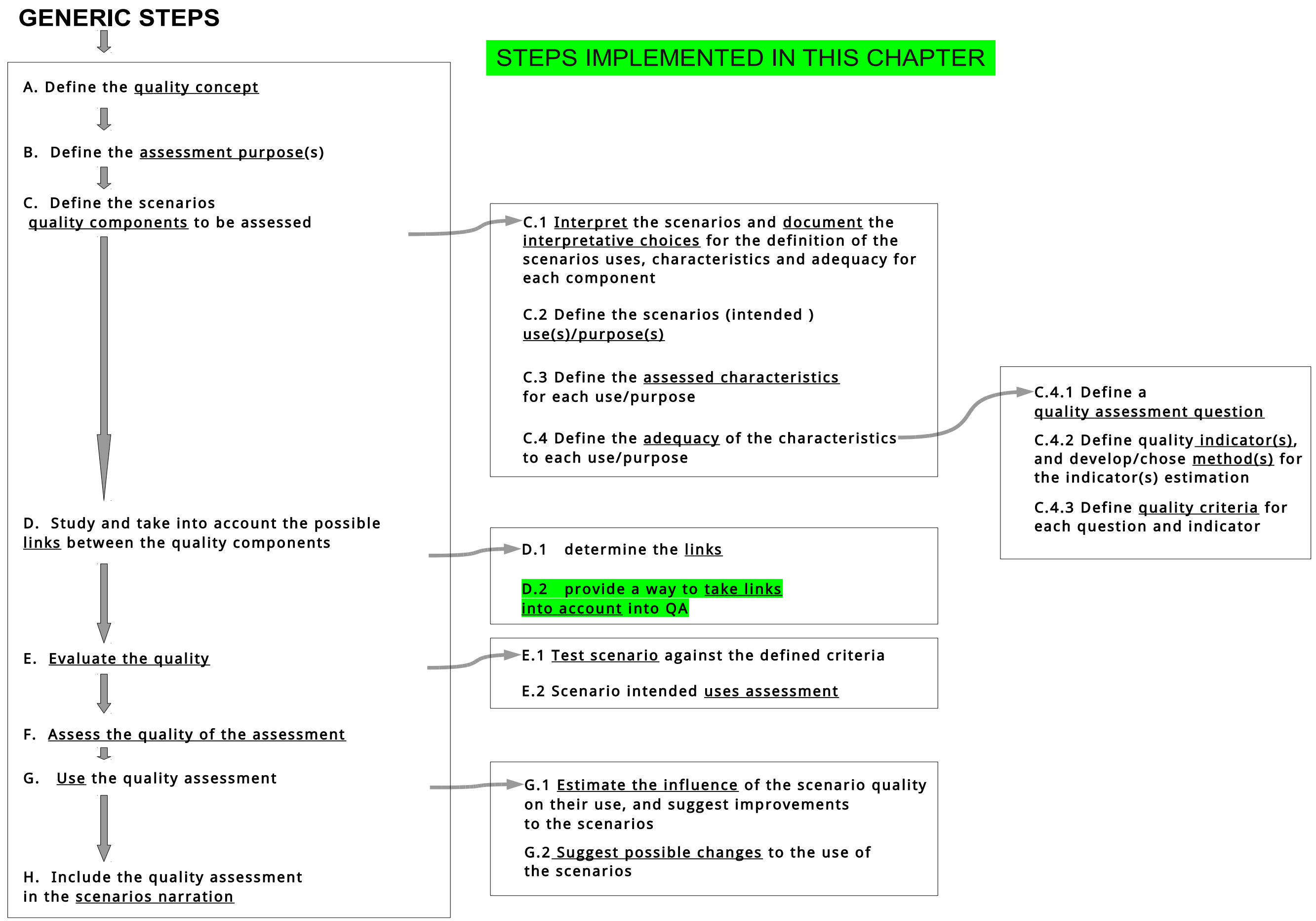


Figure 5.1: Reading guide of chapter 5. In chapter 3, various quality components where defined, and in chapter 4, the links between those quality components where analysed. This chapter implements the generic steps for quality assessment with the concept of interpretation, in order to take into account, in assessments, the links between quality components.

Figure 5.2: Reading guide: contribution of chapter 5 to the implementation of the quality assessment framework generic steps. This chapter implements the generic steps with the concept of interpretation, in order to take into account, in assessments, the links between quality components. This is a contribution to step D.2



Summary

This chapter implements the generic steps for quality assessment with the concept of interpretation, in order to take into account, in assessments, the links between quality components.

As shown in chapter 4, interpretation choices between equally scientifically justifiable options are the result of a social translation. In this chapter we present a philosophical view about those links. Our concept of scenarios representation allows room : (1) for a constructed interpretation of the scenarios (through translation involving actants from the political, economical, and scientific fields) into a scenario representing concept; (2) for the establishment of a referential link between the scenario representing concept and the phenomena; (3) and, if needed, for the establishment of a link between the scenario representing concept and a mathematical structure that may provide access to the empirical adequacy of the representation, and from there to the phenomena.

The role that interpretation should have in the quality assessment framework is made explicit in terms of hermeneutics, criteria and norms of interpretation, transparency, and scenarios narration. Interpretation is present in all the proposed steps for quality assessment, and for all the studied components. Interpretation of a text is not the action of finding what the author had in mind, but a live performance done by the interpreter taking interpretants (i.e. the results of other interpretations) into account. If interpretation is the basis of quality assessment, then the quality of the interpretation must be evaluated. Some criteria were listed for that, and they need of the interpreter's judgement and experience. Moreover, two different but complementary routes for taking interpretation into account are distinguished: the expertise route, and the participatory route. The expert route should at least consider that various interpretations are possible and should thus transparently document the interpretation choices that are made. The participatory route opens the interpretation to citizens participation, which may lead to the mapping of a range of possible interpretations that the expert can take into account when performing its assessment of the scientific component.

5.1 Linking the quality assessment components through interpretation. A philosophical approach.

This section presents the development of a holistic concept of representation using simulation scenarios that is philosophically based, takes into account the anthropological observations of chapter 4 into account, and is designed to be used in the scientific practice through the use of mathematical structures.

Such a view is the basis for proposing ways to take into account the fact that interpretation is present in quality assessments as empirically shown by our anthropological observations in chapter 4. As we shall see, this leads us to have a multidisciplinary approach, with elements from mathematics, sociology, anthropology and philosophy.

Why should we be interested in philosophical considerations about representation? As shown in chapter 3, our informants (i.e. scientists) are making use of a representation concept when expressing opinions about scenarios and models. Moreover, "representation" is a widely shared concept amongst scientists as shown in scientific literature[1]. Once the representation concept is to be formalized, we are entering in the field of philosophy. Based on a review of various concepts of representation, we develop a formalized concept of representation for the case of the simulation scenarios. Our main interest in formalizing comes from the observation (in chapter 4) of a contradiction between a concept of scenarios seen as "representation of", and "distance to", a thing in itself, versus the scenarios practice which shows that the referent of the representation is fluctuating according to the evolution of a controversy and a translation process. Such a contradiction must be solved if a criterion for representativeness quality assessment is to be proposed.

In general terms, representation is a generic term to designate conceptions that, in some way, deal with a relation between a subjectivity or consciousness, and an objectivity or object(Barsotti, 2001). Representation can be understood in various ways in the current language: the act of represent something, to have a representation of something in our consciousness or mind, or the representor with which the act is performed (either in mind or physical). Often, representation (Utaker, 2002) in modelling is defined in terms of the relation between a model

[1] As shown by a simple search of that term in current scientific literature databases.

and an object, leaving the designer (i.e. the subject, the consciousness) out of such a relation as an impartial observer. However, from the philosophical point of view, scenarios and models are complex and a view that conceives them as simplified representations has been criticised in various aspects. A reflexion on representation in modelling leads to fundamental philosophical questions about objectivity versus subjectivity, the possibility of experimental validation of models and scenarios, the relations between theory, models and measurements. For those not familiar with some of the philosophical background used here, appendix H presents an overview of the representation concept in philosophy.

Representation does not imply that the represented is a pre-existing object, but it is however true that many common conceptions of representation includes pre-existing objects. While representation is always representation of “something”, it can be seen in appendix H that in philosophy, the “thing in itself” does not necessarily pre-exists or even exists, and also does not necessarily refers to “real entities found in the nature” as conceived by positivists. Moreover, some philosophers deny the very possibility of a representation (see examples in Barsotti (2001), and in appendix H).

Various philosophical conceptions are grouped here under the umbrella of representation: Husserl’s intentionality[2] (Husserl, 1976), van Fraassen’s appearances, indexicality, measurements, empirical structures (van Fraassen, 2008), Latour’s circulating reference (Latour, 2001), as well as the concept of individuals in Diday’s symbolic data (Bock and Diday, 2000)[3]. Such a grouping is justified by the fact that they all, in some way, deal with a relation between a subjectivity or consciousness, and an objectivity or object. Representation is a controversial subject in

[2]It is necessary to distinguish between intention, intension and intentionality. Intention is referring to the purpose, aim or goal, while intentionality is central in analytic philosophy and phenomenology: intentionality is the power of minds to be about, to represent, to be oriented towards or to stand for, objects, things, properties and states of affairs. Intentional objects may be physical objects, but also *abstracta* (e.g. numbers), *ficta* (e.g. Santa Claus) or *impossibilia* (such as the square circles). Intentions are a particular type of intentional phenomena (others are e.g. believes, desires, perceptions). Intension is in contrast encountered in logic and sets theory. In the sets theory, the intension is a definition of a set by mentioning a defining property, while the extension is the definition of a set by enumerating its members. See Weisstein (2008a), Weisstein (2008b), Jacob (2010), Lecourt (2003), Makins (1991). See also intentionality in appendix H.

[3]See more about Husserl, van Fraassen, Latour, Diday in appendix H.

philosophy, and many debates took and takes place around it. It is central in philosophy since at least the ancient Greece, as well as in present philosophy of science, as shown by the recent book ‘‘Scientific representation: paradoxes of perspective’’ by Baas C. van Fraassen, one of the most influential contemporary philosophers of science (van Fraassen, 2008). Phenomenology and hermeneutics also contribute significantly to the debate (see e.g.Kockelmans (1997)).

The goal of the conceptual view of representation using simulation scenarios that we are proposing is to be used as the basis for a representativeness quality assessment criteria. The requirements we put on the design of such a view are that it must take into account (1) the results of measurements and experimentation if available; (2) the observation that choices, interpretation and negotiation of scenarios takes place, along with the social and intersubjective aspects of it (chapter 4); (3) the role of the social, economical and political fields on representation using scenarios;(4) the existence of the so-called ‘‘virtual scenarios’’ (e.g. FOCUS (2000)) in which it is explicit that they do not represent existing fields; (5) the existence of scenarios that are not constructed from measurements; (6) the need for a mathematical quantification of representativeness; (7) the need to obtain a concept that is useful in scientific practice (and thus not only a kind of meta-definition).

We will introduce in our concept aspects of van Fraassen’s constructive empiricism and its notions of *empirical adequacy*, *indexicality* and *use* of a representation (van Fraassen, 2008), adapted to the representation using simulation scenarios. Two mathematical structures will be compared for adequacy, one of them may have empirical content if needed when the empirical adequacy is to be assessed. Regarding the negotiation aspects, we will take it into account through the *actor-network theory*. *Hermeneutics*, *indexicality* and *use* notions will allow us to define the way the reference of the scenarios is defined by *subjects performing a representation*.

All concepts necessary for presenting our proposal can not be defined at once, so we are building on top of view 1 presented in 3.2.2 to introduce increasingly complex views about ‘‘how scenarios represent’’ (views 2 and 3) to arrive at our full proposal. Then, some aspects of our full proposal are discussed more in depth, i.e. the role interpretation, intentionality, empirical adequacy, indexicality, social translation, use, and measurement are playing in our representation concept. Additionally, the so-called ‘‘virtual scenarios’’ are elucidated in philosophical terms.

5.1.1 Scenarios represent by having a scenario representing concept interpreted from narrations

In the View 1 about representation, presented in chapter 3, scenario's intension SI is a mathematical structure. It must be recalled[4] from chapter 3 that the mathematical structure SI is used to find in a symbolic data array $\underline{X}$ (a database or a data model) the descriptive scenario's extension DSE, using a description vector d and a mathematical relation R. We will build here over View 1, by taking into account the results of chapter 4 that showed that a process of interpretation of the scenarios' representing concept is taking place when the scenarios are used. Such process leads to the mathematical structure SI, derived from scenarios' narrations and parameters, by interpretation: this is missing in view 1.

Another thing to note in view 1 is that ESE is termed "enumerative" to highlight that it enumerates the objects ω_i and does not refer directly to the phenomena, this link is also missing in view 1. As contended by van Fraassen (2008), nothing represents without the intervention of a *user* that takes something to represent something else, as thus or so. van Fraassen states that a representation needs an essentially indexical[5] operation, that can not be done without a user intervention:

> *Is there something that I could know to be the case, and which is not expressed by a proposition that could be part of some scientific theory? The answer is YES: something expressed only by an indexical proposition. For to use a theory or model, to base predictions on it, we have to locate ourselves with respect to it* (van Fraassen, 2008).

[4]We recall here the significance of some of the symbols used in View 1: SI: scenario's intension: the scenario's intension is a symbolic object; ω_i: an object: the symbol used to indexically locate (represent) a phenomenon in the logical space of the scenarios; ϕ: the description domain: the space of all the possible levels, alternatives, numbers of the symbolic variables; $\underline{X}$: the symbolic data array: the database (or data model) available for the descriptions; ESE: enumerative scenario extension: is the set of the objects that are represented by SI and is termed "enumerative" to highlight that it enumerates the objects' symbols (the i^{th}) and does not refer directly to the phenomena; a: membership function: allows either to decide which individuals ω are members of ESE or to estimate the degree of membership of a given ω_i to ESE; R: a mathematical relation (e.g. $\{<, >\}$); d: a description vector (e.g. $\{25°C, 100\ mm\ mean\ annual\ rainfall\}$); SI is defined by a,R and d. See p. 58 and thereafter, as well as fig. 3.3, and the list of symbols p.339.

[5]An indexical term is (Makins, 1991) in linguistics and logics a term whose reference depends on the context of utterance, such as I, here, now, you, tomorrow.

The subject is thus introduced in a triangle with the representor and the represented. van Fraassen (2008) gives an example: a sentence like *''the exponential function represents [...] the growth of the investigated bacterial colony''* receives its reference from the context in which it is used.

Thus, according to us, the link between the objects ω_i and the phenomena is an indexical operation. This is also the case with the choice of a description domain ϕ (the space of all the possible levels, alternatives, numbers of the symbolic variables that describes the objects ω_i) and a symbolic data array $\underline{X}$ inside ϕ. Such a choice is an indexical operation that locates a scenario into a logical space. The same is true for the description (representation) of a scenario as a given SI. Moreover, all those operations that now we term as indexical needs an interpretation as shown in chapter 4.

In consequence, we have to introduce *use, indexicality, and interpretation* in our scenarios' view: this is done in fig. 5.3. Interpretation of the scenarios allows to define an *intension* and from there the correspondent extensions.

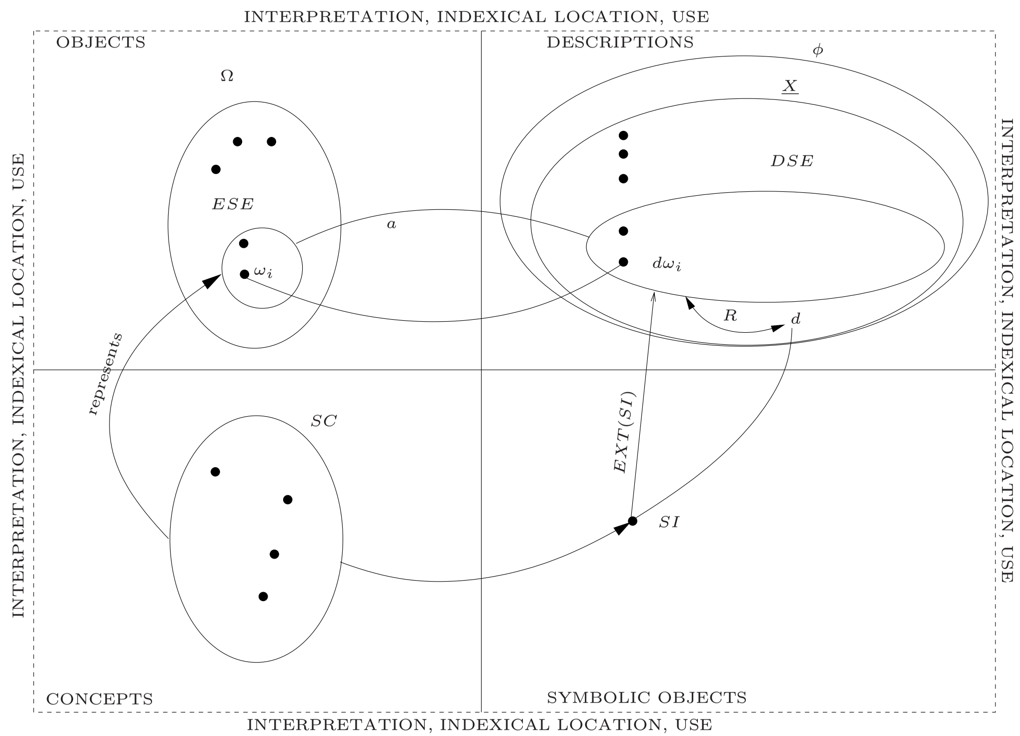


Figure 5.3: *SC, the representing concept of a scenario,* represents *by having an enumerative extension ESE (a table of symbols can be found p.339).* The symbolic objects framework (inside the dotted line) is having a representative function through a use, an indexical location and an interpretation. The dotted line indicates that interpretation, use and indexical location takes place in various places in the figure (e.g., the definition of ESE is an indexical operation that needs the interpretation of the scenario's representing concept into a scenario's intension and description).

The interpretation of the scenario's representing concept SC is made using scenarios' narrations, which are the texts that often comes along with the scenarios to describe them (see chapter 1). Scenario' description d is derived from the parameters values of the scenario, possibly transformed (interpreted) into the logical space ϕ.

How is such a framework applicable to FOCUS, and PRIBEL scenarios? A combination of choices, expert judgement and measurements in various proportions is used when constructing the FOCUS scenarios narration, as well as their description vector d and $\underline{X}$, the database (or data model) available for the descriptions. PRIBEL scenarios have a greater proportion of measurement in their construction, compared to FOCUS scenarios. d in the FOCUS scenarios are constructed by expert judgement, and their $\underline{X}$ is (for their geographical component) a hand-drawn sketch map of EU climatic areas.

Another difference between the FOCUS and PRIBEL scenarios is that FOCUS scenarios are not meant to represent existing fields: their reference is given by the way they are used. They represent realistic worst cases of leaching, *when they are used* with the FOCUS models, for a particular pesticide, and more specifically, they represent the 90$^{\text{th}}$ percentile of leaching inside agro-climatic areas. In consequence, FOCUS scenarios represent through a geographical and a leaching component, with ϕ being (for the leaching component) the set of all possible leaching values for a given pesticide. Another thing that must be emphasized, is that the pesticide characteristics are not considered as being part of d, this is also determined by the particular use context of those scenarios (i.e. the test of pesticide by modelling simulations) which makes that pesticide characteristics are considered as variables.

Integrating all those new elements with our view 1 leads to view 2, presented in fig. 5.4. View 2 integrates a concept of scenarios as choices and/or as measured parameters, with an interpretation view of the way a scenario represents. Interpretation is seen as an intersubjective process. Our view of scenarios brings interpretation into play as an integral part of the intension definition. In consequence, there is more than one intension per scenario, the intension is not thus an essence characteristic of a scenario but rather the result of a series of operations to re-start, to perform, each once the extension of a scenario is sought.

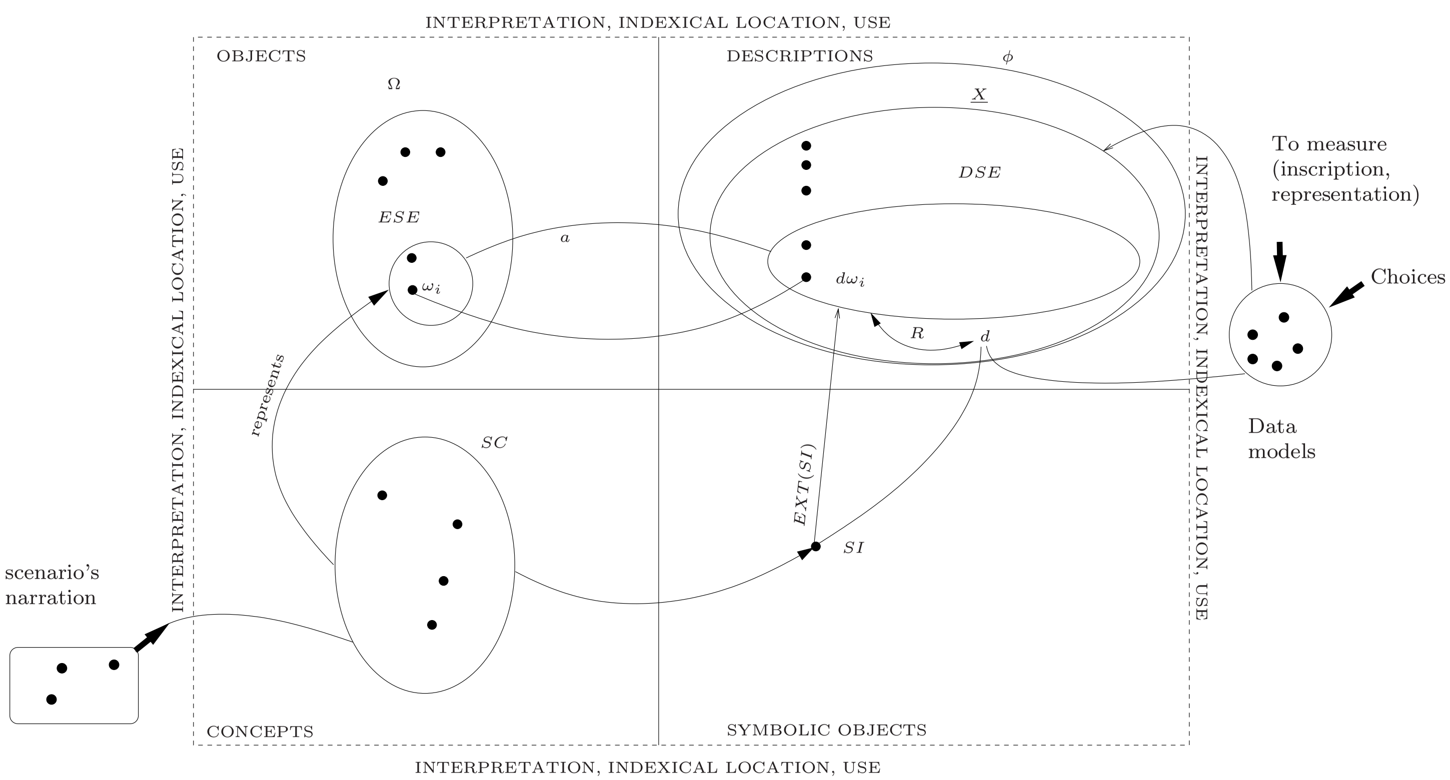


Figure 5.4: *Scenario's view 2: SC (the representing concept of a scenario)* represents *by having an enumerative extension ESE (a table of symbols can be found p.339). The definition of ESE is an indexical operation that needs the interpretation of the scenario's representing concept into a scenario's intension and description.* Interpretation is performed using scenarios' narrations, data or surface models and choices. The dotted line indicates that interpretation and indexical location takes place in more than one place in the figure.

5.1.2 Scenarios represent by having an interpretation resulting from a translation

If we come back to the conclusions of our chapter 4, measurements (data models), choices, narrations can be considered as actants in the actor-network theory[6], along with other actants, actants coming e.g. from the scientific, economic, or regulatory fields. Those various actants enter into a translation process in order to determine the scenario's interpretation. Integrating this with our View 2 leads us to our View 3, presented in fig. 5.5.

Such a description is in itself a representation of the functioning of a scenario as representor. Our aim (i.e. the use we make of such a representation) is to make such a description as general as possible in a theoretical use context. We are however conditioned in our performance by our first practical use context: the validation of scenarios (chapter 6). To asses the generality of such a definition, we confront it, here after, to other scenarios' uses: FOCUS scenarios as used by industry, as well as PRIBEL scenarios' use.

5.1.3 Scenarios "constructed from" vs "used to"

We saw in chapter 1 that PRIBEL scenarios are different in comparison to FOCUS scenarios: FOCUS scenarios are claimed to not represent the characteristics of specific fields. They represent the 90th percentile of leaching[7] *when used* with FOCUS models. Moreover, the FOCUS scenarios parameters are mainly constructed from expert judgement, because databases at the EU level were not available at that time. In the case of PRIBEL scenarios, we are confronted with tables of numbers that comes from procedures that can be assimilated to the results of measurements (van Fraassens's "data models"). As an example, PRIBEL scenarios involves data coming from agricultural practices surveys, or soil analysis. PRIBEL scenarios are used for picturing in van Fraassen's sense[8].

We have thus scenarios constructed from measurements or expert judgment, and then scenarios that represent through picturing or through the results of a modelling exercise.

---

[6]Actor-network theory: see appendix D.

[7]See chapter 1 for a complete narration of what FOCUS scenarios represent.

[8]Picturing is selective likeness with perspective, see van Fraassen (2008).

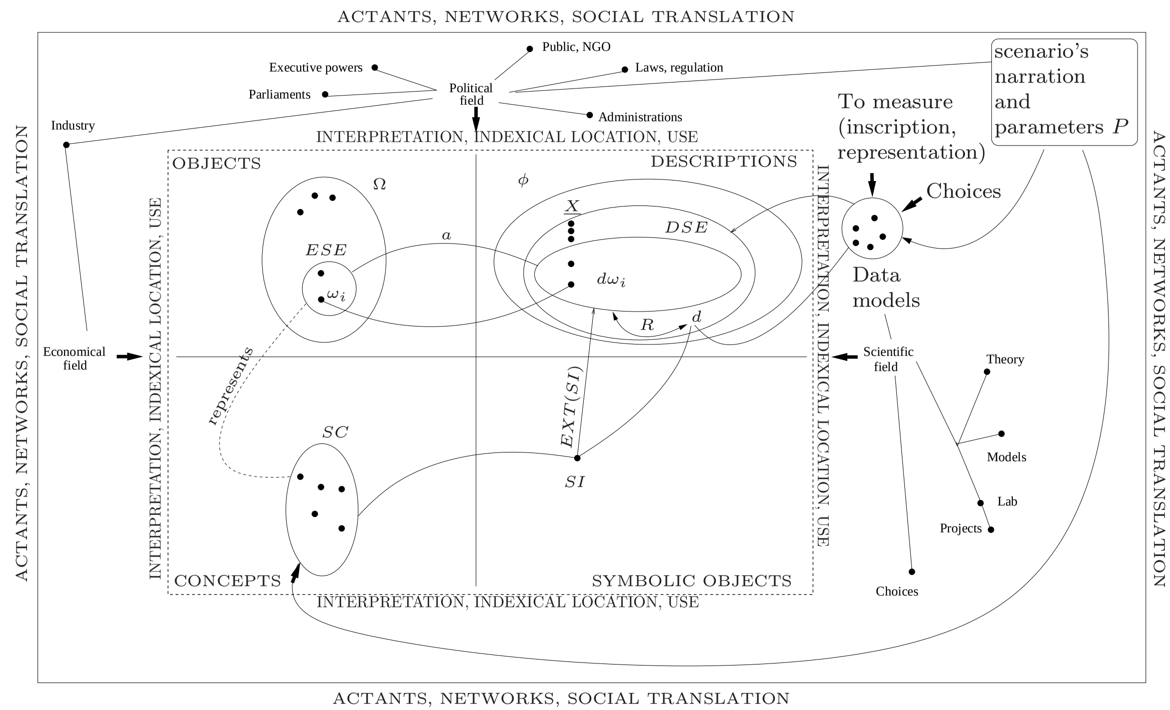


Figure 5.5: *Scenario's view3. Integrating the symbolics objects, interpretation, use and indexicality, measurement and translation aspects of representation using scenarios (a table of symbols can be found p.339).* A scenario is composed of a scenario's representing concept SC, of an extension ESE and DSE, of an intension SI, of a description d, parameters P, and of a narration. d is the result of measurements (product of a circulation reference, inscription) and/or of choice. ESE is a sub-set of Ω, defined with the aid of SI. DSE is e.g. a selection made within a symbolic data array $\underline{X}$ (a database). SI is the product of negotiations during a process of translation and interpretation in which heterogeneous actants are involved. Once SI has been stabilised in a network, it is possible to apply it to $\underline{X}$, in order to know ''what scenarios represent''. In this view, scenarios are border objects: they are one of the meeting points between well stabilised networks (scientific, politic, economic). The dotted line indicates that interpretation and indexical location takes place in more than one place in the figure. A direct link between SC and ESE can be established by simple denotation, but going through DSE to ESE is what opens the possibility to a representativeness quality assessment with some empirical content.

In the case of scenarios constructed from measurements, *if* they are used for picturing, *then* their reference points to what they are supposed to resemble, and the degree of resemblance gives an assessment of the quality of such a representation. If the use of the scenario does not involve picturing, then the fact that the scenario parameters are constructed from measurements may not play any role in the way they represent.

van Fraassen (2008) provides a conception of how something can represent (more details on that conception can be found in appendix H). Partially inspired by his work, we propose the following sentence:

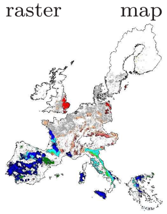
$$\textit{Z1 uses X to refer to Y as F for purpose P to Z2} \tag{5.1}$$

In (5.1), Z1 and Z2 are the users of the representor X. Z1 is having the performer rule, while Z2 is the rest of the actants involved in a translation process. Y is the represented object, F is the way Y is represented, and P is the purpose of the representation performance.

We advocate that, when performing the interpretation of a scenario, it is useful to determine (i.e. interpret) each term of such a sentence, as a guidance to interpretation. This is what we did, for the FOCUS scenarios case, in table 5.1, which is in page 126. We constructed such an interpretation on the basis of FOCUS scenarios' narrations as presented in chapter 1. At the moment of the FOCUS scenarios' development, Z1 is "FOCUS". However other performers of the scenarios interpretation exist (e.g. once the scenarios are used by APECOP). This highlights that the interpretation of the scenarios is not absolute, but rather relative to a particular situation, to a particular *use*. A temporal dimension is also introduced in the interpretation as shown in table 5.1, because various uses are encountered through the time: scenarios' development (1995), scenarios' use in the pesticide registration (2002) and scenarios' "validation" (2003). The interpretations in table 5.1 are either for the whole scenarios (for FOCUS (2002)), or for parts of it (for FOCUS (1995) and Vanclooster et al. (2003)). The values of soil and climate parameters that are mentioned in table 5.1 are provided in table 5.2, page 127. From table 5.1, we can deduce the necessary elements for a description, in terms of our View 2, of APECOP's interpretation of FOCUS Thiva-Sevilla scenario: see fig. 5.6, page 128.

Figure 5.6 depicts, according to our View 2, how FOCUS scenarios represent when used by APECOP for "validation" purposes. The actants participating in APECOP's interpretation of FOCUS scenarios are not depicted when using our View 2 scheme. Those actants could be extracted from the table in appendix G, and we could then obtain a full view according to our View 3.

Table 5.1: *How do FOCUS scenarios represent, according to our interpretation of the scenarios' narration in FOCUS (1995), FOCUS (2002), and Vanclooster et al. (2003).* Details about the FOCUS scenarios were given in chapter 1. In the sentence "Z1 uses X to refer to Y as F for purpose P to Z2", Z1 and Z2 are the users of the representor X. Z1 is having the performer rule, while Z2 is the rest of the actants involved in a translation process, Y is the represented object, F is the way Y is represented, and P is the purpose of the representation performance.

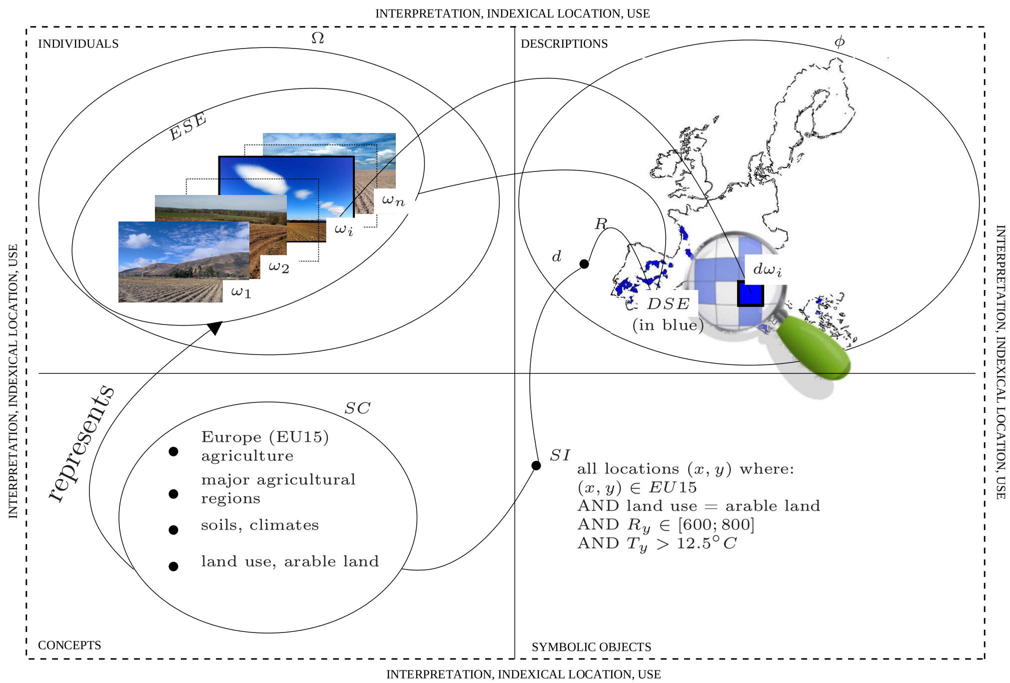
Z1	uses	X	to refer to	Y	as	F	for purpose	P	to	Z2
FOCUS (1995)		a hand-drawn sketch map 		the agricultural areas of Europe		divided in climatic zones		show the geographical borders where each scenario should be used		industry, regulators
FOCUS (2002)		scenarios		realistic worst-cases of vulnerability to pesticide leaching in major agricultural areas in the European Union		(1) geographically located in climatic areas (2) arable lands (3) having a series of values of soil and climate parameters		tier 1 leaching potential assessment using deterministic models		industry, regulators
Vanclooster et al. (2003)		a GIS-derived raster map 		the FOCUS scenarios		representing arable climatic areas		do a validation exercise		industry, regulators

climatic area	mean annual rainfall	mean annual temperature	Scenario's name
1	<600	<5	Jokionen
2	<600	[5-12.5]	Châteaudun
3	[600-800]	[5-12.5]	Hambourg/ Châteaudun
4	]800-1000]	[5-12.5]	Kremsmunster
5	]1000-1400]	[5-12.5]	Okehampton
6	<600	>12.5	Thiva - Sevilla
7	[600-800]	>12.5	Thiva - Sevilla
8	]800-1000]	>12.5	Piacenza
9	]1000-1400]	>12.5	Porto

Table 5.2: *The climatic areas that are represented by the FOCUS scenarios, according to FOCUS (2000)*

Figure 5.6: *Description, in terms of our View 2, of APECOP's interpretation of FOCUS Thiva-Sevilla scenario. According to APECOP (Vanclooster et al., 2003), the geographical component of the Thiva-Sevilla FOCUS scenario (with crops irrigation) represents a major agricultural area, that can be depicted into a geographical information system.*

ω: an object: the symbol used to indexically locate (represent) a given $5x5km^2$ portion of the landscape (the phenomenon) in the logical space of the scenarios as interpreted here; Ω: extension universe: the set of all ω pertaining to the first 15 countries of the European Union (i.e. EU15); ESE: enumerative scenario's extension: the set of all the ω represented by the scenario; ϕ: description domain: an object ω is represented in a GIS as having coordinates, soil types present in a given proportion, rainfall, temperature; DSE: descriptive scenario's extension (i.e. the GIS map of the areas covered by the scenario); SC: scenario's representing concept that is being interpreted as a symbolic object SI; d: scenario's description: $\{EU15; arableland; [600; 800]; 12.5\}$; $d\omega$: object description: a $5x5km^2$ pixel located at coordinates (x,y) described by $\psi(\omega) = \{x, y, soils, \%soils, R_y, T_y\}$; R_y: mean annual rainfall; T_y: mean annual temperature; R: relation: $\{\in, <, =, AND\}$.



5.1.4 Discussion of our philosophical concepts

Our proposal of a concept of representation using scenarios haves elements in common with philosophies which gives room, in their concepts, to *use*, *indexicality* (van Fraassen, 2008), *intentionality* (Husserl, 1976), and *interpretation* (Ricœur, 1986). We also took elements from sociologists studying the relations between science and *social translations* (B. Latour, J.Law, M.Callon and others, see appendix D). Relations between our views about representation using scenarios and some philosophical and sociological concepts are discussed here below.

Interpretation

In our proposal, "How do scenarios represent" can not be defined without both the *use* of the representation, and the intervention of an actor *performing* a representation each time he tries *to interpret* a scenario. This is in fact similar to what Ricœur (Ricœur (1986)) explains when comparing the act of interpretation to the act of playing music on the basis of a score. Moreover Ricœur explains (about the act of reading) that:

> *[T]he text is not without reference; the task of reading, qua interpretation, will be precisely to fulfill the reference. The suspense that defers the reference merely leaves the text, as it were, "in the air", outside or without a world. [...] underline the "present" character of interpretation. Reading is like the execution of a musical score; it marks the realization, the enactment, of the semantic possibilities of the text. [...] The "actualized" text finds a surrounding and an audience; it resumes the referential movement - intercepted and suspended - toward a world and toward subjects. This world is that of the reader, this subject is the reader himself. [...] Initially the text had only a sense, that is, internal relations or a structure; now it has meaning, that is, a realization in the discourse of the reading subject. By virtue of its sense, the text had only a semiological dimension; now it has, by virtue of its meaning, a semantic dimension. [...] So far [...] to interpret, we said, is to appropriate here and now the intention of the text. [...] [T]he depth semantics unveiled by the structural analysis of the text invites us to say that the intended meaning of the text is not essentially the presumed intention of the author [...]*

> *but rather what the text means for whoever complies its injunction. [T]o explain is to bring out the structure, that is the internal relations of dependence that constitutes the statics of the text; to interpret is to follow the path of thought opened up by the text [...]. [T]he object is the text itself; the sign is the depth semantics disclosed by structural analysis; and the series of interpretants is the chain of interpretations produced by the interpreting community and incorporated into the dynamics of the text [...] [T]he first interpretants serve as tradition for the final interpretants, which are the interpretation in the true sense of the term. [T]he entire theory of hermeneutics consists in mediating [the] interpretation-appropriation by the series of interpretants. [...] Appropriation loses its arbitrariness.* Ricœur (1991-1986).

For Latour (Latour (2001), Latour (2005)), an interpretation is the result of a social translation process, and is an actant itself. To interpret is to participate in a translation process. Quine (according to Vinck (1995)) advocates for a recreation of meaning each time the language rules are used: use is never repeated, it is a transformation; while Kuhn uses the term "reasonable agreements" between scientists about the interpretation of experiments (however, Kuhn was hostile to the sociological analysis of it).

What is the link between hermeneutics and phenomenology? Ricœur (1991-1986) contends the necessity for phenomenology to conceive of its method as an *Auslegung*, an exegesis, an explication, an interpretation. He explains one of the ways Husserl's phenomenology sees representation and interpretation:

> *[...] Husserl's use of the term* Vorstellung *- "representation" - to encompass both consciousness of generality and the consciousness of singularity [...] the two forms of consciousness refer respectively to "specific representations"*[9] *and "singular representations" [...] [P]erception "represents" because it is already the seat of work of interpretation, and it is because it represents that it can, in spite of singularity, serve as a "support" for specific representations.*

[9] I.e. of generality.

Interpretation vs empirical adequacy

Trying to make (as we did) a scheme were indexicality, empirical adequacy and interpretation coexists leads us to the hearth of the debate between constructive empiricism, phenomenology and hermeneutics, as stated by van Fraassen (1994) and the response to it given by Kockelmans (1997). van Fraassen (1994) recognises first of all that the question of the possibility of cognition as posed by Husserl can not be addressed in a scientific context, and that Husserl is not proposing a debate about the reality of the "external word". The question is about our attitude towards science and the status of science. van Fraassen (1994) proposes a paradigm for rational inquiry. He is however opposed to scientific realism[10]: he replaces it by empirical adequacy. van Fraassen (2008) states however that *"to be an empiricist [...] involves a common sense realism in which reference to observable phenomena is unproblematic"*. According to Kockelmans (1997), van Fraassen position is that philosophical positions can not be captured in dogmas that asks for beliefs about what the world is like, that a philosophical position may be an attitude or an approach, he rejects any form of transcendent dimension[11] as well as metaphysical arguments and any *a priori*. Kockelmans (1997) concludes from there that empiricism is unable to define what experience and phenomena are. Indeed, van Fraassen (2008) gives no further definition of phenomena or experience besides the "common sense realism". Moreover, Kockelmans (1997) states that van Fraassen does not discusses the process of thematization in and through which scientists make things become objects of research. One of the central arguments of Kockelmans (1997) against constructive empiricism is that van Fraassen's statements about science do not have their origin in experience. He considers thus that van Fraassen's arguments are indeed transcendental, in other words, he considers that van Fraassen is self-contradictory and uses undetermined *a priori*.

[10]Realism seen as implying (1) that the acceptance of a theory involves the belief that the theory is true; (2) that sciences gives a true story about what the natural word looks like.

[11]Understood by Kockelmans (1997) as the conditions of the possibility of all finite knowledge, including Husserl's transcendental phenomenology and Heidegger's ontology.

Kockelmans considers that there is an ontological projection typical to each science:

> *In this projection a certain domain of natural things is staked out, the approach to the domain is given its particular methodological direction, the structure of the conceptual and discursive explanation receives its orientation, and a specific "language" is constituted. The purpose of the thematization is to free the relevant natural things [...] in such a manner that they can become the objects of a purely theoretical discovery and, therefore, be examined objectively* (Kockelmans, 1997).

Regarding the phenomena, (that he conceives as the things insofar as they manifests themselves), Kockelmans (1997) refuses the dilemma between absolute truth and radical skepticism:

> *for humans, there is nothing wrong in accepting that to know always means to understand within a context of meaning that is already given in advance. This acceptance brings indeed an interpretative dimension into the picture. [...] Such contexts always contain prejudgements which need critical examination from even larger perspectives of meaning. [...] Such prejudgements [...] nonetheless are often not arbitrary insofar as they rest on very rational and reasonable grounds.*

In such a conception, truth does have a meaning, and the correspondence between a representation and a real thing, as well as the concept of empirical adequacy are replaced by an agreement *"between the content of a claim about a thing and the thing insofar as this can be discovered by us independently of our claim, by means of some process of confirmation which is intersubjectively acceptable. [...] a statement is true really means that [it] reveals the thing as it manifests itself to be in some process of confirmation that logically antecedes the claim"*. Other activities of man, for instance in perception, observation and experiments, are part of such confirmation process, and are conceived as *"intentional relations or inter-actions"* with the phenomena. Kockelmans (1997) thus considers that *"[w]e must move beyond semantics, mathematics, logic, and epistemology, to an ontological domain where issues concerning meaning and truth can be discussed."*. If we want to keep this concept of agreement under the umbrella of the concept of representation, a claim thus represents well a phenomenon when there is *agreement*, as defined by Kockelmans (1997).

Intentionality vs indexicality

In our view 3, the phenomena are indexically located as objects ω_i (which are symbols) into a mathematical structure. Indexicality, when seen as a link established between an object and a mathematical structure, seems problematic from a phenomenological point of view where the essence of a phenomenon is to appear and the duality natural object / subject is avoided. We argue that this is only an apparent contradiction, if we accept that the phenomena we are talking about here are, from the phenomenological point of view, constituted by intentionnality. The indexicality described by van Fraassen (2008) is then an operation that occurs (phenomenologically) *after* the constitution of the intentional phenomena. This view allows the coexistence of van Fraassen's constructive empiricism with a phenomenological approach. Someway, this is not contradictory with van Fraassen's views, when he says that metaphysics is not necessary in his system, but can be useful to stimulate the scientific imagination, or when he states that he is not offering a philosophy about what nature is, but rather a point of view about what science is (van Fraassen, 2008).

It seems to us that the fact that Ricœur's reference comes from the act of interpretation has something in common with what van Fraassen calls the *indexicality* of the scientific representation: the user's role. Moreover, in our framework, indexical statements can not be done without interpretation in a use context.

Interpretation, translation and use

Models in general do have various uses (see e.g. Nouvel (2002)). Models can be used e.g. as a heuristic tool in fundamental sciences, as a tool for action in applied and engineering sciences, as a tool to arrive at consensus, as support for (regulatory) decisions, as a communication language/tool between a diversity of actors, as prediction tools, as ways for a laboratory to obtain financial means, etc. This means that scenarios, when used in modelling also do have a variety of uses.

The use of a scenario determines the way a scenario represents. Our View 3 tries to capture as much as possible of the variety of uses, and is able at least to describe two general categories of uses: (1) scenarios used to be in a relation of perspectival selective resemblances (i.e. "picturing"[12]) to the conditions of the postulated phenomenon (as in PRIBEL); (2) scenarios used as inputs for a model, and receiving their reference from the outputs of the simulations. Moreover, our View 3 and our test cases cover the particularities of scenarios used in applied sciences, regulation and prediction. However, confrontation to other cases could be useful in order to test more in-depth the generality of our View 3.

We are thus at this stage in a situation where we recognize the existence and the rule of indexicality, interpretation and use. From there, we can follow the path shown in chapter 4 were the use of a scenario and the interpretation of a scenario's representing concept are actants resulting from a social process.

Use and measurement

Measurements have a function in representation using a scenario when the use of it (or its links to other actants - "articulation" in Latour's terms) is such that it has a picturing function.

The fact that scenarios enter into socio-technical translations does not mean that data models are necessarily purely arbitrarily or purely socially determined objects: Latour defines its philosophy of measurement through the processes of "inscription" and "circulating reference", which is in fact, according to us, very compatible with van Fraassen's measurement views (measurement as representation)[13].

It seems to us that the intension and extension that make up a scenario constructed from measurements may usefully be seen (fig. 5.7) as one of the multiple ways in which the process called "inscription" by Latour may be performed. By starting from matter, a symbolic data array $\underline{X}$ is progressively constituted (e.g. the map or another database) inside a

[12]According to van Fraassen (2008), representation through likeness is called "imagery", and when perspective is involved in imagery, it is called "picturing". Perspective implies e.g. occlusion, marginal distortion, grain, angle,... Picturing is a selective likeness. More over, imagery of non-existing things is possible and has a purpose in the context where they function. See more details in appendix H.

[13]See appendix H for a description of Latour's inscription concept, and van Fraassen's measurement concept.

description domain (a form) and then a second form is applied (the intension) determined intersubjectively (via translation and interpretation), to determine the descriptive extension of the scenario DSE (for example, the area represented by the scenarios or by the list of parameters).

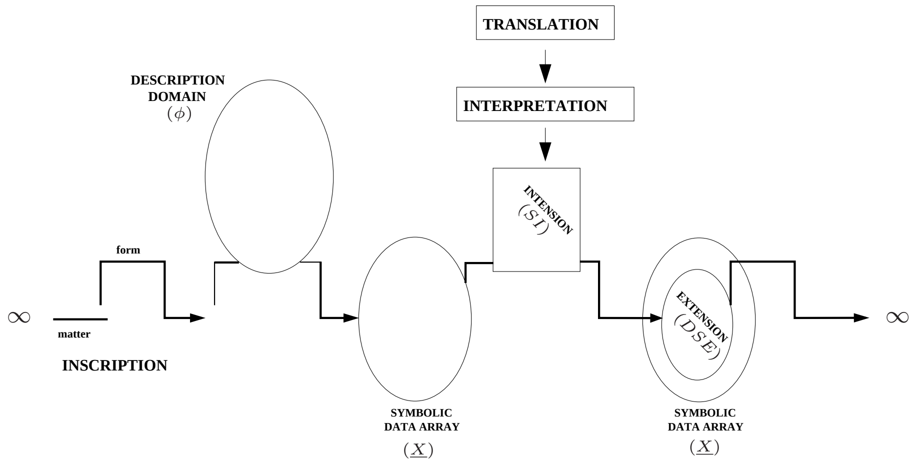


Figure 5.7: *Link between translation, intension/extension and inscription/measurement.* Intension is determined by a process of translation and interpretation. This intension is applied to a symbolic data array $\underline{X}$ (i.e. a data model, or a database), and thus determines the descriptive extension DSE. DSE is itself the result of a chain of inscriptions (inscription understood as in Latour's circulating reference concept) with an infinite succession of form / matter.

The so-called ''virtual scenarios'': a red herring

We emphasize that the two scenarios' uses just described (as picturing and as input for modelling) *are not necessarily encountered simultaneously in modelling practice.* Let us take our examples: the PRIBEL and FOCUS scenarios. PRIBEL scenarios are used as ''pictures'' of the system conditions (this is shown by the fact that scenarios' narrations claims that they were build from measurements and surveys), and at the same time, they are used in conjunction with a model for making predictions. However, this is not the case with FOCUS scenarios, as their narration states that:

> *[...] the nine scenarios should be viewed collectively as representing major agricultural areas in the EU. [...] FOCUS groundwater scenarios are virtual, in that each is a combination of data from various sources designed to be representative of a regional crop, climate and soil situation. As such, none can be experimentally validated.* (FOCUS, 2000)

Thus, as seen in chapter 1, FOCUS scenarios' designers refer to their scenarios as ‘‘virtual’’ AND as representing, at the same time. Are we here in search of a concept of representation of non-existing, virtual or fictitious objects that could solve this apparent paradox? How is this approached in philosophy?

For Husserl (according to Joumier (2007)), the objectifying acts are those in which something is constituted as object for us, either a thing or a state of affairs. Any intentional experience is an objectifying act or has its basis on it. Husserl divides the objectifying acts in two categories: signitive and intuitive. The former points towards something empty, while the latter can be imaginative or perceptive acts. Representation can be understood phenomenologically as following (Barsotti, 2001): ‘‘to know’’ is to fulfill an intention that was empty or partially empty. This is a generalised semiology supported by objectifying acts, and could be used as a basis for a concept of representation by ‘‘virtual’’ scenarios.

For Goodman (1990-1968), a symbol system consists of a symbol scheme (syntactic, with ‘‘inscriptions’’) correlated with a field of reference (semantic, with ‘‘compliants’’). For him, there exists *vacant* inscriptions, i.e. without a compliant, either because none was assigned, or because it was decided so, or none exists which corresponds to what is asked for. However, a vacant inscription can have a rule as well as others.

On Zalta's account (Zalta (1988) and Zalta (1983)), some objects (the ‘‘concrete ones’’) "exemplify" properties, while others (abstract objects like numbers, and what others would call "non-existent objects", like the round square, and the mountain made entirely of gold) merely "encode" them. Zalta's logic and axiomatic allows reasoning on such objects by extending classical propositional, predicate, and modal logic, and to describe the laws governing universal entities such as properties, relations, and propositions. This kind of approach could also be the basis for a logical approach of ‘‘virtual’’ scenarios. Moreover, Zalta (1998) makes a link between intensional logic and Husserl's intentionality, and more precisely between its concepts of exemplification and encoding[14] and Husserl's noemata[15]:

> *[...] for the purposes of building a logic of noematic senses, the suggestion that noematic senses encode properties may be the best way to proceed. [...] By encoding properties, the noematic sense provides a mechanism by which the noema directs us towards the world, towards objects that may exemplify the properties that are encoded in the noematic sense. Of course, in the cases of dreams, hallucinations, and presentations of 'impossible' objects, there are no objects that exemplify the properties encoded in the noematic sense.*

van Fraassen (2008) discussed also imagery of non-existing things. According to him, such an imagery can have a purpose in the context where they function, and thus according to their *use*.

We will go along the lines of van Fraassen (2008)[16]. The paradox of "virtual scenarios" that, at the same time, represents turns then to be a fallacy, a red herring. FOCUS scenarios' parameters are not as such a "data model", they are not taken to represent any specific field conditions (albeit it is not excluded they can have resemblance to a field): their use does not imply that they are "picturing". It is only *when used for modelling* that FOCUS represent: they represent the 90$^{\text{th}}$ percentile of leaching when used with a FOCUS model. In fact, *what is important for the use of the FOCUS scenarios is that the model outputs obtained from them, using the model, are adequate.*

[14] However, Jacquette (1991) criticises this link. This shows that there is an entire field of research on how to conceive the scenarios from the logical, semantic and philosophical points of view.

[15] Husserl distinguishes between noesis and noema: the intentional process of consciousness is called noesis, while its ideal content(s) is(are) called noema(ta). See appendix H.

[16] See appendix H for van Fraassen's picturing, data model and use concepts.

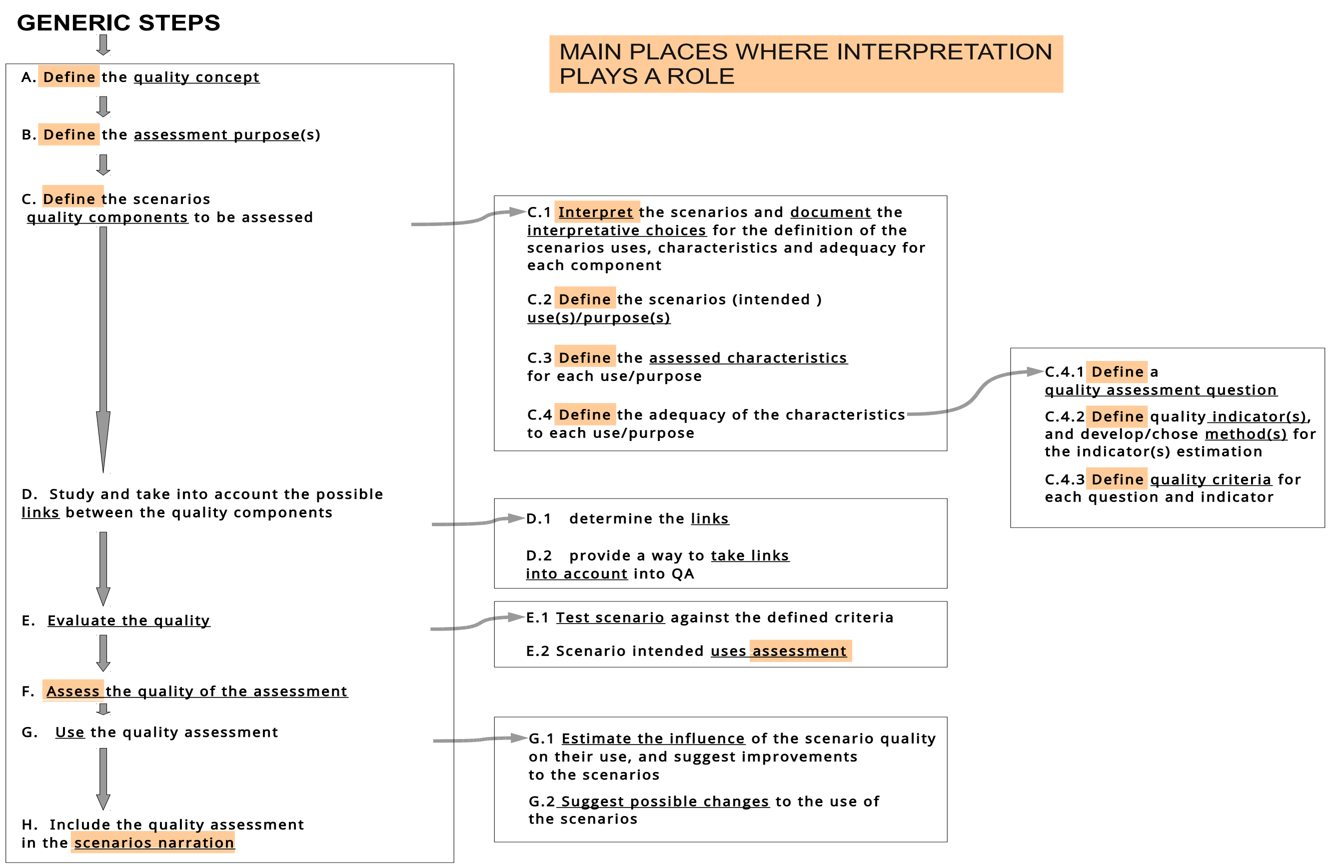
5.2 The role of interpretation in quality assessment

5.2.1 Quality assessment generic steps in which interpretation is involved

Our anthropological observation of the APECOP case (in chapter 4) shows that intepretation is involved in representativeness assessment, and our holistic view about representation, presented in section 5.1, proposes a way to conceive how interpretation links the various components of the scenarios representativeness quality, including the political component. The representativeness is not the only place where interpretation is involved in quality assessment. Figure 5.8 (p.139) presents the main steps of our assessment framework in which interpretation plays a role. Indeed, at those steps corresponds questions that, when answered, involves interpretation:

- What is quality? When performing an assessment, it is necessary to consider if our generic definition of quality is to be interpreted for a particular context (step A).
- What is the purpose of the assessment? (step B.) Our chapter 4 showed, using the APECOP case, that the purpose of the assessment can be subjected to discussion and interpretation (e.g. should-it be performed in view of obtaining a bad/good answer about the quality, or should-it be rather oriented to suggesting improvements?)
- How to define the quality components to be assessed? Interpretation plays a role at all stages in step C. This was discussed in chapter 3, that is devoted to the definition of quality components.
- What is a "good" use of a scenario? Step E.2 implies the necessity to understand how the scenario is used, or is intended to be used. According to the quality component that is being assessed, this will have various interpretations. This may e.g. refer to political considerations about the role the scenarios are playing in regulation, and for this various persons may have various views.
- What is a good quality assessment? Step F is in fact the whole subject of this thesis, and various interpretations may exist, e.g. the case of APECOP researchers, that only assessed the scientific quality component, versus our proposal of a multicomponent (thus multicriteria) assessment.

Figure 5.8: Interpretation in quality assessment of regulatory forecasting simulation scenarios. The highlights in the figure indicates the steps where interpretation is playing a major rule in determining the assessment results.



5.2.2 Tools for taking interpretation into account

Having recognised the importance of interpretation, the next question is how to take it into account. Our answer, constructed from developments in this chapter and previous ones is composed in six parts that are described here below:

1. hermeneutics
2. criteria and norms of interpretation
3. intersubjectivity
4. interpretation in representativeness assessment
5. transparency
6. narration

Hermeneutics. Hermeneutics gives us the basis for conceiving interpretation. In section 5.1, and in appendix H, we stated that interpretation of a text is not the action of finding what the author had in mind, but a performance done by the interpreter taking interpretants into account. Moreover, is to be performed each time an assessment is used. In this sense, we are really very close to the way Ricœur (1986) conceives interpretation, when he compares it to the act of playing music on the basis of a score: it is a *live performance.* Such a conception helps us to avoid problematic discussions such those that took place in APECOP (see chapter 4), where the scientists tried to find "what the authors of the scenarios thought about what the scenarios represent", which turned into a discussion with strange questions like: "What do you think the author thought?". When questioned, the authors did not agree about what they thought when writing the documents describing the scenarios! The answer is that interpretation is a live performance, to be done each time an assessment is performed, and that must be based on the available scenarios narrations.

Criteria and norms of interpretation. If interpretation is the basis of quality assessment, then the quality of the interpretation must be evaluated. Hermeneutics does not rejects the existence of criteria and norms of interpretation, but rather rejects the existence of *exact* criteria without the need of the interpreter's judgement and experience. Interpretation rules are rules derived from experience.

Romano (2010) mentions some criteria for such a performance:

- the interpretation's coherence;
- the interpretation's power (i.e. the ability to show a maximum of the object being interpreted);
- the consideration of the context (historical, political,...) in order to avoid e.g. anachronisms;
- the radicalism of the questions raised;
- the interpretation's originality, based on the knowledge and consideration of other interpretations.

Intersubjectivity Hermeneutics and norms of interpretation do not say much about the social mechanisms of interpretation. Our anthropological participant observation (in chapter 4) shows that the scenario concept SC is the result of the interpretation of the scenarios' narration, parameters, and various sources describing how the scenarios were constructed, their intended use, their legal framework. Interpretation is also found when SC is translated into a symbolic object SI, and in the choice of a given data model in which DSE is defined. This introduces the question of the objectivity in relation to interpretation. Objectivity can be considered as intersubjectivity mediated by language, following on this an interpretation of phenomenology by Mutelesi (1998). However, the content and social mechanisms of this intersubjectivity must be described. We have seen in chapter 4 how the definition of the scenarios' representing concept fluctuates depending on changes (translations) in the network in which they are inscribed. From this socio-anthropological viewpoint, a scenario's representing concept is socially constructed by the processes of translation of interconnected networks: interpretation is an intersubjective process, and from there is linked to social translation.

Intepretation in representativeness assessment In order to implement our representativeness criteria as defined in chapter 3, we propose here a concept of *representation using modelling scenarios* that sees representation as a user linking phenomena to mathematical structures through interpretation. In that concept, representation is the *performance* of a *user* that allows something to be *used* "instead of" something else. When performing a *representation using a scenario*, a user uses a scenario's narration and scenario's parameters to represent phenomena, as having properties described according to a description domain, for a given purpose, to someone. The link between the represented phenomena and the representor (i.e. narration and parameters) and thus the determination of the target of the representation, is made through an indexical statement by the user. The indexical statement is performed by the user through an interpretation of the scenarios parameters and narration. Such an interpretation, called here a *scenario concept*, is dependent on a use context and social translations between actants. The action of *performing* is to be taken as in ‘‘live performance’’: an action that is done every time something is used to represent. In our view, scenarios’ *use* is part of Ricœr’s interpretants to take into account when interpreting, while use and interpretation are themselves the result of a social translation.

The place given to *use* in that concept allows us to avoid red herrings in the understanding of the way the so-called ‘‘virtual scenarios’’. The FOCUS groundwater scenarios are claimed to represent, amongst others, agriculture, and at the same time they do not ‘‘mimic specific fields’’. Their parameters are taken to represent the 90th percentile of leaching inside their corresponding area, only *when used in combination with a FOCUS groundwater model*. This is sorted out by considering that no representation is possible without the intervention of a user that takes something to represent something else, and that it is the *use* of the representation that gives the criteria for a good representation. In the case of PRIBEL scenarios, they are explicitly used through their *resemblance* to specific farms or agricultural practices, so they represent in a different way than FOCUS scenarios. This gives philosophical grounds to the distinction between functional representativeness and measurement representativeness that was made in chapter 3. In the case of a functional representation, measurements as such may not play any role in the way scenarios represent.

In order to take into account the use of the scenarios when interpreting them, we propose that each term of the following sentence be interpreted for each case: *Z1 uses X to refer to Y as F for purpose P to Z2*, where $Z1$ and $Z2$ are the users of the representor X. $Z1$ is having the performer rule, while $Z2$ is the rest of the actants involved in a translation process. Y is the represented object, F is the way Y is represented, and P is the purpose of the representation performance. This is central in the way FOCUS scenarios can be interpreted into the symbolic objects mathematical formalism, and from there to interpret them when those same scenarios are used in a quality assessment process.

Something can be taken to represent something else, without the need for any detailed description of the represented objects. However, one of our goals is to introduce a concept of representation using scenarios that allows for empirical assessment of the scenarios' representativeness. Thus, in particular use contexts, i.e. when it is necessary to have a mathematical version (SI, i.e. scenario intension) of the scenario concept SC, as presented in chapter 3 for the needs of quality assessment, the indexical statement is performed through the interpretation of the scenario representing concept. In that mathematical formalism, ESE, the enumerative scenario extension is the set of the mathematical objects ω_i that are represented by SI and is termed ''enumerative'' to highlight that it enumerates the objects' symbols and does not refer directly to the phenomena. The descriptive scenario intension DSE is a representation of the phenomena to which reference is made by an indexical statement that links the ω_i in the ESE to those phenomena. The phenomena are indexically located as objects ω_i (which are symbols) into a mathematical structure. We see the phenomena themselves as constituted through intentionality. DSE is the structure that can be quantitatively compared to a reference extension in order to assess the scenario' representativeness. In such a concept, DSE plays—for scenarios—the role of van Fraassen's model's empirical mathematical substructures.

In our view, scenario's representing concept is socially constructed by the processes of translation of interconnected networks, and our concept indicates a link between interpretation and social translation. We gave names to the networks that are involved in the translation (scientific, political, and economical fields). In our conception (and in this we follow the sociology of translation) this does not imply that these networks pre-exist to the analysis: it is rather a matter of a posteriori constructions of the sociologist, but constructions that simply show fields where alliances are dense and strong, and whose constituent actants are often mobilised together.

The links between the components of our representation concept are thus as follows:

- The values of the parameters may come from measurements or/and from choices (e.g. in the case of scenarios from expert judgement).
- Measurements are conceived as circulating references (Latour, 2001), and we indicated similarities between this and the views of

van Fraassen (2008) regarding data models and indexicality.
- DSE, the descriptive extension, is the result of the application of a form (the intension) to a matter (the symbolic data array $\underline{X}$).
- The link between objects ω_i and phenomena is indexical, while phenomena are constituted through intentionality.
- Intension and intentionality are linked to intersubjectivity by interpretation, and interpretation is formed by this in objectivity.
- The process essential to the work of interpretation is translation as illustrated by our anthropological observations. The use of a scenario results from translation. The scenarios that we have studied are an actant on the border of highly institutionalised networks.
- Interpretation must take into account the use that is made of a scenario.
- In order to avoid red herrings, a measurement representation must be distinguished from a functional representation.

Our representation concept is quite general, and is applicable to regulatory forecasting simulation scenarios, which are an important part of the scenarios, both in number and in importance of their application domains. However, its generality and ability to describe in a useful way other types and uses of scenarios should be further tested, e.g. in the field of future studies and scenario planning.

Transparency. Given the importance of interpretation in scenarios quality assessment, our framework suggests to document the results of the interpretative process, for the seek of transparency. The need for transparency can be justified by the need for comparability between quality assessments of scenarios, e.g. the interpretation being a boundary condition for the scientific assessment of representativeness, two representativeness studies may lead to two different results due to either the different interpretations of the scenarios, or to different data or methods used. Making transparent the interpretation allows to distinguish the scientific component of quality from the other components. Another justification for the need of transparency is that equally scientifically justifiable interpretations of the scenarios may lead to different regulatory decisions, and thus the choice between various interpretations has a political character. The public, the regulators and industry need to know which interpretation is chosen, respectively for e.g. democratic reasons, for making informed regulatory decisions, or for avoiding arbitrary decisions with economical consequences.

Narration Assessment is based on interpretation, and intrepretation uses interpretatants that comes from other persons performing interpretations. An interpretation is better if it takes into account the previously made interpretations. Our interpretants should thus be included as an equipment of the scenarios for further interpretations performances: that equipment is what we call a *narration* (see section 1.1). The importance of having good narrations will be also highlighted in chapter 7, were we will see the difficulties related to the understanding of scenarios without a pre-existing narration. Moreover a narration is a contribution to and a tool for transparency.

5.3 Chapter conclusion

Chapter 4 showed that interpretation is involved in quality assessments, and that interpretation choices produces the links between quality components: choices between equally scientifically justifiable options are made based e.g. on juridical or political arguments. In this chapter we provided philosophical concepts that allows to conceive that role of interpretation, in a way that links the quality components and at the same time preserve the specificity of the approaches to each of them. A particularly important question for a scientist involved in quality assessment is thus the way to define the specificity of a scientific approach in a context of interpretation. In our proposal, a certain "hard core" of knowledge is governed by logics, measurements and empirical adequacy but coupled to (and not determined by) premises socially constructed. The link between the two is made through translation and interpretation. The example of the representativeness shows that the socially determined choices are the initial condition, the premises, from which the scientific work can start. Those premises are not questionable from a scientific point of view. To take an example from our case studies, do we consider that FOCUS scenarios represents climate areas, or agricultural areas? Both are scientifically justifiable. Those choices are then translated into mathematical formalisms that may provide access to an estimation of the empirical adequacy of the representation, or to plausible or credible estimations of that adequacy through modelling and simulations if experiments are not accessible. By being linked to interpretation in quality assessments, science may give scientific answers, but it is of paramount importance to know to which questions it is answering. We advocate that scientists assessing scenarios should acknowledge the existence of interpretation and take it into account, otherwise they run the risk of being the plaything of political interests, or to mislead the public about the conclusions that can be drawn from their scientific work.

We proposed a series of tools to take interpretation into account in quality assessment, starting from an inventory of the steps of quality assessment at which attention should be paid to interpretation. A philosophically based concept of interpretation allows to avoid to fall into the traps of the quest for the "true" interpretation or "the intention of the authors". Our answer is that interpretation is a live performance, to be done each time an assessment is performed, based on available scenarios narrations. The interpreter may use the criteria and norms for interpretation that we listed, either for performing his interpretation, or to justify it. However, our approach also acknowledges that interpretations are intersubjectively, socially negotiated. The interpretation norms are thus guides for a practice, but not means to reach "the true" interpretation. Various interpretations are possible, and thus the one that is chosen should be transparently documented, for various reasons: comparability between studies, to distinguish the scientific component of quality from the other components, for democratic control, etc. That documentation should be included in narrations as an equipment of the scenarios for further interpretation performances.

Now comes the important choice about who should perform the interpretation when in a situation of quality assessment. We distinguish two routes that will deserve our attention:

- A first one, that we call here *the scientific expertise route*, is to consider that an expert, or a group of experts, are the ones who should do the interpretation. It must be highlighted that scientific expertise here does not mean that no interpretation is needed: this chapter and chapter 4 showed the importance of interpretation is a scientific assessment. The classical scientific expertise that already involves interpretation is thus equipped here with a justification of the interpretation, written into a narration to make it transparent. Chapter 6 presents a test and an illustration of this for the FOCUS scenarios case.
- Another route sees intersubjectivity at the heart of interpretation, and a practical way of implementing the role of intersubjectivity into quality assessment, termed here *the participatory route*, is to design a participatory process that may have broadly speaking two objectives: to arrive at a consensus about the interpretation, and/or to map the a diversity of possible interpretations. Chapter 7 will test a partic-

ipatory scenarios assessment, for the case of the PRIBEL scenarios. In this route the expert interpretation is rendered explicit in order to be discussed, and others may propose other interpretations that the expert may not have seen, or for which the expert may not have seen the consequences for the decision making.

Part III

Framework testing: two case studies

Summary of Part III

In this part, the FOCUS and PRIBEL cases are used to test the quality assessment framework. The term *test* is understood here as an account of the applicability, limits and interests of the proposed framework. The two routes that were defined in chapter 5, are chosen to perform the test of our framework: the *expert route*, and the *participatory route* for performing quality assessments. Both routes take interpretation into account, and are complementary.

In chapter 6, FOCUS scenarios go through the expert route, for assessing the representativeness aspect of their scientific quality component. Chapter 6 is a test of steps C,E,F,G,H of the quality assessment. Succeeding to the assessment made by APECOP, a fresh new quality assessment of the FOCUS scenarios is presented here, as an example of a scientific quality assessment that at least makes transparent the central role of interpretation as a factor that conditions the output of the assessment. Transparency renders possible a scrutiny of the interpretation choices by stakeholders and the public.

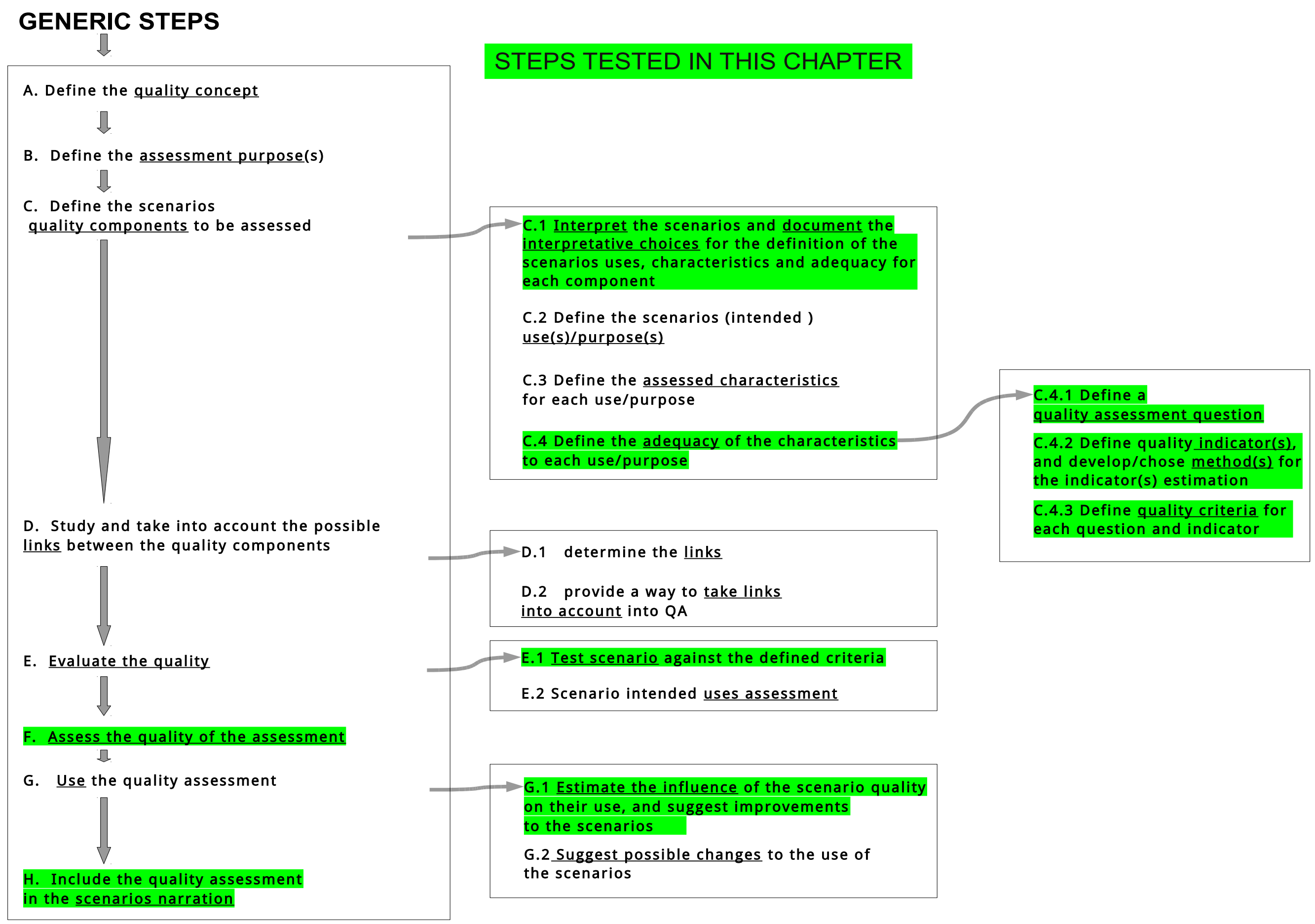
In chapter 7, the PRIBEL scenarios go through the participatory route, for an assessment of their political quality component. Acknowledging that, when performing a quality assessment, the political component is linked to the scientific component through the scenarios interpretation choices, the quality of the PRIBEL interpretation by scientists is subjected to a participatory assessment by lay-people. PRIBEL is taken as a test of a step further towards the full use of the framework (all steps, including the quality definition, except steps C.4,D.1,E.1) through the development and test of methods for assessing the political quality (in the sense defined in chapter 3), through the opening to discussion and documentation of the interpretative choices made by scientists when using scenarios in regulatory contexts, and through the use of the quality assessment for suggesting improvements to the scenarios and to the use of the scenarios.

CHAPTER 6

Making interpretation choices transparent through the scientific expertise route

Figure 6.1: Reading guide of chapter 6. This chapter contributes to the test of the proposed framework through an assessment of the representativeness component of the scientific quality of the FOCUS scenarios, in a way that takes interpretation and transparency into account.

Figure 6.2: Reading guide: contribution of chapter 6 to the test of the proposed framework generic steps. The FOCUS representativeness assessment requires (and thus is here a test for) steps C.1, C.4, E.1, F, G.1 and H.



Summary

This chapter tests our generic novel conceptual framework for scenarios assessment, for the case of the FOCUS groundwater leaching scenarios representativeness. Thus this chapter leads to two types of results that are linked but should not be confounded: (1) results about the FOCUS scenarios themselves, and (2) the lessons from it about our conceptual framework. The test is done through the expert route equipped with interpretation and transparency, as proposed in chapter 5.

(1) FOCUS scenarios are interpreted as representing the 90th percentile of leaching inside their corresponding climatic area. The results about the FOCUS scenarios representativeness show that 5 out of the 8 FOCUS scenarios have a poor representativeness, and should thus be used with great care. Those results are obtained from simulations since experimentation is not possible, and are thus to be considered only as plausible estimations. A transparent report of each step of the assessment processes is provided.

(2) This test case displays positive results for the applicability of our assessment framework, and our representativeness concept is operational. However, the expert route, even supplemented by transparency, leaves uncertainties, and this opens questions, that are examined in the following chapter, about the social acceptability of those uncertainties.

6.1 Introduction

The PEC calculations using the FOCUS-GW scenarios may be biased in representing worst cases of leaching, due to a potential under-sampling and/or wrong sampling, or wrong choice of the population of European soil types, climatic conditions and agricultural practices. Therefore, to assure the quality of the PEC assessment, it is of paramount importance to make progress in the quality assessment of the ground-water scenarios representativeness and to quantify and reduce this potential bias.

Regarding the representativeness, chapter 3 distinguishes between a functional and a measurement representativeness. For the case of the FOCUS ground-water scenarios, chapter 3 demonstrated that the *measurement representativeness* of the scenarios is not an issue (see section 3.2.3). In consequence, the present chapter is focused on the quality assessment of the functional representativeness of the FOCUS-GW scenarios.

Here below, the generic steps as presented in chapter 2 are thus applied to the assessment of the FOCUS scenarios. Steps A and B where already done in chapter 2 and chapter 3. As stated in the presentation of the quality assessment framework (chapter 2), the first step of the representativeness quality assessment must be an interpretation of the scenarios, which is thus presented here below.

6.2 Documentation of our interpretation (step C.1)

6.2.1 Scenarios narrations used for our interpretation

FOCUS-GW scenarios (see chapter 1 for a narration of these scenarios) are used for assessing diffuse pollution of groundwater by plant protection products. More precisely, they are used to calculate PECs (predicted environment concentrations), and are supposed to represent realistic worst cases of vulnerability.

FOCUS considered expert judgement to characterise some of the properties of its standard groundwater scenarios. Moreover, given the limited availability of pan-European environmental data at the start of the FOCUS activity, a statistical approach to infer optimal scenarios that represents worst-cases of leaching when used with FOCUS models could not be adopted (FOCUS, 2000). The worst-case character of the scenarios was thus deduced from an expert-judgement estimation of the variability at EU-15 scale of the soil and climate factors, rather than being deduced from an examination of the statistical distribution of leaching (in other words, the distribution of the response variable was deduced by expert judgement from the distribution of the identified factors). The PEC calculations using the FOCUS-GW scenarios may therefore be biased in representing worst cases of leaching.

FOCUS (2000) presented various uncertainties sources related to the FOCUS-GW model inputs (i.e. the scenarios):

- future weather versus past weather used in the scenarios;
- weather interpolation;

- weather data coming from heterogeneous datasets;
- soil pedotransfer functions;
- local soil parameters used instead of field scale effective parameters;
- crop representativeness for the scenario's areas (called here ‘‘FOCUS areas’’);
- crop parameters uncertainty (e.g. rooting depth); and
- substance parametrisation.

FOCUS (2000) suggested ways to improve their scenarios:

- a review of the appropriateness of the scenario's selection procedures;
- a data set improvement (soil profile analysis, and real weather data); and
- improving the ‘‘validation status’’ of the modelling for the different scenarios, through a comparison of model's results when used with the scenarios, vs the reality.

Although only Vanclooster et al. (2003) made an explicit and systematic assessment of FOCUS-GW scenarios, partial indications about their quality can be found in various publications. Kah and Brown (2007) presented results giving information about the ‘‘degree of protection associated with exposure estimates from the FOCUS scenario tools’’, which can be viewed as a partial indication about the groundwater scenarios' quality for a particular aspect (their results show that the pH chosen in the scenarios are not adapted to 35% of the soils found in England, and they discuss the consequences on ionisable pesticide sorption). The Advisory Comittee on pesticides, UK (2008) gives statements about the FOCUS surface water scenarios that are of importance for the groundwater scenarios also (many shortcomings were observed, particularly with respect to the coverage for certain crops and partial relevance of some scenarios). They presented modifications to the scenarios and suggested that the modifications would comply with regulatory outcomes. Hence they considered that the modified scenarios were sufficiently protective.

6.2.2 Interpretation of the scenarios

The translation of the scenarios narration into the symbolic objects analysis framework, as prescribed by our method for representativeness assessment, see section 3.2, requires an interpretation of those narrations. In the following sections our interpretation performance is presented and documented, for the seek of transparency and repeatability. This implies that the definition of the scenarios concept (SC), scenarios intension (SI), description domain ϕ, and descriptive extension DSE, assessment question, quality indicators, and assessment criteria are carefully documented here below.

6.3 Determination of the assessment question (step C.4.1)

According to § 3.2.3 the assessment question is found by translating the scenario's intension into the interrogative mode; this can be done in various ways. For this work, the following is chosen as the assessment question:

"Are FOCUS-GW scenarios combinations of parameter values selected in such a way that, when used in combination with a FOCUS PEC ground-water model, the calculated leaching values approximates satisfactorily the 90th percentile of leaching inside the corresponding scenario's area?"

The term "satisfactorily" is here to be evaluated through the matching between the scenario's extension DSE and the scenario's reference extension DRE. The matching function will be defined below.

Other interpretations, could have lead to other assessment questions, e.g. "Do scenarios' areas represent major agricultural areas? (from the pesticide usage point of view)"; "Do scenarios' areas span the range of temperature and rainfall in EU arable agriculture?"; or "Are the scenarios' areas distributed across EU?". Those assessment questions were answered in an unpublished deliverable of the APECOP project (Vanclooster et al., 2003) and will not be presented here.

6.4 Defining the quality indicator (step C.4.2)

6.4.1 FOCUS-GW scenarios concept (SC) and intension (SI)

The scenario's intension is determined by answering to the question *"What are the FOCUS-GW scenarios representing?"* The answer to that question is based on an interpretation of various sources describing how the scenarios were constructed, which is their intended use, and which is the legal framework, and is the result of the social translation described in chapter 4. Table 6.1 presents our interpretation of various bibliographic sources.

Table 6.1: *Our interpretation of various bibliographic sources in order to answer to the question "What are the FOCUS-GW scenarios representing?", and thus defining the scenario representing concept SC and scenario intension SI.*

<table>
<tr><th>SOURCE</th><th>OUR INTERPRETATION</th></tr>
<tr><td>FOCUS (1995)</td><td>A scenario in the context of the FOCUS procedure is defined as a representative combination of crop, soil, climate and agronomic parameters to be used in modelling. In this context representative means that the selected scenarios should represent physical sites known "to exist", i.e. the combination of crop, soil, climate and agronomic conditions should be "realistic".</td></tr>
<tr><td>FOCUS (2000)</td><td>Scenarios should represent major agricultural regions (as much as possible) and span the range of temperature and rainfall occurring in EU arable agriculture.</td></tr>
<tr><td>FOCUS (2000)</td><td>FOCUS-GW scenarios are representative for 9 corresponding agroclimatic areas, termed here "FOCUS areas", described as follows:
<table>
<tr><th>scenario</th><th>mean annual rainfall</th><th>mean annual temperature</th><th>land use</th><th>Scenario's name</th></tr>
<tr><td>S1</td><td><600</td><td><5</td><td>arable</td><td>Jokionen</td></tr>
<tr><td>S2</td><td><600</td><td>[5-12.5]</td><td>arable</td><td>Châteaudun</td></tr>
<tr><td>S3</td><td>[600-800]</td><td>[5-12.5]</td><td>arable</td><td>Hambourg- Châteaudun</td></tr>
<tr><td>S4</td><td>]800-1000]</td><td>[5-12.5]</td><td>arable</td><td>Kremsmunster</td></tr>
<tr><td>S5</td><td>]1000-1400]</td><td>[5-12.5]</td><td>arable</td><td>Okehampton</td></tr>
<tr><td>S6</td><td><600</td><td>>12.5</td><td>arable</td><td>Thiva - Sevilla</td></tr>
<tr><td>S7</td><td>[600-800]</td><td>>12.5</td><td>arable</td><td>Thiva - Sevilla</td></tr>
<tr><td>S8</td><td>]800-1000]</td><td>>12.5</td><td>arable</td><td>Piacenza</td></tr>
<tr><td>S9</td><td>]1000-1400]</td><td>>12.5</td><td>arable</td><td>Porto</td></tr>
</table></td></tr>
<tr><td>FOCUS (2000)</td><td>Scenarios are as "realistic" as possible by including most major European crops.</td></tr>
<tr><td>FOCUS (2000)</td><td>For PEC ground-water calculations, FOCUS intends to construct scenarios that represent an overall vulnerability approximating the 90th percentile leaching of all possible situations (this percentile is often referred to as a realistic worst case).</td></tr>
<tr><td>FOCUS (2002)</td><td>FOCUS groundwater scenarios are virtual, in that each is a combination of data from various sources designed to be representative of a regional crop, climate and soil situation.</td></tr>
<tr><td>FOCUS (2002)</td><td>Scenarios do not mimic specific fields, and nor are they necessarily representative of the agriculture at the location after which they are named. This is to be compared with FOCUS (1995) here above (the "physical sites known to exist"). We interpret this by saying that in principle, a field that resembles to each of the scenarios should exist, but this is not what is important for their function in the framework of the 91/414/EC directive. The FOCUS-GW scenarios' functional representativeness is more important than their measurement representativeness.</td></tr>
</table>

Having considered the texts in table 6.1, it can be concluded that various FOCUS scenarios representing concepts can be defined. For example, the scenarios should represent major European crops. From there, a scenario's intension could be derived in order to test, from agricultural statistics, if the major crops are included in the scenarios. Our choice is to define the scenario's intension based on the following representing concept (which is one of the possible assertions that can be derived from table 6.1):

> *FOCUS-GW scenarios' representing concept SC* ≡
>
> *A FOCUS-GW scenario can be described by (1) a combination of soil, climate, agricultural and crop parameter values and (2) a scenario's geographic area which it represents, located in the EU15, and defined by climate and land use characteristics.*
>
> *The climate and land use characteristics are defined according to table 6.1. The scenario's parameters are in such a way that when used in combination with a FOCUS PEC ground-water model, the calculated leaching values correspond to an accurate estimation of the 90th percentile of leaching inside the corresponding scenario's area.*

Moreover, we choose not to investigate the representativeness of the scenario nr.7 (Thiva-Sevilla irrigated), because it is an irrigation scenario that requires a different parametrization that the other non-irrigated scenarios.

SC can be translated into a mathematical form, taking as an example the Jokionen scenario:

$$\begin{aligned} & SI_1 \equiv \\ & locations \in EU15 \quad \wedge \quad LU(locations) = arable \\ & \wedge \quad R_y(locations) < 600mm \quad \wedge \quad T_y(locations) < 5°C \\ & \wedge \quad F(p_{90}^{FOCUS}(P_1)) = 0.9 \end{aligned} \tag{6.1}$$

where $EU15$ is the EU former territory (EU with 15 countries, from 1st January 1995 to the 30th April 2004), LU is the land use, R_y is the mean annual rainfall, T_y is the annual mean temperature, $p_{90}^{FOCUS}(P_1)$ is the p_{90} calculated with the FOCUS deterministic model PEARL[1] with the parameters P_1 of the scenario 1, and F is the cumulative distribution function (cdf) of the leaching.

[1]PEARL is described in (Tiktak et al., 2004).

In eq. (6.1), $(EU15,\ arable,\ 600,\ 5,)$ are the first four components of the vector d_1, which is the description of the scenario 1, and $\{\in, =, <, <\}$ are the relations R_1 that applies to d_1 as defined in fig. 3.3. Eq. (6.1) can thus be re-written in a more concise form:

$$SI_1 \equiv \quad F(p_{90}^{FOCUS}(P_1), R_1 d_1) = 0.9 \tag{6.2}$$

This means that $p_{90}^{FOCUS}(P_i)$ is supposed to be equal to the p90 of leaching inside the ith scenario's area described by $R_i d_i$, with $i = 1$.

6.4.2 Description domain ϕ

From the FOCUS-GW scenarios' intension, it can be derived that the description domain ϕ has two components. The first one ϕ_g is a geographical component: FOCUS-GW scenarios are representing worst cases inside their correspondent area, the latter being located in the EU15. In consequence, the description domain comprises a geographical data model[2] for describing those areas. In this work, a raster model is chosen. The second component ϕ_{cdf} of the description domain is any real number between $[0, 1]$ and is describing the value of the cumulative probability of leaching inside a given scenario's area.

6.4.3 Scenario's descriptive extension DSE

FOCUS-GW scenarios are representing a 90th percentile leaching inside their corresponding scenario's area (intension). This is interpreted into two DSE for each of the 9 scenarios: (1) DSE_g: the geographic extension is a scenario's area inside a raster map of Europe; (2) DSE_{cdf}: the cumulative probability of annual leaching. Corresponding to those extensions, two extension mappings must be defined: EXT_g and EXT_{cdf}.

Since the geographical component of the scenario's extension is a raster map, then *location* is defined by its coordinates (x, y) and the pixel size. If we take scenario 1 as an example, EXT_g is as follows:

$$\begin{aligned} EXT_g(SI_1) \equiv \quad & (x, y) \in d_1^1 \quad \wedge \quad LU(x, y) = d_1^2 \\ & \wedge \quad R_y(x, y) < d_1^3 \quad \wedge \quad T_y(x, y) < d_1^4 \end{aligned} \tag{6.3}$$

where d_1^j is the component j of the description of the scenario 1, LU is the land use, T_y and R_y are the mean annual temperature and annual rainfall. We have thus $DSE_{g,1} = EXT_g(SI_1)$.

[2]data model in van Fraassen's sense van Fraassen (2008).

In consequence, in a first step for assessment, DSE_g must be constructed in a GIS. We did it using Knoche et al. (1998) maps, for temperature and rainfall, while land use is from the map of PELCOM (2000). Fig. 6.3 presents the geographic extension DSE_g we constructed using EXT_g, for the 9 FOCUS-GW scenarios. $DSE_{g,i}$ is called the FOCUS area of the scenario i.

Chapter 4 described how various choices were made to arrive to such a map, through negotiations (social translation of the scenarios' interpretation). Chapter 5 showed how a conception of scenarios representation in terms of intension and extension can be linked to Latour's circulating reference. In fact, fig. 6.3 here is the illustration of the process described in fig. 5.7 page 135. More generally, fig. 5.5 page 123 shows how the DSE_g is linked to interpretation and translation.

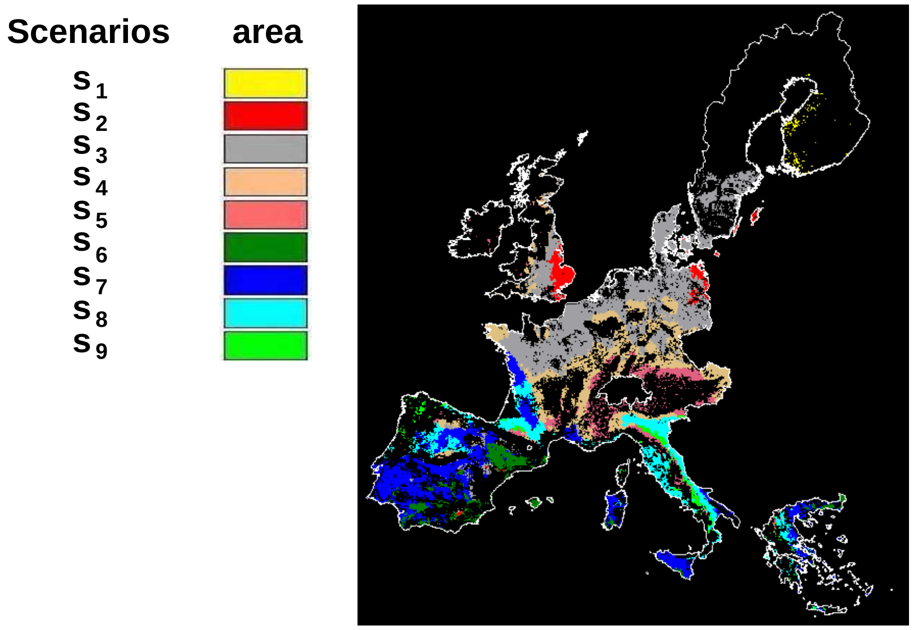


Figure 6.3: *The geographical component of the scenarios' extension DSE_g is a raster map, with 5x5 km^2 pixels.* The areas are derived by querying into 3 maps: the year rainfall, year temperatures (Knoche et al., 1998) and landuse (Mucher et al. (2000); PELCOM (2000)). Climate in each area is selected according to table 6.1 values. Regarding landuse, arable land is selected, according to the scenarios' intension derived in § 6.4.1. The frontiers correspond to the EU frontiers in 2003 (EU-15).

EXT_{cdf}, the extension mapping for the cumulative probability component of the extension, is defined here as the expert judgement (EJ) involved in the construction of the scenarios. In our interpretation of FOCUS (2000), a scenario, when used with a FOCUS model, gives an estimation of the p90 of leaching inside the corresponding scenario's area. In consequence:

$$DSE_{cdf,i} = EXT_{cdf}(SI_i) = EJ(SI_i) = 0.9 \tag{6.4}$$

6.5 Defining the quality criteria (step C.4.3)

6.5.1 Upscaling procedure for calculating the reference extension DRE

To assess the representativeness of the scenarios, it is essential to define the descriptive reference extension DRE. As for the DSE, DRE can have also a geographic component. However, the assessment question chosen does not ask for the assessment of the geographical component of DSE. In consequence, DRE is here a reference estimation of the cumulative probability (F_R) of the FOCUS leaching $p_{90}^{FOCUS}(P_i)$:

$$DRE_{cdf,i} = EXTR(S_i) = F_R(p_{90}^{FOCUS}(P_i), R_i d_i) \tag{6.5}$$

DRE must be as accurate as possible in order to have a high quality assessment, since DRE is to be compared to DSE. In consequence, the mapping $EXTR$ must be a more accurate procedure to estimate DRE than the FOCUS procedure. In an ideal situation, the 90th percentile is obtained from a detailed monitoring of the presence of active substance in ground-waters. However, our observational skills are limited in time and space, and monitoring is not available for new substances to be registered. Therefore the measured 90th percentile is unknown. If the 90th percentile of leaching cannot be obtained from direct measurements, an alternative is offered by simulating these leaching values by means of a spatially distributed leaching model in combination with pan-European soil, climate and agricultural databases. This type of assessment has a lower power than an assessment in which the 90th percentile of leaching is estimated from direct measurements, but it is the only pragmatic way to proceed with scenario's assessment at this time.

Tiktak et al. (2004) presented the EuroPEARL deterministic model, that could be used as *DRE* (EuroPEARL is a conceptually based, spatially distributed deterministic model). EuroPEARL used SPADE-1 soil profiles database (see the Soil Profile Analytical Database, SPADE-1 and SPADE-2 in: Hiederer et al. (2006), JRC (2009)). SPADE-2, a newer version of SPADE was not available at the time when EuroPEARL simulations of Tiktak et al. (2004) were performed. Since 2003, EU grew from 15 up to 27 countries (EU-27), thus the representativeness assessment here presented is to be revised if FOCUS-GW scenarios are used to represent the EU-27 territory. However, SPADE-2 can still not be used for such a revision. Still presently (in 2011) SPADE-2 does not cover all EU-15 countries, and in consequence, SPADE-2 can not be used for FOCUS-GW scenarios' assessment.

EuroPEARL suffers from some limitations. First the conceptually based model approach is a computer-intensive exercise. Second, the model ignores the variability of soil properties within Soil Mapping Units (SMUs) by using only the dominant soil typological unit (STU) within each SMU[3]. Third, in order to use the EuroPearl, an uncertain link needed to be established between the soil map and SPADE-1. Finally, the modelling procedure can not presently be implemented for approximately 25 to 48% of the European agricultural area, since nor SPADE-1 or SPADE-2 covers all EU-15 and all STUs[4].

To overcome those limitations, we decided to develop a metamodel, which is a model of a model. Metamodels map modelling codes in statistical significant relationships, thereby limiting the model to the description of the significant relationships within the considered system, with a significant reduction in calculation time and parameter requirements. The metamodel is to be used as part of an upscaling procedure. In our approach, three spatial levels are considered (figure 6.4): (i) the point scale corresponding to the mean and standard deviation of leaching at the pedon level, calculated with a metamodel, which includes climate variability; (ii) the grid scale corresponding to a regular grid of 5x5 km2

[3]The final aim of SPADE-2 is to expand the "estimated" soil profile database to include "primary soil properties" for all Soil Typological Units in the SGDBE v 3.2.8 and for both the designated dominant and secondary land uses (USE1 and USE2) for all the EU Member States JRC (2009).

[4]The actual % of coverage of EuroPEARL depends on the consideration of not of the soil variability of STU inside the soil mapping units SMU. 48% of the agricultural area can not be covered by EuroPEARL if the STU variability is taken into account.

and which includes soil and climate variability influence on leaching; and (iii) the regional scale for which a mean and standard deviation are estimated, and for which a DRE_i is calculated for each scenario's area: $DRE_i = F_R(p_{90}^{FOCUS}, R_i d_i)$.

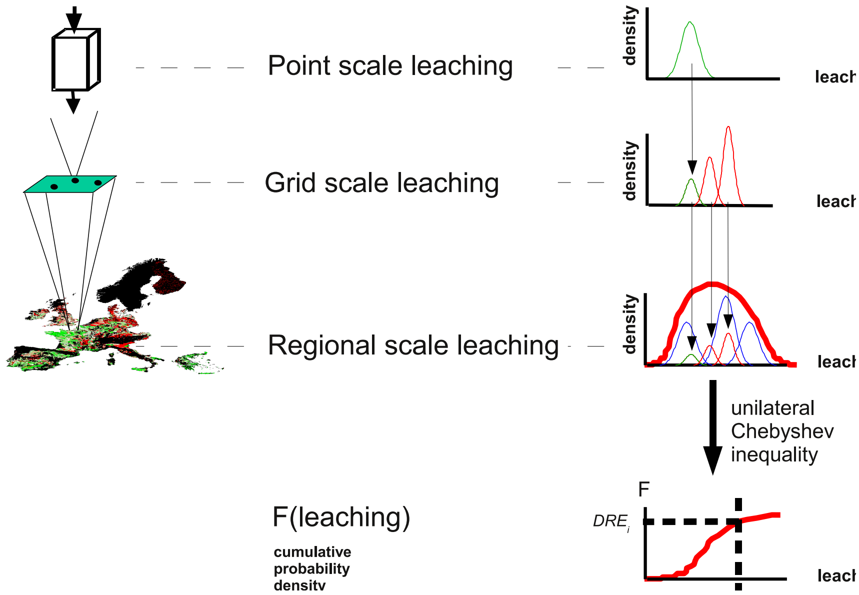


Figure 6.4: *The multi-scale spatially distributed leaching modelling approach.* Three scales are considered: the point scale, the grid scale and regional scale. For each scale, mean and standard deviation of the pdf can be calculated without making hypothesis about the form of the distribution. From the regional scale mean and standard deviation, the reference extension $DRE_i = F_R(p_{90}^{FOCUS}, R_i d_i)$ can be calculated using the unilateral Chebyshev inequality.

At each spatial level, the leaching concentration pdf is characterised by the mean and standard deviation, without making assumptions about the form of the distribution. The local scale pdf of the leaching concentration reflects the temporal variability of the annual leaching concentration due to weather. The grid scale pdf additionally considers the variability of the soil parameters within a 5x5 km2 grid. The regional scale pdf finally encodes the variability of the grid scale parameters within the region. The upscaling from the point scale to the grid scale is done by considering that the grid scale is a mixed distribution[5] from each indi-

[5]Mixed distributions: see e.g. Bishop (1995).

vidual point scale pdf. In the same way, the regional scale pdf is a mixed distribution of the grid scale pdfs. Finally, from the regional-scale pdf, the cumulative probability of leaching can be calculated.

This multi-scale hierarchical spatial modelling approach is based on the possibility of estimating some parameters of the point scale leaching concentrations pdf, from a reduced number of parameters present in the soil and climate maps. For such a purpose, a metamodel is necessary.

As presented in appendix K, according to the available pan-European spatial databases, the following input parameters can be made available for the metamodel:

- long term average mean seasonal and annual rainfall data;
- mean annual temperature; and
- top- and subsoil parameters (available water, organic matter, sand and clay content, packing density and pH).

As presented in fig.6.4, the upscaling procedure is based on the assumption that the distribution at the regional scale is a mixture distribution resulting from the leaching distribution at the point scale. Marin et al. (2005)presented an equation for estimating the moments X^m of mixture distributions:

$$E(X^m) = \sum_{j=1}^{k} p_j E^{f_j}[X^m] \tag{6.6}$$

where p_j is the weight of the f_j probability density function in the mixture, and k is the total number of pdf's in the mixture.

From that equation, the mean and standard deviation of the mixture distribution can be calculated:

$$\mu_X = \sum_{j=1}^{k} p_j E^{f_j}[X] = \sum_{j=1}^{k} p_j \mu_j \tag{6.7}$$

$$\sigma_X = \sqrt{\sum_{j=1}^{k} [p_j(\sigma_j^2 + \mu_j^2)] - [\sum (p_j \mu_j)]^2} \tag{6.8}$$

In a first step, the mean and standard deviation of the annual leaching at the regional scale is estimated from the point scale mean and standard deviation, using the upscaling procedure presented above (eq. (6.7) and (6.8)). Then, the obtained means and standard deviation are used to estimate $DRE_i = F_R(p_{90}^{FOCUS}, R_i d_i)$, using a unilateral Chebyshev equation described in Zwillinger and Kokoska (2000), and Tijms (2007):

$$\text{for } b \leq \mu_X, F(b) \leq \frac{\sigma_X^2}{\sigma_X^2 + (\mu_X - b)^2} \tag{6.9}$$

From eq. (6.9) we know that $F(b)$ is lower or equal to $\frac{\sigma_X^2}{\sigma_X^2+(\mu_X-b)^2}$, thus $F(b) \in [0, \frac{\sigma_X^2}{\sigma_X^2+(\mu_X-b)^2}]$, and in consequence, $\frac{\sigma_X^2}{\sigma_X^2+(\mu_X-b)^2}$ is an upper bound estimation for $F(b)$, the cumulative probability of b.

Eq. (6.9) is translated as following in our case, for each scenario i:

$$\text{for } p_{90}^{FOCUS}(P_i) \leq \mu_{X,i}, \quad DRE_i \leq \frac{\sigma_{X,i}^2}{\sigma_{X,i}^2 + (\mu_{X,i} - p_{90}^{FOCUS})^2} \tag{6.10}$$

Zwillinger and Kokoska (2000)and Tijms (2007) gives another version of the unilateral Chebyshev equation which is valid when $b \geq \mu_X$ and allows for a lower bound estimation of $F(b)$. Equation 6.10 is valid if $p_{90}^{FOCUS}(P_i) \leq \mu_{X,i}$. This in fact the case because very high values of leaching at the point scale are leading to a high mean at the regional scale, as it will be shown here below.

In consequence, according to our upscaling procedure, the outputs of the metamodel must be: the point scale mean μ_j and standard deviation σ_j of the annual leaching concentration at 1 m depth. Point scale mean and standard deviation of annual leaching (μ_j and σ_j) are to be obtained from 20 years PEARL[6] simulations of mean annual leaching concentrations.

For obtaining σ_j, we use:

$$E(S^2) = \frac{1}{n-1}(\sum_{t=1}^{n}(\mu_t - \overline{\mu})^2 \tag{6.11}$$

[6] PEARL is the point scale modelling unit of EuroPEARL.

which is the classical estimator for the variance of a sample. Possible time correlations between years are thus not taken into account.

Regarding the mean, we use the usual mean estimator $\sum_{t=1}^{n}(\frac{y_{NM,t}}{n})$.

In the proposed methodology, the mean and standard deviation of annual leaching at the point scale are used as the description of the components of a mixture distribution describing leaching at the EU level. The parameters of the mixed distribution are used in section 6.5.3 with an unilateral version of Chebyshev inequality to estimate F_{Rmax}, a maximal reference estimation of the cumulative probability of the concentration given by FOCUS.

6.5.2 A pesticide leaching metamodel that is used as the basis for the upscaling

The next step in the FOCUS-GW scenarios representativeness assessment is the development of a pesticide leaching metamodel (MM) for the agricultural soils of the EU-15, using calibration and assessment (‘‘validation’’) data from the EuroPEARL leaching deterministic model. The results of the MM are used to calculate DRE. The detailed development of the MM is presented in appendix K.

The MM is based on a radial basis neural network. The outputs of the metamodel are the point-scale inter-annual mean (μ_j) and standard deviation of the annual leaching (σ_j), for the D dummy FOCUS pesticide, which is a low persistent and low adsorbing compound, and for winter wheat. The metamodel used is a RBFMM (Radial Basis Functions Metamodel). A MKM (multidimensional kriging metamodel, as in Piñeros Garcet et al. (2006c)) can not be used because some input parameters vectors are close to each other, leading to singular matrices during MKM calibration. Moreover, the MKM can not be used in cases where the same inputs lead to different outputs, as it is potentially the case with the pesticide calibration data due to the inputs dimensionality reduction from PEARL to MM inputs[7]. The RBFMM presents a high model efficiency for the two chosen outputs, and is able to run over pan-European soil databases taking into account intra-cell soil variability, while it is not the case with the EuroPEARL[8] model alone. Moreover, the RBFMM

[7]In order to develop a generic tool, such possibility is sufficient to discard MKM here, even if only similar but not equal inputs where encountered in our dataset.

[8]Tiktak et al. (2004).

can use directly pan-European climate databases, without time-series interpolation, as opposed to the EuroPEARL parametrization scheme.

We use here the RBFMM to run it on European scale soil and climate databases, in order to calculate the inter-annual mean and standard deviations of annual leaching concentration, at the point scale.

The main databases used in the present work as inputs for the metamodel are: (1) for climate, Knoche et al. (1998) provides $5x5$ km^2 raster maps for mean annual temperatures, mean annual precipitation, mean rainfall in spring, and mean rainfall in autumn (spring means: March, April, May, and autumn means: Sep, Oct., Nov.); (2) for soil parameters, we used the European soil database (JRC, 2011), which includes a 1:1.000.000 soil map (Daroussin et al. (2006), Jamagne et al. (2002), JRC (2001)) which was provided to us as a $10x10$ km^2 raster map.

Figure 6.5 presents the map of mean leaching at the grid scale as calculated from the metamodel. The European soil map is originally a vector map composed of SMU (soil mapping units) with information about their STU (soil typological units). The STU represents the variability of the soils inside the SMUs. When converted to a raster map, the following assumption is made: the soil variability inside a cell is equivalent to the variability of the correspondent SMU. In consequence, the map presents the weighted mean of leaching of the STUs inside each cell (weighted by the area of the STU).

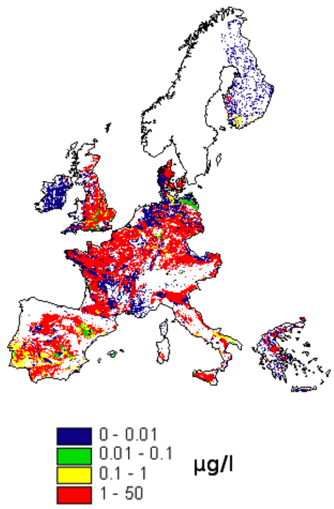


Figure 6.5: *Inter-annual mean of the annual leaching at the grid scale as calculated by the metamodel, taking into account intra-cell variability (i.e. taking into account the STUs inside SMUs).* Climate and soil input parameters are taken from Knoche et al. (1998) and JRC (2011).

When intracell variability is taken into account, the proportion of the arable land modeled by the RBFMM is greater than the surface modeled by EuroPEARL alone: 71% instead of 52%. The combination of PEARL with a metamodel is thus an improvement over the use of EuroPEARL alone. It must be emphasized that the modelling exercise presented here is a combination of a deterministic model with a neural network and a statistical upscaling procedure, and thus the metamodel and the deterministic model are complementary rather than opposed methods.

Significant differences exists between the EU-15 leaching maps generated by EuroPEARL and RBFMM as presented in appendix K. The RBFMM model efficiency being high (0.93 for the mean, and 0.97 for the standard deviation as calculated with data that where not used for calibration), those differences are not to be attributed to metamodel fitting errors. Those differences can not be attributed to extrapolation errors, because the metamodel incorporates a novelty detection of the inputs. Differences between the RBFMM and EuroPEARL are partially due to the fact that EuroPEARL uses dominant STUs (dominant soils in soil mapping units), while the metamodel uses all STUs thanks to its simplified parametrization. The incorporation of variability at lower spatial scales usually increases the mean leaching concentration (der Zee and Boesten, 1991), and this is also verified in the present work for the majority of the modelled area.

Significant uncertainties and errors exists in the soil databases used both by EuroPEARL and the RBFMM, and this leads to unquantified errors in both approaches when estimating leaching at the point scale. EuroPEARL uses the SPADE-1 database (JRC, 1999) to estimate soil properties, while the metamodel directly uses the 1:1.000.000 European Soil Map (Daroussin et al., 2006). One would expect a good correlation between soil properties in the European Soil Map and SPADE-1, but unfortunately this is not true, as shown in K. This is particularly the case for the extremely poor relationship between organic matter derived from the soil map and organic matter derived from SPADE-1. We conclude that the improvement of the pan-European databases should be a main research task for the upcoming years. The future results of the SPADE-2 project (JRC, 2009) will hopefully improve the available soil databases.

In section 6.5.1, it is stated that the upscaling procedure, and thus the representativeness test, is based on the assumption that $p_{90}^{FOCUS}(P_i) \leq \mu_{X,i}$ (eq.6.10 page 169). Table 6.2 presents a comparison between $\mu_{X,i}$ and $p_{90}^{FOCUS}(P_i)$, which shows that such an assumption holds in the present case.

Table 6.2: *Regional scale inter-annual mean of the annual leaching ($\mu_{X,i}$) as calculated using the RBFMM, and 90th percentile as calculated by FOCUS ($p_{90}^{FOCUS}(P_i)$), per scenario's area i, in ($\mu g/l$). Scenario number 7 is not modelled (scenario with irrigation).*

Scenario	$area_i$	μ_i ($\mu g/l$)	p_{90}^{FOCUS}
Jokionen	1	2.75	0.190
Châteaudun	2	6.61	0.140
Hamburg	3	18.13	1.100
Kremsmünster	4	19.92	0.510
Okehampton	5	18.83	1.900
Sevilla/Thiva	6	7.76	0.140
Piacenza	8	14.99	1.600
Porto	9	30.92	0.008

6.5.3 The assessment criteria: the MATCH function

To evaluate the representativeness, once we have the reference extension DRE_i and the scenario's extension DSE_i, the assessment question must be translated in mathematical terms to be able to compare DSE with DRE. The comparison is made by a matching function $MATCH$, according to the method described in § 3.2.3 (p.60).

We propose to compare p_{90}^{MM}, the 90th percentile of leaching concentration obtained by a reference assessment technique (which is here a metamodel combined to a distribution mixture model) with the 90th percentile of leaching concentration obtained with a FOCUS-GW scenario and FOCUS model, i.e. p_{90}^{FOCUS}.

This can be translated as follows:

$$MATCH(DSE_i, DRE_i) = P(p_{90}^{\mathrm{FOCUS}}(P_i) < p_{90}^{\mathrm{MM}}(S_i)) \tag{6.12}$$

i.e. the probability that the FOCUS-GW scenario underestimates the 90th percentile of leaching inside the corresponding scenario's area, if the metamodel is used as reference.

The probability presented in the test equation 6.12 can be calculated if a full assessment of the uncertainties is available, or at least a bootstrapping estimation of the uncertainties. This was not the case in the present work due to time constraints, in consequence a simplified version can be derived:

$$MATCH(DSE_i, DRE_i) = truthvalue[p_{90}^{\mathrm{FOCUS}}(P_i) < p_{90}^{\mathrm{MM}}(S_i)] \tag{6.13}$$

i.e. a Yes or No answer to the assessment question.

Equation 6.13 can be further transformed into an equivalent equation constructed in terms of a cumulative distribution function F_{Rmax} (estimated with the metamodel and the mixture model), instead of percentiles:

$$MATCH(DSE_i, DRE_i) = truthvalue[0.9 < F_{Rmax}(p_{90}^{\mathrm{FOCUS}}(P_i), R_i d_i)] \tag{6.14}$$

Eq. (6.9) gives an upper bound estimation of the cumulative probability, which is in fact a maximal estimation of DRE_i.

$$DRE_{max,i} = \frac{\sigma^2_{X,i}}{\sigma^2_{X,i} + (\mu_{X,i} - p_{90}^{FOCUS}(P_i))^2} \tag{6.15}$$

If $DRE_{max,i} < 0.9$, then we consider that p_{90}^{FOCUS} is below the 90th percentile of the reference extension, and that the corresponding scenario is not representative of a 90th percentile vulnerability. However, if $DRE_{max,i} \geq 0.9$, we consider that nothing can be concluded. Indeed, if $DRE_{max,i} \geq 0.9$, then the leaching calculated by FOCUS can correspond to a cumulative probability equal to any number between 0 and $DRE_{max,i}$, either below, equal or above 0.9. This is due to the fact that eq. (6.15) gives only an upper bound estimation of F_R, and thus an estimation of a lower bound for DRE_i can not be obtained.

Modifying equation 6.14 and using eq.6.15, the following matching function can be constructed:

$$MATCH(DSE_i, DRE_i) \equiv \begin{cases} \underline{\text{if } DRE_{max,i} \geq 0.9 \text{ then}} \text{ quality}(S_i) \text{ is unknow: might be good, or underestimation, or overestimation of p90} \\ \underline{\text{if } DRE_{max,i} < 0.9 \text{ then}} \\ \underline{\text{quality}(S_i) =} \text{ poor: underestimation of p90} \end{cases} \tag{6.16}$$

Such an equation could be improved if a full estimation of the uncertainties of the estimation of $DRE_{max,i}$ becomes available.

In summary, to perform the representativeness test, the following steps are followed: (i) point scale metamodel calibration and assessment; (ii) application of the metamodel at the point scale for each scenario area; (iii) upscaling for generation of the leaching mean and standard deviation for each scenario area; (iv) estimation of F_{Rmax}, the maximal estimation of the cumulative probability for the leaching calculated by FOCUS (i.e. $F_{Rmax}[p_{90}^{\mathrm{FOCUS}}(P_i), R_i d_i]$) for each scenario area; and (v) apply the matching function.

The FOCUS-GW scenarios were conceived in order to represent a 90th percentile of vulnerability inside 9 areas, and we assess 8 of the 9 scenarios (all except the 7th scenario, Sevilla-Thiva irrigated). There is one scenario per area, and in consequence 8 scenarios' areas for our assessment. The steps (iii) to (v) presented above are in consequence to be applied for each area separately. This procedure is presented in fig.6.6. Fig. 6.7 presents the conceptual framework in terms of the sets theory (compare to fig. 3.4).

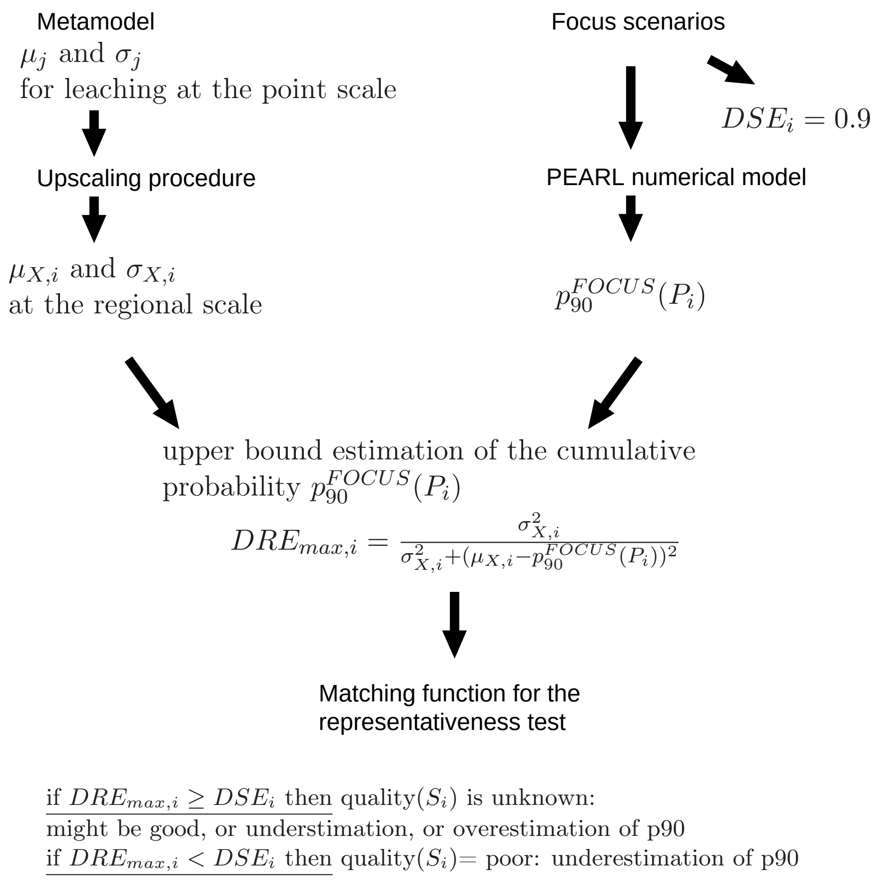


Figure 6.6: *The representativeness test of the FOCUS scenarios.* A metamodel is used to obtain an estimation of mean and standard deviation of annual leaching at the point scale. Using an upscaling procedure, an estimation of the mean and standard deviation of the pdf of annual leaching at the regional scale is obtained. In parallel, the FOCUS-GW scenarios are used with the deterministic model PEARL to calculate the 90th percentile of leaching at the regional scale. Then, using an unilateral version of the Chebyshev's inequality, an upper bound estimation of the cumulative probability of p_{90}^{FOCUS} is obtained, and used to perform the quality assessment test for the representativeness.

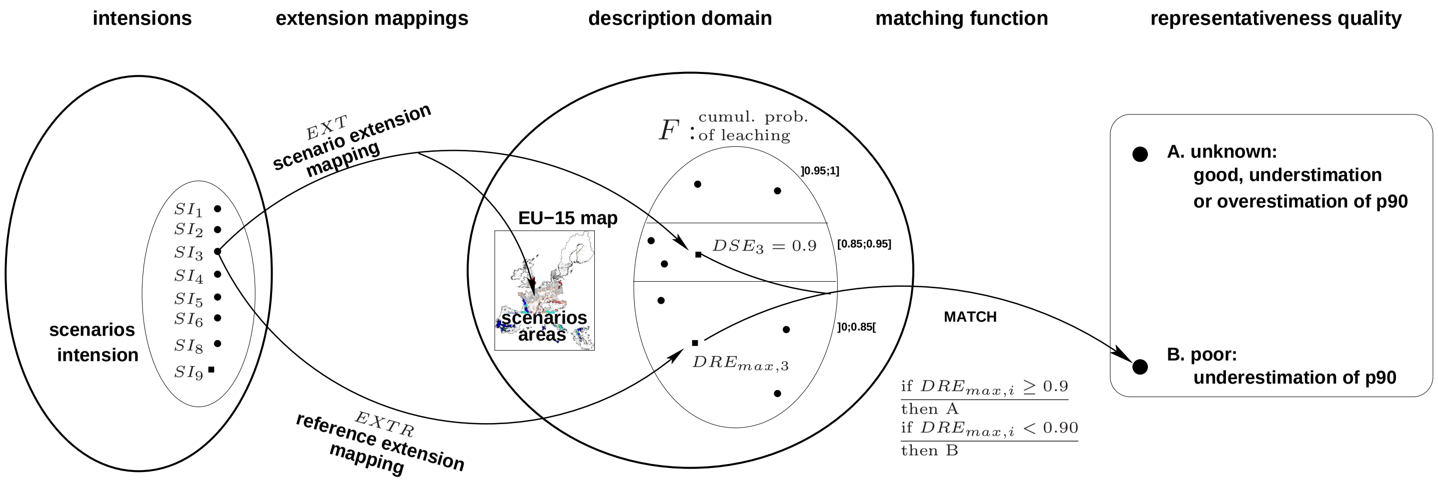


Figure 6.7: *Definition of the intension, extension mapping, reference and scenario's extension, matching function and representativeness quality, for FOCUS-GW scenarios.* A: quality(S_i) is uncertain:
might be good, or understimation, or overestimation of p90. B: quality(S_i)= poor: underestimation of p90. Other symbols: see table of symbols p.339

6.6 Evaluating the FOCUS scenarios representativeness (step E)

The next steps in the assessment procedure are the upscaling from the point and grid scale to the regional scale (i.e. to the scenario's area) for the estimation of the descriptive reference extension $DRE_{max,i}$ and the application of the $MATCH$ function. Using equations 6.7, 6.8, and 6.15, a maximal estimation of the reference extension $DRE_{max,i}$ ($= F_{Rmax}(p_{90}^{FOCUS}, R_i d_i)$) is obtained, and compared with the scenarios' extension DSE_i (=0.9) using eq. 6.16. This result is presented in table 6.3. Five out of the 8 maximal estimation of the reference extensions $DRE_{max,i}$ are below 0.90. This means that 5 out of the 8 leaching concentrations calculated by FOCUS are percentiles below 0.9, according to the metamodel, databases and upscaling procedure. If taken literally, as such, the results of the assessment test with the metamodel leads to say that 5 out of the 8 tested FOCUS ground-water scenarios are of poor quality regarding the representativeness criterion.

Focus area	Scenario	$DRE_{max,i}$	< 0.90 ?	test result
1	Jokionen	0.96	No	representativeness unknown
2	Châteaudun	0.90	No	representativeness unknown
3	Hamburg	0.79	Yes	poor representativeness
4	Kremsmünster	0.78	Yes	poor representativeness
5	Okehampton	0.84	Yes	poor representativeness
6	Sevilla/Thiva	0.90	No	representativeness unknown
7	Sevilla/Thiva irrigated			not modelled (scenario with irrigation)
8	Piacenza	0.83	Yes	poor representativeness
9	Porto	0.78	Yes	poor representativeness

Table 6.3: *FOCUS-GW scenarios' representativeness test results.* $DSE_i = 0.9$. If $DRE_{max,i} \geq DSE_i$ then quality(S_i) is unknown; if $DRE_{max,i} < DSE_i$ then quality(S_i)= poor: that means an underestimation of p90

6.7 Assessing the quality of the assessment (step F)

Using the available databases, EuroPEARL, and the MM, the representativeness of the FOCUS-GW scenarios has been found to be low for 5 out of the 8 tested FOCUS ground-water scenarios. The present work is assessing the representativeness only for one pesticide (a low persistent, low adsorbing one), and one crop (winter wheat). In principle, this question can be examined for a large variety of pesticides and crops, but a low representativeness for one case is sufficient for saying that the representativeness of the scenarios must be further examined. However, there are potentially considerable uncertainties in the prediction of the leaching concentration by both the metamodel and EuroPEARL, due to uncertainties in the prediction of Pearl itself, but also to databases uncertainties and errors. There are also uncertainties because time and space correlations where neglected in the upscaling procedure. As long as the effects of the different error sources on the leaching concentration have not been quantified, results of pan-European leaching studies, including the FOCUS, EuroPEARL and metamodel results, should be treated with extreme care.

The development of a statistical estimation for $DRE_{max,i}$ is an improvement over the method used by FOCUS to select their scenarios: FOCUS made the hypothesis that selecting a 80th percentile of soil vulnerability and then an 80th percentile of the vulnerability due to climate lead to an overall 90th percentile vulnerability scenario. This hypothesis can be shown to be false by a simple counter example with uniform distributions, but the result (the FOCUS 90th percentile estimation) could have been correct by chance. The present work does not infirm that in a definitive manner, but gives elements to have doubts about FOCUS-GW scenarios' representativeness of an overall 90th percentile of leaching.

6.8 Adequacy for the intended use, suggestions for scenarios improvement (step G)

A vulnerability that can be estimated without the characteristics of a particular pollutant (as claimed by FOCUS) seems hard to defend from a theoretical point of view: the scenarios defined by FOCUS may approximate the 90th percentile of leaching for some pesticides and not for others. We advocate rather that the FOCUS scenarios should be constructed as geographically distributed parameters into a GIS that stores soil, climate and crops parameters. Each new pesticide should be evaluated through the application of a leaching model or metamodel to the maps stored into the GIS. Such a scheme still depends however on the quality of the spatial data, that is presently poor, and on the quality of the (meta)models.

Our results means also that presently the statistical representativeness of the FOCUS-GW scenarios can not be taken for granted, and the present work shows that at least legitimate doubts exists regarding their representativeness. This is in line with Vanclooster et al. (2003))

who concluded that the Sevilla, Porto and Thiva scenarios may underestimate the 90th percentile of the leaching concentration, which would imply that these scenarios are not strict enough.

It is worthwhile, in conclusion regarding the FOCUS-GW scenarios, to state that both the present study and Vanclooster et al. (2003) points to potential problems of underestimation of leaching by some of the FOCUS-GW scenarios, but those conclusions are still preliminary mainly due to databases uncertainties. Without other data to complete, confirm or unconfirm those FOCUS scenarios quality assessments, from the scientific point of view, the FOCUS-GW scenarios must be used with great care in a regulatory context, and must be completed with a weight of evidence approach, including field and monoliths experiments as well as expert advice. The quality of the pan-European soil databases needs to be improved, particularly if representativeness of scenarios is to be assessed.

6.9 Inclusion into scenarios narration (step H)

The preceding sections showed that narrations are very important as they are the basis for the interpretation that is needed for translating the scenarios into the symbolic data analysis framework and from there for assessing the representativeness. The existence of this thesis, and in particular this chapter is a narration of the FOCUS scenarios that should be taken into account in future assessments of the FOCUS scenarios.

6.10 Chapter conclusion

This chapter tests our quality assessment framework, for the case of the FOCUS groundwater scenarios representativeness. Previous chapters implemented the necessary tools for estimating the FOCUS scenarios representativeness: chapter 2 gives the assessment steps, chapter 3 gives the definition and the symbolic objects formalism for representativeness assessment, chapter 4 showed the necessity to interpret the scenarios, and chapter 5 gives the foundations and criteria for interpreting the scenarios.

This chapter leads to two types of results that are linked but should not be confounded: (1) results about the FOCUS scenarios themselves, and (2) the lessons from it about our conceptual framework. The test is done through the expert route equipped with interpretation and transparency, as proposed in chapter 5.

Application of the framework to the FOCUS scenarios

FOCUS scenarios are interpreted (step C.1) as used to represent worst-cases of pesticides leaching. The chosen assessment question is: *''Are FOCUS-GW scenarios combinations of parameter values selected in such a way that, when used in combination with a FOCUS PEC ground-water model, the calculated leaching values approximates satisfactorily the 90th percentile of leaching inside the corresponding scenario's area?''*

This is interpreted into the symbolic objects formalism, as an estimation (step C.4.2) of the 90th percentile of leaching inside the area represented by a scenario. Such an estimation can not be obtained presently from experiments or surveys due to the difficulty of extending them up to the continental scale. We thus used simulations results from a metamodel (methods in step C.4.2) to approximate the 90th percentile of leaching. For that purpose, a radial basis functions neural network metamodel was developed. An upscaling procedure allows to obtain for eight FOCUS-GW scenarios the mean and standard deviation of the interannual leaching, at the corresponding scenario area scale. An unilateral Chebyshev equation is used in order to calculate a maximal estimation of the scenarios descriptive reference extension, to be compared (steps C.4.4 and C.4.3) with the leaching as given by FOCUS.

The results (step E.1) show that 5 out of the 8 assessed FOCUS scenarios have a poor representativeness. However, there are potentially considerable uncertainties in the prediction of the leaching concentration by both the metamodel and EuroPEARL, due to uncertainties in the prediction of Pearl itself, but also to databases uncertainties and errors. There are also uncertainties because time and space correlations where neglected in the upscaling procedure. What is the value (step F) of such a comparison? FOCUS used mainly expert opinions in order to chose the scenarios parameters. Our hypothesis is that the use of a metamodel with spatialized databases gives a better approximation of the 90th percentile of leaching. In consequence, while it is agreed that simulation results are not experiments, and thus that the conclusions regarding FOCUS scenarios are to be considered only as plausible estimations, our results imply (step G.1) that FOCUS scenarios must be used with great care, that five of them should be examined for further possible improvement of their representativeness, and that EU soil databases must be improved. Moreover, our advice is that while five FOCUS scenarios are useful as a way to extrapolate experimental results in view of providing indications about the existence (or not) of safe uses on a significant area in the EU, the uncertainties about their representativeness may prevent their use in a tiered process to decide if lysimeter or field experiments are needed. The assessment is based on modelling: to come to a more powerful assessment statement, results should be assimilated with detailed monitoring of active substances in European ground-water systems. It is expected that through the implementation of the European water framework Directive, reference monitoring data will become available and that direct comparison with measurements using techniques such as presented by Leterme et al. (2006) and Worrall (2002) may become possible in the future. Monitoring data are mainly useful, however, after having a pesticide already on the market. Modelling and other pre-market evidence are thus necessary if pesticide use is to be continued.

A transparent report of each step of the assessment processes was provided (step H). It must be stated that there is no unique FOCUS scenarios interpretation. Other results can be obtained with other choices, it is thus very important to state clearly which is the interpretation used and which are the references used to justify the interpretation.

Lessons about the conceptual framework

What can be concluded from the FOCUS representativeness assessment, regarding our framework? The proposed concept for representativeness assessment is currently operational as shown by its application to the FOCUS-GW scenarios. In our opinion, the method proposed presents a high degree of transparency, due to two main characteristics. First, it is based on a theoretical framework which obliges the person doing the assessment to define in a very detailed way the different steps that are followed, and to formalize it in a mathematical way. Secondly, the subjectivity and social premises that are inherent to assessment are not hidden, they are fully taken into account by the fact that the interpretation of the scenarios must be documented and described.

It must be noted that our framework can take into account cases where measurements or experiments are available: the lack of a possibility for an experimental assessment is particular to the FOCUS case. We can on the contrary conclude that the whole framework is applicable to regulatory scenarios. Moreover, the fact that different equally scientifically justifiable interpretations are possible is an argument in favour of the central role of interpretation in our framework, because they can lead to different impacts on the regulatory decisions that are taken. It is expected that a documentation of the choices and the respect of the steps will contribute also to the comparability of assessments of the same scenarios.

The application of the framework, that requires an assessment of the assessment, helps in setting the limits of the conclusions regarding the FOCUS scenarios representativeness. Indeed, if we come back to the objectives that we set to our framework, we find that the present test results are positive: the assessment lead to an evaluation of the influence of the quality on the use of the scenarios, to suggestions for improvement of the scenarios, and also to suggestions for changes to the scenarios use. The whole assessment displays thus positive results that supports steps C, E.1,F,G.1 and H of our generic assessment framework.

More generally, this chapter shows the conjunction of three difficulties that are encountered in many assessments that are done through the expert route in environmental regulatory situations: (1) a difficulty to extend experimental results up to the scale at which the regulated products are used, that calls for the use of models and scenarios; (2) incompressible uncertainties in the simulation results; (3) various equally scientifically justifiable interpretations may lead to different assessment results. Our framework helped in taking those difficulties into account by requiring an assessment of the assessment, that helps in setting the limits of the conclusions, and by requiring a transparency about the intepretation choices made by the experts. The limits of expert route in terms of uncertainties, even supplemented by transparency, opens however various questions that are examined in the following chapter, about the social acceptability of the consequences of those uncertainties, about the acceptability of the use of scenarios, about the acceptability of the intepretation choices made by experts.

CHAPTER 7

Opening interpretation choices through the participatory route

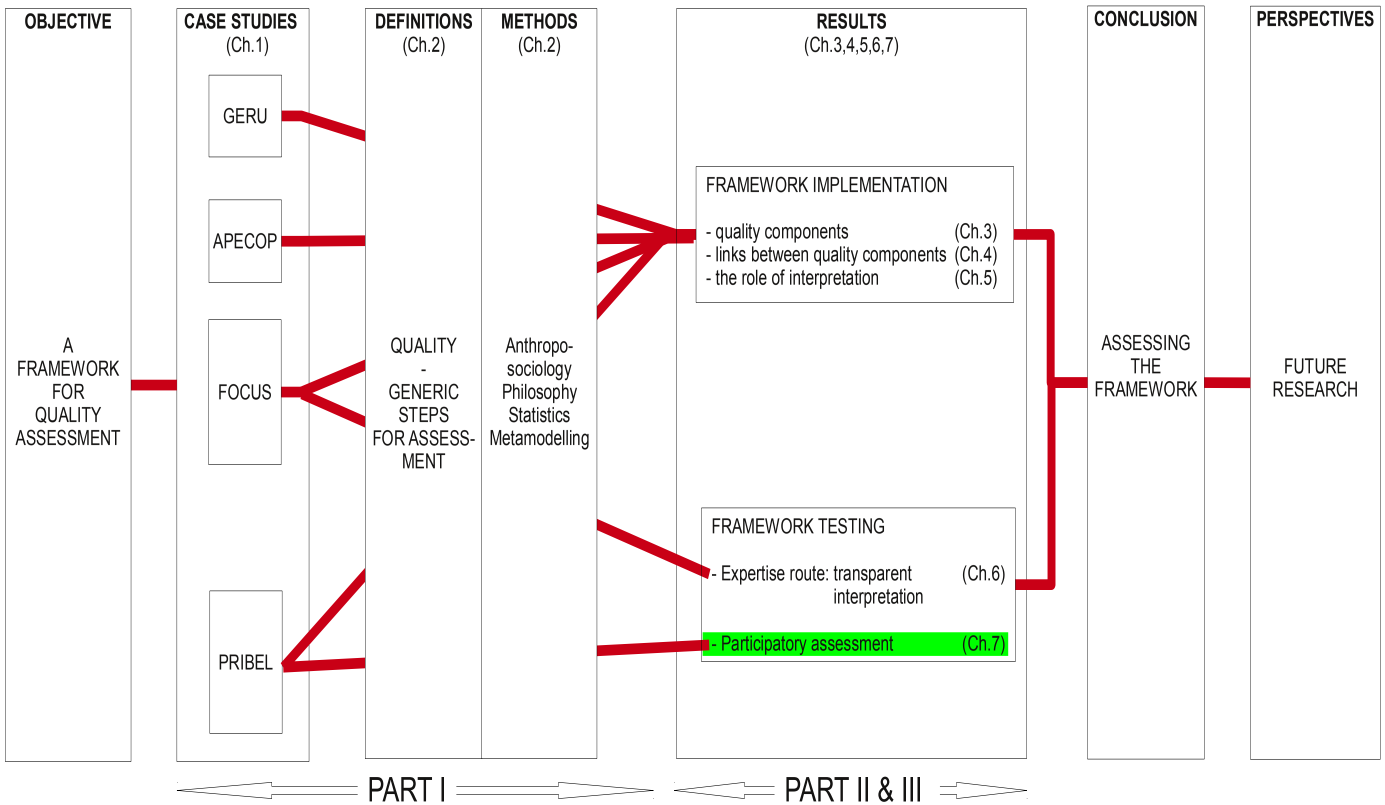
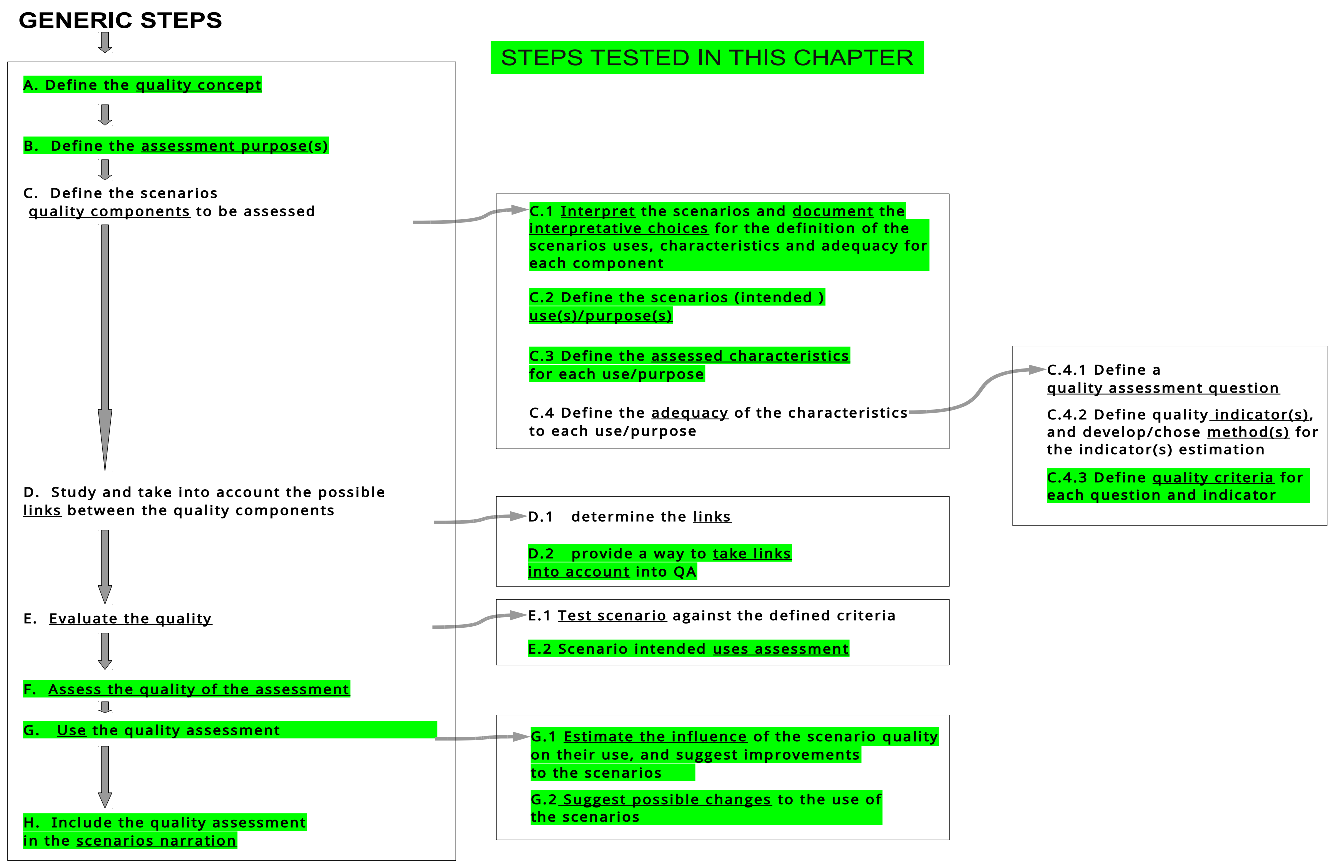


Figure 7.1: Reading guide of chapter 7. This chapter contributes to the test of the proposed framework by assessing the quality of the political component of the PRIBEL scenarios, through the participatory route. It applies the steps defined in chapter 2, the political quality component defined in chapter 3, and the role given for interpretation in chapter 5.

Figure 7.2: Reading guide: contribution of chapter 7 to the test of the proposed framework generic steps. A citizens participatory assessment of the quality of the political component of the PRIBEL scenarios is proposed. The questions to the citizens are sufficiently open in order to let the dispositive explore many of the steps, including the suggestions for improvement of the scenarios. It is thus also a first step towards a scenarios co-construction.



Summary

Chapter 3 defines what is meant by ‘‘political choices embedded into scenarios’’; chapter 4 showed from anthropological observations that such choices exist; and chapter 5 proposes to see those choices as interpretations that links the political and the scientific dimensions of scenarios quality. Acknowledging these links, this chapter presents the development and test of a method for opening scenarios to a participatory assessment with citizens. The contribution of that method to the generic steps is highlighted.

Cognitive insights where obtained about the way citizens understand the scenarios, and about the criteria they use for assessing them. The test showed also the possibilities and limits of a co-construction of the scenarios with citizens, and gives insights about the relations between uncertainties and confidence, the limited changes induced in expertise, and about the difficulty of interaction with a model under construction. The analysis of the participatory set-up points towards the need for an equipment of the model and scenarios' narration with two matrices: one with the choices in models and scenarios, the other with their role in the regulatory framework.

The test was performed without a political mandate, thus the extrapolation of the method to situations that binds the politicians or the regulators to the results is still to be studied.

7.1 Introduction

Scenarios construction often involves stakeholders in a traditional and well known manner, i.e. through a project steering committee, and without citizens involvement. FOCUS scenarios (presented in chapter 1) were developed through an expert group involving stakeholders (the group was constituted by academic researchers, industry and regulators). This kind of method had the advantage to lead to very specific scenarios. However, it is clear that this was also a method that had as result to *avoid* to consider some alternative practices in the modelling exercise (e.g. conservational tillage) and leads to an interpretation of the regulatory requirement of ‘‘realistic worst cases’’ that is only one of the possible choices amongst various equally scientifically justifiable interpretations (see chapter 4). This is a point where participatory exercises with lay persons are expected to arrive to different results.

In the area of pesticide risk assessment, a large number of scenarios, models and risk indicators are used for supporting political decisions, at various levels (see, e.g., Russell et al. (1994), Vanclooster and Boesten (1998), Jones and Mangels (2002), Vanclooster et al. (2003), Tiktak et al. (2004), Sahoo et al. (2005), Tiktak et al. (2006)). None of these models or scenarios were, or are, currently developed using a participatory approach with citizens[1]. Moreover, many uncertainties remain in the field of pesticide risk evaluation, and this rises questions, e.g. about the way to take into account the cumulative risk of various active substances on the same organism, or about the possible greater sensitiveness of particular populations. Various questions rise also when a synthetic view of many risks must be given (i.e. aggregation for a whole region/country or the entire EU, or for various species considered together). What are the weights of a given species or environmental compartments with respect to others? Does every body agree about giving weights? The answers and the choices made are not entirely scientifically determined. It is particularly in such a context, in presence of choices and uncertainties largely present, that choices should be made according to a participatory approach. This is needed not only to ensure social acceptability of the risk for decision-making, but also to take into account various kinds of knowledge and points of view not considered in classical scientific risk assessments.

In this chapter, we document the development and implementation test of a participatory approach for constructing and evaluating a model and its scenarios. The interest in getting involved in public participation can be shared by the modellers themselves, but, in actual fact, this is seldom the case. On the contrary, the present proposal started from the initiative of a modeller.

According to the literature, models and scenarios are submitted to the public *after* their construction, while we are proposing to explore the possibilities of co-construction of the model and scenarios. This is a central part in the originality of our proposal.

The work of Yearley concerning the study of a predictive model for atmospheric pollution in an urban environment (Yearley, 1999) attempt to inquire the issue of the appropriation of a modelling tool by citizens when citizens have difficulties in perceiving the reality of pollution through their own senses. The purpose of the study in question concerns citizens'

[1] A participatory approach with citizens is termed here below "participatory approach''.

interpretation and understanding of the process of constructing knowledge about the issue of air pollution and the use and relevance of the alert device. The conclusion drawn from the study is interesting in relation to the approach that we adopt in our own study, since it consists of affirming that: ‘‘to build robust and legitimate models, public bodies will need to devise methods of consultation and participation not only when the model is running, but also in setting out the objectives and parameters[2] of the model in its earliest stages’’ (Yearley, 1999).

The participatory approach presented here was developed by Piñeros Garcet et al. (2006a), within the framework of several ongoing modelling projects related to the modelling and aggregation of risks from pesticide use. These modelling projects are described in Piñeros Garcet et al. (2006b), Vergucht et al. (2007), and Piñeros Garcet et al. (2007). The participatory exercise presented here is conceived as a small scale test, performed in order to evaluate the feasibility of the developed method, as well as to test the basic hypothesis supporting it.

Piñeros Garcet et al. (2006b) evaluate the acute toxicity risk for bees associated with the use of pesticide. In this case, use was made of the PRIBEL model, embedded within a risk aggregation procedure[3]. The project was performed in the framework of the Belgian pesticide reduction plan (PRPB). The PRPB and the modelling work required the construction of scenarios (see section 1.3 for a presentation of the PRPB, the scenarios and the PRIBEL model). Vergucht et al. (2007) presents the development of a pesticide risk indicator, also in the framework of the PRPB. Finally, Piñeros Garcet et al. (2007) presents a method for risks aggregation, taking into account various environmental and health compartments at the same time.

Both the scenarios and the model presented in Piñeros Garcet et al. (2006b) and Vergucht et al. (2007) as well as the aggregation method developed in Piñeros Garcet et al. (2007) were submitted by a multidisciplinary team (social and modelling scientists, participation practitioners) to a participatory exercise with lay people. The exercise was performed

[2]The model's parameters values are what we call the scenarios.

[3]The PRIBEL indicators are called here the PRIBEL model: an indicator is in our view a simple model used to monitor the evolution of a situation, and involves often a comparison with a desired situation.

involving stakeholders (NGO, industry, regulators, scientists[4], farmers and beekeepers associations). The context of the Belgian reduction plan offered good examples of how scientific tools (PRIBEL and the statistical aggregation method) can be highly articulated to a political process; it was therefore ideal for testing how a participatory approach adds value to classical expertise and scientific work.

In a first stage, the proposal to develop a participatory method was made by myself, with the hope to contribute to the transparency and social legitimacy of regulatory simulation scenarios, in a participatory democracy perspective. I was thus involved both as a modeller and as organiser of the participatory exercise. My initial impression was that political choices were involved in the scenarios I was developing, while at the same time, my modelling work was presented as a ‘‘risk assessment’’[5]. I felt that situation as uncomfortable, and the proposal for a participatory exercise was made to find a way to elucidate the political choices, and to test their political quality (i.e. social acceptability and robustness[6]). Another initial hypothesis I had was that not all the political choices embedded in the scenarios were known by me (due e.g. to their integration into work habits and unquestioned traditions, to a lack of knowledge of the whole context, to a lack of experience in other areas not related to my work, to a lack of time to investigate it, to an organisation of the work in projects with narrow, precise and short term goals with rapidly changing tasks, etc). I suspected that it is a common situation to many modellers. I was thus ready to open the model and scenarios to suggestions for changes, with the hope that a participatory modelling will introduce choices in the model and scenarios that were acceptable democratically. Those initial feelings and hypothesis were reinforced by my previous participation to the APECOP project and the conclusions drawn from it (see chapter 4) showing that political choices are embedded in regulatory scenarios.

[4] I.e. scientists having developed the scenarios and model, as well as scientists from institutions not involved in the models and scenarios being examined in the exercise.

[5] Risk assessment is supposed to be "purely scientific" and to involve only scientific choices, as opposed to ‘‘risk management’’ which is supposed to involve political decisions.

[6] I.e. the citizens’ acceptance of their technical functioning, of the choices embedded in the scenarios, their objective – namely monitoring the risks of pesticide use in function of a pesticide reduction plan - and their use in their policy context, the ability of the scenarios to maintain themselves through time.

As stated in chapter 3, a scenarios' quality assessment performed exclusively by scientists / experts can be interpreted as a dysfunction of the collective as defined by Latour, since they are at the same time part of the perplexity, consultation and hierarchization processes. This can be seen has a reason for questioning the expert's legitimacy. But who do we consider as a legitimate expert? How can participation contribute to legitimacy? I consider that a necessary (albeit not sufficient) condition for establishing legitimacy is that important decisions (e.g. regarding major potential consequences on health or environment) are taken with a maximum of transparency towards the public. To present a model and its scenarios as "very complicated" black boxes and to say that "the model predicts this or that" is not acceptable, since the embedded choices and hypothesis are not presented. Transparency is therefore a matter of existence of a communicable information, of information availability, and of information understandability. But is the transparency a final guarantee of legitimacy and democracy? I had, when beginning the project, the hypothesis that power relations remain between the actors of decision making, despite the participatory exercise. This was however counterbalanced by the hypothesis that at least transparency and participation can avoid scientists to be the unconscious instruments of power relations, and provide the public with a starting point for questioning the decisions.

Those questions are in fact mainly related to the possibility of having value added to democratic processes and cognition by a participatory modelling exercise. The added value for the democratic process is expected to come from changes to the decision-making process, from an improvement of the transparency, and from and improvement of the acceptability of the model and scenarios. The added value regarding cognition is expected to come from some sort of cross-learning and knowledge co-production, e.g. when it is expected that the model and scenarios may be modified through a participatory exercise.

"The public" (whether directly concerned or not), evolve in some cases towards a full-fledged actor in the environmental controversies (Fischer, 2000), rather than being only consumers or citizens willingly delegate their actions to their political representatives. However, "the public", in its role as an actor, is very far from appearing in a homogeneous, monolithic and stable form. This so-called actor status should be imagined more as a posture, which can take highly varied forms.

The adopted posture cannot be dissociated from the means implemented to solicit and involve the public.

Cases show that lay persons, ‘‘ordinary’’ citizens, placed in a participatory posture have real aptitudes and skills in posing problems and envisaging solutions (Fiorino (1990), Fischer (2000), Funtowics and Ravetz (1991), Wynne (1999)), but different from those of persons in a scientific expertise posture. The aptitudes that were pinpointed show, among other things, that ‘‘ordinary’’ citizens place the problem in a wider setting than experts, because they are not constrained by disciplinary boundaries. They also show that citizens allow to mobilise a broader spectrum of expertise upstream of decision making. They show, further, that citizens' participatory involvement provides an opportunity to define the limits of the experts' models to be tested. As regards solutions, citizens are more inclined to identify alternatives than experts are; citizens can contribute to decisions in case of uncertainty and take into consideration possibilities of errors in decisions; their opinions lead to greater sensitivity to values and common sense. Citizens are attentive to questions that often proves to be crucial for decision making, namely the question of trust and the reliable nature of the behaviours adopted by the various stakeholders involved in a given issue (Alsop, 1999). This aptitude is precisely what makes it possible for them to link dimensions of the problem that are often considered to be separate.

The possibility of cross-learning between a public and, in this case, scientists, is based on the possibility of implementing new methods of collaboration giving the priority to the co-production of knowledge (Callon (1998), Epstein (1995)). This co-produced knowledge acquires thereby a hybrid status, but it is hoped that it will serve to reinforce and legitimize the decision-making processes. A participatory process can aim at something more than just constituting a panel of citizens learning the sometimes complex dimensions of the impact of a technology on the environment and expressing a considered opinion about this. A participatory method should also incite the initiators (in this case, modellers who are members of a public research institution) to engage in reflexive dynamics that question both their modelling practices and their perceptions of the wider issue of which they form part.

A precondition to a participatory exercise with models and scenarios, when they are used to assess large scale environmental problems, is that they can be actually appropriated by the public or a citizens' panel.

This precondition is both a cognitive and political one. The difficulties involved in the public assessment of expert models are not alien to the very format of the knowledge being manipulated and to the IT architecture, where choices are based on hypotheses made by the modellers, concerning the state of the social world (i.e. the behaviour of agents, their preferences, etc.). The format and architecture do indeed pose real problems of physical and intellectual accessibility when a participatory method is being developed.

The participatory methods are diversely accepted within political or administrative institutions, industry or the scientific community (Murdock et al. (2005); Pearson (2001); Weber (1999); Van Cauwenbergh (2008)), each having constructed ways to be distinguishable from the citizen, consumer or lay person respectively. The use of modelling as an environmental risk management practice is flourishing (studies of the dispersal of pollutants, noise, smells, floods, etc.). It, however, needs to be appropriated with considerable technical expertise (a good command of IT and mathematical tools). This way of appropriating is, unfortunately, at odds with civil society's – politicians' or citizens' - way. As a corollary, the model as a scientific and political assessment, management and decision-making tool has been the subject of very little research into its social and public accountability (Yearley, 1999). Moreover, the scenarios are even less studied. It is nonetheless important to point out that, unlike an idealised view of the benefits of public participation in the social assessment of new technologies or public environmental risk management policies, the involvement of the public in a participatory method is not an easy matter. The reasons for that are many and varied and should be sought not only on the ‘‘social’’ side with the disillusioned, critical or cynical perception that citizens may have concerning their role in the consultation or decision-making process (Garland, 1999), but also in the concrete technical procedures employed to implement the actual participatory methods and, above all, in the difficulty involved in assessing them (Rowe and Frewer (2000), Mormont and Mélard (2003), Rowe et al. (2004), Rowe et al. (2005)).

There exists a large variety of participatory methods (see e.g. Elliott et al. (2005)). Following Rowe and Frewer (2005), we can summarise approaches and methods for public engagement into three key concepts: a) public consultation, b) public communication, and c) public participation. These differentiations are mostly due to the nature of the information circulating between the organisers and the participants, to the status and competences for which the participants are acknowledged, and to the participants' positioning in relation to the definition of the considered issue.

Hisschemöller and Midden (1999) studied the advantages and drawbacks of the main approaches to public engagement. They concluded that the choice and implementation of a given approach require clarification of the place that the initiators reserve for the public concerned[7]. Choice and implementation depends also on the initiators' more or less marked ability or willingness to allow themselves to be questioned (Stengers, 1999) or even to question their own definition of the issue, especially for the case of public participation in Rowe & Frewer's meaning (Rowe and Frewer, 2005) of the concept.

It appears that one of the keys to success of a participatory process lies not only (and typically) in the capacity of the method to reflect the public's perceptions, but also for describing, documenting and confronting that perception with the perception that the initiators may (often implicitly) hold about the public in question, whether the initiators are political, scientific or industrial. This reflective approach requires the circularity of perceptions and representations and makes their interdependence intelligible and productive. It is the focus of the recent concerns of researchers working in the area of the relationship between stakeholders and citizens (Marris et al., 2001) and more precisely the relation between scientists and citizens (Cook et al., 2004).

Since equally scientifically justifiable choices exists for constructing scenarios in a regulatory context (see chapter 4 and chapter 6), leading possibly to different regulatory decisions, those choices are considered to be political. In consequence we propose here a method that corresponds to step D.2 in our framework (take links between quality components into account when performing quality assessments). The link is made here between the scientific and political components by asking citizens to assess the choices that were made in scenarios between various scientifically justifiable principles, hypothesis and parameters, and about their applications. However, as we will see, citizens went a step further,

[7] The place reserved depends on the power relations, democracy (participatory/direct vs representative), to the rule of the state administrations, regulators, and politicians, NGOs, etc.

and assessed also the effect of the use of the scenarios for them and for the whole society, and suggested modifications to the scenarios.

7.2 Objectives

The general objective of this chapter is to test our quality assessment framework steps, as defined in chapter 2, implemented through the participatory route as defined in chapter 5, for the political component.

Our specific objectives are thus to:

- develop a participatory method for scenarios assessment;
- test the method using the PRIBEL case;
- study the contribution of the participatory method to the generic steps of the framework;
- evaluate the potential of the method to contribute to the goals we set to our framework: it should evaluate the quality of a scenario in its multiple aspects in a transparent way, and give insights about (1) the influence of the scenarios quality on the use of the scenarios, (2) potential improvements to the scenarios, and (3) potential improvements or changes to the scenarios use.

Our objectives concerns thus the feasibility and cognitive aspects of the method (i.e. the information that can be expected from such an exercise). Moreover, since the present work is a small scale test, and given that our aim is methodological, more attention is paid to the potential of the method than to the particular results regarding the quality assessment of the PRIBEL model and scenarios.

7.3 The method for a participatory assessment (step D.2)

The proposed method involves four categories of participants: (1) a multidisciplinary team of organisers; (2) persons involved in the development and use of the model (scientists, regulators, politicians; (3) stakeholders; (4) citizens recruited thorough campaigning efforts on the local market and phone calls. The method includes three phases: (1) preparation; (2) participation; (3) analysis. The stakeholders and users of the model were involved in preparation and participation, citizens in the participatory mainly but also up to a certain level in analysis during the participatory phase. The participatory phase comprises three functional parts (presentation of the model and scenarios, stakeholders' opinions, citizens' deliberation) organised in 6 phases: (1) welcome presentation; (2) presentation of the model and scenarios and question/answers session; (3) stakeholders opinion, and question/answers session; (4) organisers analyse citizen's discourses, and elaborate questions to be submitted to citizens; (5) citizens' perspectives (discussions by-subgroups and together); (6) evaluation of the test (both citizens and organizers impressions).

In its practical implementation, the set-up took the form of a two-days workshop, with 26 participants. The persons who participated can be classified into four categories: (1) organisers of the exercise: five social and natural scientists and a facilitator; (2) persons involved in the use of the PRIBEL model (1 representative of Belgian federal authorities responsible for the reduction plan, 1 representative of the cabinet of the Minister of Public Health, 1 representative of VAR[8] i.e. the institution that developed the model and scenarios); (3) stakeholders (1 representative of farmer's associations, 1 representative of a beekeeper organisation, 1 scientist with expertise regarding bees, 1 representative of the pesticide industry, 1 representative of an environmental NGO); (4) citizens with diverse interests in the pesticide risk model and scenarios for bees: 4 consumers, 1 conventional farmer, 1 organic farmer, 1 integrated pest management farmer, 3 beekeepers, 2 amateur gardeners.

Citizens were recruited through campaigning efforts on the local market and phone calls. Since the objective of the participation is here to test the feasibility of the method, in a small scale, and particularly in the cognitive aspect of knowing if citizens are able to appropriate themselves the complex scenarios and from there to know if useful opinions can be obtained about the choices embedded in them, the statistical or democratic representativeness of the citizens is not required at that stage. Moreover, the objective was not to obtain a consensus. The criteria for selecting them was rather to obtain a diversity of points of view relatively to the scenarios, the bees, the pesticides and the regulation. Such a diversity is the condition to map the diversity of the potential opinions: it is in that sense that the citizens are representative. The fact that we have consumers, organic and conventional farmers, beekeepers and gardeners expresses the diversity of interests, while keeping the number of participants manageable.

[8]Veterinary and Agrochemical Research Centre.

The scientists involved in developing and the persons involved in using the PRIBEL model presented this model and scenarios to the citizens so that they can start with forming their own opinion. Secondly (in time), we invited stakeholders to comment and criticize the model/scenarios, so that citizens can reconsider their first opinion. The third part of the test is mainly devoted to deliberation between the citizens themselves so that they can test and adjust their opinion via a confrontation with the opinions of other citizens. The goal of this third item is not to reach consensus, but to make the various opinions and underlying arguments more explicit.

This three functional parts were organized as following during the workshop:

1. Welcome, General Presentation, 9:30 – 10:00
2. The policy goals and the functioning of the model and scenarios, 10:00 – 13:00
3. Stakeholders' perspectives, 14:00 – 16:50
4. 17h00-23h00: organisers analysed citizen's discourses, and elaborate questions to be submitted to citizens dived in sub-groups
5. Citizens' perspectives (discussions by-subgroups and together), 9:30 – 11.30
6. Evaluation of the test (both citizens and organizers impressions), 11.30 – 12.30

The whole participatory exercise was recorded in order to be further analysed. When analysing the observations of the participatory exercise, we pay attention to:

- social and political choices that are possibly made explicit by the exercise;
- opinions, advice, insights provided to the modellers, as well as to public authorities, decision makers and stakeholders;
- opinions about the technical functioning of the scenarios and model, their objective, and their use in their policy context;
- possible changes occurred to scenarios, to their regulatory use, to expertise, and to the decision-making, that are attributable to the participatory exercise;
- possible proposals for improvements or alternatives to the model and scenarios or complementary tools useful for the envisaged objectives;
- ability of the organisers and modellers to be questioned and to learn

both in terms of content (improvement of their model and scenarios), but also in terms of their positions and responsibilities in the broader context of their practices;
- the efficiency of the method to make the scenarios, the model and the whole context accessible, both scientifically and as a social construction, to lay people.

The physical set-up is presented in fig. 7.3. Since the interaction with the citizens took place during one week-end, it was necessary to make them feel comfortable, and so a kitchen and a dining room were necessary, as well as catering. During the event, the stakeholders were required not to listen to each other's presentation, because this could have given unequal opportunities to some of them to respond to each others arguments[9]. The location of the organisers table made the citizens a little bit uncomfortable at the beginning of the exercise, because they felt themselves observed, but this was rapidly solved by a presentation of each ones rules and functions on the set-up.

7.4 Results obtained when testing the participatory route

7.4.1 Quality definition and definition of the assessment purposes (steps A and B)

The initial set-up considered, regarding the model and scenarios, the presentation of the principles, the known hypothesis and choices, and some examples of applications. The objective was to evaluate the acceptability of such choices by the citizens. However, through our discussion, it became evident that an essential part of the process was devoted to understand "what the model and scenarios do". Moreover, the participatory method revealed that citizens are suggesting modifications to the model and scenarios, after having understood "what the model/scenarios do". Ideally this is a cyclic co-construction process, where modifications are made and then the consequences of the modification (scientific, social, economic,... consequences) are explored to determine again "what the model/scenarios do".

[9]Because there was no time to give a right to reply to each stakeholder. In practice some were very insistent and managed to see some of the presentations.

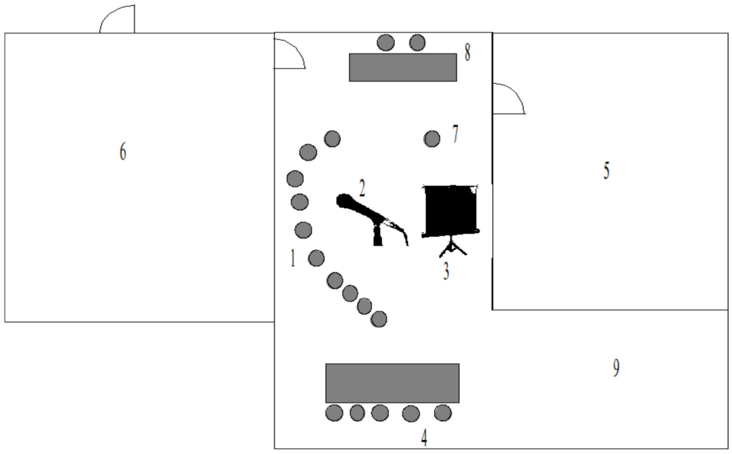


Figure 7.3: *The workshop physical set-up.* 1: citizen's seats; 2: microphone and recorder; 3: screen for slide-shows by the modeller and the stakeholders; 4: organizer's seats and table; 5: kitchen; 6: dinning-room and place where citizens were asked to discuss between them in subgroups to answer to questions framed by the organizers; 7: facilitator seat; 8: seats and table for persons in charge of the written transcription of citizen's discourses; 9: storing place for printer and other material needed.

7.4.2 Organizers and stakeholders interpreting the scenarios (step C.1)

Explaining and understanding the indicator set within the research consortium

From the beginning of the research project, we were convinced that, in order to discuss the PRIBEL indicator set with citizens, we first had to make explicit for ourselves the choices that are made within the indicator set as much as possible. The first step in the method is thus a reflexive work of the modeller (in this case, myself), trying to make explicit the choices and hypothesis he is aware of. This exercise allowed me already, even before the participatory phase, to question new elements of the model and scenarios, and to find some choices that were made implicitly.

Our research project was an interdisciplinary research project. Part of the researchers from the consortium had at the beginning few knowledge regarding PRIBEL. Since the researchers have different disciplinary backgrounds, the next step to be taken during the research process is to create a common understanding of the indicator set that is intended to be the object of the participatory exercise. Therefore, the first meeting of the consortium was mainly devoted to a presentation of and discussion about the PRIBEL model and scenarios. The slides I prepared for this were judged rather technical by the team[10] but essential. It was also concluded that my presentation to the citizens must be more accessible.

The PRIBEL model and scenarios were presented in section 1.3, thus we do not repeat it here in extenso. Topics that are dealt with during the introductory meeting of the research consortium are a) the objective of the indicator set; b) the choices, extrapolations, simplifications and selections in order to arrive workable a model with the data available; c) the aggregation problem.

Though the formulas in PRIBEL are not that complex, this does not imply that the model and scenarios are easy to understand. We experienced, that the research project was too short to reach a sufficient understanding —by the social researchers themselves —of the technical functioning of the PRIBEL indicator set.

We experienced that understanding the indicator set on a technical level, i.e. understanding the functioning of the formulas used, is not sufficient or satisfactory. In order to gain a more complete insight into the indicator set, we also need sufficient insight into (1) the (policy) context in and for which the indicator set is developed and —linked with this aspect—into (2) the reasons for and the meaning and impacts of the reductions and choices made within the indicator set. We concluded from this topic that a complete understanding of the political and regulatory context is needed. It is interesting to note that, during the participatory exercise, citizens expressed the same observation.

Discussing the model and scenarios with the stakeholders

Such a discussion was not performed in a systematic way for reasons of time limits. However, stakeholders we invited to present their perspective during the participative meeting asked themselves for a preparatory meeting with the research team in order to get more clarity regarding their role and the researchers' expectations towards them.

[10]Despite the fact that I felt at that time that it was a big simplification in comparison to my actual work.

Preparing the presentation of the model and scenarios to the citizens

Starting from the presentation of the tacit assumptions, simplifications and preconditions I made to the research team, I prepared a presentation of the model and scenarios directed towards the citizens. I took into account the remark that it had to be less technical and more accessible. The presentation was tested with a small group of relatives (no scientists).

Presentations to the citizens

Various presentations took place during the first day, with room for question from the citizens' side:

- one by a representative of VAR to present the functioning of a bee-indicator and a short history of its coming into existence;
- the one I presented about the use of that bee-indicator as a starting base to develop the model and scenarios for risks aggregation (with an account of the tacit assumptions, simplifications and preconditions);
- one by a representative of the administrative authorities in charge of the PRPB in order to present the legal context of the PRIBEL model;
- one by a representative of the cabinet of the Minister of Public Health in order to present the policy context and the political perspective;
- four others, prepared by some stakeholders (the representative of industry, of an environmental organisation, of a bee-keepers organisation, and a entomologist).

7.4.3 Citizens interpreting (step C.1), evaluating (step E) and suggesting improvements (step G) to the scenarios

Amongst other questions, citizens were asked for opinions and questions about the model and scenarios, their strengths and weaknesses, their contribution to pesticide risk reduction, and how scenarios can contribute or not to risk reduction[11]. In order to answer those questions, they were divided in two sub-groups and asked to arrive to a common text to be read by one person of each subgroup.

[11] See in Piñeros Garcet et al. (2006a) for a complete description of the questions asked to the citizens and their answers.

Group 1 was in favour of the use of the model and the scenarios, but by improving them. According to group 1, improvement can be achieved by consulting practitioners (farmers, beekeepers, people "*with knowledge about nature*"). For them, the model does not allow to reduce risk but to calculate it. More than one model must be taken into account, and one cannot restrict itself to bees (because bees may go well and at the same time groundwater can be at risk). The various models must be taken into account separately and together also, through a "*mean*" of them. More than one model for the same environmental compartment must be used, and one must make a "*mean*" of the results, in order to avoid a bias towards the model presenting the "*nicest results*". Moreover, the calculations must be done by two institutions that are independent of each other.

Group 2 said that PRIBEL has the merit to exist, but wandered if other models exist. They also wandered if ''*everything all is already done*'', if there was still a chance for adapting or modifying the model and scenarios. They also asked for the follow-up they can expect for the work they are doing. They judged that the model and scenarios are *''an oversimplification of reality''*, regarding (amongst others) aspects they consider essential, and that makes a problem for them. Group 2 asked about the correctness of the data used, especially when they are given by industry. They wandered if an additional verification is necessary, if 25 parameters are enough, if 100 or more are necessary. They also asked if an observation of the pesticide effects is necessary (on the field) and the same for an observation of the effects on the bees behaviour. They wandered if effects on the behaviour *''can remain even after all pesticide traces disappeared''*. For them, such an observation must be done in collaboration with people *"in the field"*, i.e. with on-site and practical experience. Group 2 also gave an opinion about the fact that the model is designed to arrive to one single number aggregating the various risks. They agree to start with a single number. However, *''if the model is over-simplified, some highly toxic substances will not be taken into account''*. The model must be *''adapted to the severity of the problem and to the complexity of the reality''*. There is a risk of *''let go''* toxic pesticides, like a sieve letting things go through the holes. The model is *''too simple in comparison to the complexity''*. Citizens from group 2 also questioned the scientists by asking them *''what they really think about their model''*. *''Is the method really scientific, or not?''*, or do the scientists have *''to give a number for the sake of politicians?''*. *''Is the cumulative effect of* *doses through the food chain taken into account by the model and scenarios, or is that a big hole in the model-sieve?''*. Group 2 also asked if the politicians said to the project team what they want to do with the results of the citizens evaluation of the model and scenarios. They wandered *''if really a full-scale test of the participatory method was planned''*.

Citizens suggested modifications to the model and scenarios. However, PRIBEL was manipulated by the scientists as a complex code without a graphical user interface. Moreover, parts of the data stream were still handled manually. Additionally, the model itself can take long to produce results that can be shown, and in early stages of its construction those results can take the form of huge tables or graphs difficult to communicate. All those reasons made a "real-time" co-construction of the model and scenarios difficult, and thus prevented to realise modifications directly with the citizens.

In any case, some interactivity can be achieved if the necessary time is given for a series of questions/answers between the modeller and the citizens. I mitigated (in my modeller's role) the consequences of the lack of interactivity using a slide show presenting the scenarios and the model, along with selected results, and then tried to answer "what if" questions from the citizens regarding potential modifications and results of the model when applied to various situations. This allows questions of the type ''what if I change this in the model and/or scenarios?'', ''is such a change scientifically sound?'' or ''if I apply the model to that particular situation, what will be the results?''. A brief time for that was available in our set-up (40 minutes), but it must be extended much more in order to obtain more useful inputs from the citizens.

During the participatory experiment, difficulties arose to understand PRIBEL. These difficulties are interpreted as pertaining to two categories.

The first difficulty is epistemological: models are often presented as simplifications of a more complex reality. We saw in section 3.2.1 that a gap between the model or the scenarios and ''reality'' is continuously present in the discourse about scenarios and models. This was also the case here: the extent of the gap was conceived qualitatively through various indicators, e.g. (1) the number of factors that determines an effect (an acute risk here) but for which no data are available, and thus can not be included in scenarios; (2) the existence of factors for which our knowledge is only qualitative or very approximate; or (3) effects that are not taken into account. Starting from that point of view, citizens and beekeepers were provided with a list of simplifications made into the model in comparison with the ensemble of processes known to occur, and those simplifications were justified by the fact that, despite the knowledge of more complex factors, no measurements and data sets exist to take them into account in the indicator. So, paradoxically, the simplicity of the indicator introduces a complexity in the interpretation of the results: what are the consequences of ignoring one or another factor?

In many aspects beekeepers and citizens could confront their practical knowledge to the indicator and see the simplifications (as an example, the indicator does not consider the possibility of bees being confronted to more than one pesticide at a time, so no interactions are considered). The citizens made suggestions to improve the indicator.

During the exercise, it appeared that the consequences of simplifications on the results of the risk estimation for the reduction plan are not really known. This opens the discussion to "believe or not arguments", and also induces the need of enlarging the debate to consider *a second category of difficulties in understanding the model and scenarios. Both the modeller, the social scientists and the citizens became aware that they lack insights regarding the concrete impacts that politico-administrative uses of PRIBEL will have for the various actors in particular socio-geographical situations.* Will farmers be obliged to buy pesticides that are more expensive? Will they be requested to get an intensive training regarding pesticide use? Will their licence to farm depend on their use of a complex registration system? Will they be urged to apply more and more practices derived from organic farming? Will beekeepers be challenged to change their beekeeping practices? Will the hobby gardeners be challenged to change their gardening practices? Will producers of pesticide be requested to make their R&D strategies more transparent and to account for their commercial decisions? What consequences will walkers experience when they come to the countryside? How will local plant and animal life evolve? Will the complete responsibility to adapt rest with these different actors, or will public authorities be able to organise suitable measures, incentives, jobs, services that help citizens and enterprises to adapt to new situations?

7.4.4 Using the assessment (step G): what happened after the end of the participatory exercise

The exercise was my initiative, as a modeller directly involved in the development of PRIBEL. It was presented to my institution (VAR at that time) as a test, because the participative involvement of citizens into a modelling task was seen as innovative and risky. The institution preferred to be cautious, and suggested to wait for the results in order to see if the tested set-up could be or not repeated and included in the normal research procedures of the institution. The hierarchy (i.e. the laboratory director) and colleagues were interested in the participatory process, and took part to the set-up (they presented background theory for the model). After the exercise project was completed, a project in which the results of a full scale participatory exercise would be integrated in the PRIBEL indicators, in their aggregation, and in the PRPB was considered by the hierarchy. The project's future was seen as viable only if it was carried by somebody with a particular sensitivity towards the participative way of working. I was notified, as modeller, that a follow-up for the project was acceptable only if I was ready to take charge of the project. At the same time, the institution where I worked was in a transitional phase, in which there was no guarantee that the environmental modelling would stay a core business and no permanent jobs were available. Finally, and coincidentally, I left the institution for another job, and no follow-up was launched.

One year after the end of the exercise, the laboratory director was asked to give his opinion about the exercise. He said that he was happy to participate, because it was the first time he had the opportunity to interact directly with citizens, to work with an expanded circle of people. He was pleased to see that most citizens were open-minded and ready to go beyond their own interests. He particularly remembers that, surprisingly for him, a significant proportion of citizens was interested and positive towards the farmer who was using integrated pest management (and not, on the contrary, against any use of pesticide). More generally, his institution is confronted to the fact that research funding is increasingly oriented to joint projects between "hard" and "human" sciences. He considers that his research institution is open to such kind of projects, and that these projects help in fund raising, but that the structure of his institution is not well adapted to them. Indeed, the general mission of his research institution remained unchanged and oriented mainly towards ‘‘hard sciences’’, technical research programmes, expert advise and laboratory analysis. In his opinion, it is difficult to achieve a compromise between high level scientific outputs and multidisciplinary participatory methods. Finally, the participatory project did not change significantly the way they are working in his laboratory. Moreover, he sees the potential function of such kind of projects in the communication and vulgarisation field, because the scientific outputs for the core business of his laboratory seemed very limited to him.

7.5 Discussion: assessing the quality assessment (step F)

7.5.1 Discussion from a participatory practice perspective

The risk was present that some citizens act as stakeholders defending a pre-defined fixed position, and thus conducting to lapse into a polarised discussion. The present method did not take that into account, fortunately the group interaction allowed to avoid that. One participant clearly expressed, both verbally and physically, that he was only prepared to participate as a defender of a predefined position, namely that of a bee-keeper who accuses scientists of not recognizing particular pesticides as the main cause of bee mortality. Thanks to the presence of another bee-keeper with a more open attitude, this position could be questioned and discussed. Suggesting mechanisms to avoid this is not the subject of this thesis, but that is considered in Piñeros Garcet et al. (2006a), that suggests that persons with a direct concern should be invited to act as witnesses, i.e. to share their concrete experiences and the links between these experiences and their particular concerns. Persons without direct interests are invited to formulate more general, public concerns. Both groups of participants are, then, asked to balance the (rather short term) particular concerns of the local stakeholders against (rather long term) public concerns and to suggest ways to accommodate both.

We experience that the research project was too short to reach a sufficient understanding by the social researchers themselves - of the technical functioning of the PRIBEL model. More time, concrete illustrations of how the indicator set functions and necessity of interactions with the model (through some sort of interactive interface) will have to be taken into account for a future exercise.

A more extensive, preparatory analysis of the policy context will help the research team to define better the substantial input that is needed for the participatory exercise and to give more precise instructions to the various persons (experts of the indicator set, members of administrations, responsible politicians, ...) regarding the topics they should reflect on, present and discuss with the citizens during the participatory exercise. Moreover, a more extensive, preparatory analysis of the policy context will help the research team to define (and to explain to the citizens) more precisely what are the stakes of the participation of the citizens (which kind of impacts their participation can have on which elements). In other words, it will help to clarify the mandate of our research project.

It is worth the effort to discuss the indicator set with various stakeholders in order to receive a fuller picture of the choices that are implicit in the indicator set (and of the reasons for these choices). Therefore, we conclude that in a future, more extended research project, a separate meeting with stakeholders that focuses in a systematic and well-structured way on their perspectives on (reasons for) choices is a useful step to prepare a participatory exercise with citizens.

7.5.2 Discussion from an expertise and modelling practice perspective

One of the hypothesis to test was the possibility of some sort of co-construction of the model and scenarios between the modeller and the citizens. With respect to this, one of the citizens compares PRIBEL to a sieve, that is to say a tool designed to select the things that passes through it according to a selected criterion. I consider this metaphor as very instructive. It allowed me to look at the model and scenarios with different glasses. This metaphor allowed me, first, to see the indicator set more as a tool to select pesticides, rather than to measure something (calculate risks) or represent a portion of the reality. Consequently, the strategies of industry and the effects on agricultural practices came on the screen. Secondly, the metaphor points to the pesticides that are not retained and that go into the environment where they can interact with each other. The initial hypothesis was that removing from the market pesticides with a high acute toxicity[12] is a progress. However, since we cannot exclude the possibility that strong toxic synergies exist between pesticides with low toxicities when taken separately (either acute or chronic), one should conceive the hypothetical possibility that using the

[12] Acute toxicity is the pesticide-related variable taken into account in the scenarios.

indicator set to discard some pesticides from the market and introduce new ones that are less toxic individually, results in raising rather than lowering risks. I became fully conscious that the model itself, when used as a tool, may introduce new potential risks (although hypothetical).

Another example can be mentioned. Thanks to the participatory exercise, I had the opportunity to go to the field to talk with beekeepers and farmers about a potential case of acute intoxication of bees by a pesticide sprayed during the day on a melliferous flowering crop. According to the scenarios, this case can not happen, because the law and the instructions of use for the product prohibits such spraying during the periods of activity for the bees. However, since the meteorological conditions are variable and the time constraints are hard for the farmers, it seems that sometimes they are conducted to do applications of pesticide during the day when bees are present on the flowers. From this case, I suggested to include the consequences of accidents and misuses resulting from pesticide use into the model for estimating the risk. However those cases are difficult to study with conventional surveys since the farmers are reluctant to talk about them, as I experienced and learned during my field trips. In consequence, an alternative study could be launched that includes a participatory study, in order to quantify those cases for modelling purposes, and try to find ways to prevent those things to happen. This could include a dialogue between farmers and beekeepers to inform the beekeepers about the farmers practices and allow them for example to change the location of the hives when spraying is done.

These examples clearly shows how the participatory exercise can change the modelling practice, and potentially the model and scenarios themselves. Amongst other things that I thought about, is to try to use the model with potential future pesticides that the industry is preparing, in order to see the consequences of the model on the industrial strategy and from there to foresee "what the model is doing". This is clearly a learning from the interaction both with citizens (see e.g. the "sieve" example cited above) and with the industrial stakeholders invited for the participatory set-up.

The discussion with citizens about aggregation methods showed both the possibilities and the limits of citizens proposals for changes to the scenarios and models. Some of the suggested ideas were scientifically sound, while others came from a good problem framing but the proposed solution was not acceptable scientifically. The responsibility of the scientist here is to translate good ideas in sound scientific concepts. Co-construction takes place also in an indirect way, through the stimulation of the scientific creativity, without being bounded strictly to what the citizens expressed.

As part of the participatory exercise, I (in my role of modeller) had to present my work to both citizens and social scientists. This implies that I had to rethink the way of communicating: both translating the message, locate my work into a more general picture than the science I am expert of (so a learning process is needed), and, more importantly, the work had to be justified from a citizenship point of view. This exercise of justification proved to be important, because in applied scientific projects like the PRIBEL project, choices are made often because "we know that it works enough, and we did not receive money for doing better". This kind of arguments are not accepted by the citizen. They answered that it is the scientists' responsibility to ask what is needed and the politicians responsibilities to provide the appropriate means, because the choices are judged by them from the point of view of the social consequences.

After the participatory event, I did, as a modeller, an intent to re-think my practice in order to take into account what was learned from the exercise. I considered the possibility to adapt the indicator (for instance by trying to link it more closely to field observations) but also my scientific activity itself (by allowing time to discuss with field practitioners and going to the field). I now consider necessary to have a direct contact with the field and the practitioners, because this gives a more holistic point of view and stimulates scientific originality in the models.

One of the aims of the exercise was to develop a methodology that could be applied to new modelling cases, and in a first step, to other models developed in the institution were I was working. That aim is to be linked to more general questions. Does the involvement of a modeller not automatically bind his institution to his conclusions or objectives? Does it make a difference to hesitate individually and to hesitate institutionally? Does a participatory (or democratic) procedure towards modelling practices influence the way those practices may be conducted in the future? What does it mean to keep track of the choices induced during the different careers of a model? Does a participatory procedure have a chance to produce hesitations in the way an institution represents itself the work of its scientists?

Our observations showed that the answer to most of those questions regarding institutional hesitations can not be separated from the local conditions were the participatory exercise is taking place. The method had a chance to have a follow-up in the scientific institution where the model and scenarios are developed. However the general dynamic of the institution did not favoured it: the environment was not any more at the core business, and thus the dossier became highly dependent on a single person as well as on the possibility of obtaining external financing.

7.5.3 Discussion from a sociological perspective

In our proposal, the generation of information starts with a process of understanding the model and its scenarios. In many aspects beekeepers and citizens could confront their practical knowledge to the model and scenarios and see the simplifications. However, paradoxically, the simplicity of the indicator introduces a complexity in the interpretation of the results: what are the consequences of ignoring one or another factor? ‘‘Understanding’’ a model and its scenarios can mean different things. We can understand a model/scenario at different levels, depending on the objectives towards which the effort of understanding is oriented. We experience that understanding the model/scenarios on a scientific level, i.e. understanding the functioning of the formulas used, the results, or what the scenarios represent, is not sufficient or satisfactory. In order to gain a more complete insight into the model/scenarios, we also need sufficient insight a) into the (policy) context in, and for which, the model and its scenarios are developed, and, linked with this aspect, b) into the reasons for and the meaning and impacts of the reductions, simplifications and choices made within the model/scenarios, including possible impacts of potential politico-administrative uses of PRIBEL (which remained partially undetermined in the period that the research project was running) on the daily lives of various actors concerned. In order to understand an indicator set, it does not suffice to know how the indicator set functions in a technical way (what are the formulas and the databases used, what are the results), or to know what it represents scientifically speaking. In order to ‘‘understand’’ the indicator set on a deeper level, we need to see its ‘social construction’, i.e. the many decisions which constitute PRIBEL and which cannot be justified purely on theoretical arguments, but which stem from a variety of scientific, political, economic and policy connections.

During the preparatory phase of the exercise, it appeared clearly that the team needed to understand in detail the model and scenarios. A presentation of the model made at that time concentrates on "what the model and scenarios represents". However, during the participatory phase, it rapidly appeared that the citizens want to know "what the model and scenarios do". This changed the way I had to present the model and scenarios, and was one of the starting points for learning something out of the process, from my point of view. "What the model and scenarios do" is a political question, and this helped me, as a modeller, in broadening my view about what was at stake when different technical options where equally feasible from the technical point of view.

Here again thus (as in section 5.1 for the representativeness), assessing a scenario implies the need to understand its use. We see also here that citizens are performing an *interpretation* of the scenarios, as it was the case for scientists in APECOP for the FOCUS scenarios, as illustrated in chapter 4.

Models and scenarios have a twofold nature: on the one hand, they are often supposed to represent a necessary reduction of reality[13] and of its complexity and, on the other hand, we encounter their persuasiveness that consists of convincing people that the results of modelling actually correspond to the reality of the studied phenomena, at least scaled down to the dimensions that were considered pertinent to be taken into account. The persuasiveness is in fact a request for confidence. There is a close link between the simplifications and choices in a model or scenario, and the request for confidence: it is because simplification and choices are made that confidence is requested. This brings the scientific model, as the result of scientific (but also political) activity (perhaps even more than for other scientific products), into conflict with the public clarification of the choices involved in its construction.

This is a point where public participation to the assessment of modelling tools presents a clear advantage. While modelling is often presented as a scientific matter, confidence has a clear social character, and this is a context where involving lay-people can lead to useful assessments of the acceptability or robustness of the models and scenarios. Our observations shows that the modeller was urged to produce justifications, in such a way that they can be accepted and appropriated by

[13]Despite the problems of such a concept of representation, as we demonstrated it in chapters 4 and 5.

the citizens. Citizens judged those justifications in terms of resources needed for the modelling work, in terms of the objectives of that work, and more generally they judge the way the modelling work is articulated to the larger context in which it evolves and is used.

The problem of confidence is also linked to the supposed necessity of confidence in decision-making. Some key-actors may try to avoid the need of obtaining confidence, by developing the ability to evolve in situations of distrust. By doing this, they may gain the monopoly of the problem's definition.

The theme of the co-construction raises also the problem of the mandate given to the citizens. This is linked to power and democracy. In our particular case, the status of the participation was a test. Politicians and the regulators were not bounded to the results. If changes in the regulatory practice are to be expected, then the way the citizens inputs are taken into account must be clearly stated before and be compulsory for the regulators and the modeler. However, the initiators of the participatory setup must be aware that citizens can put into question even the necessity to construct a scenario or a model. Moreover, a participatory exercise with a clear mandate can be seen as a loss of power of the traditional representative democracy institutions, but also for the traditional ways of negotiation used by stakeholders.

7.5.4 Discussion from a quality assessment perspective

One of the goals of our quality assessment framework is to assess scenarios in a transparent way in its multiple aspects, and to generate information about the influence of the scenarios quality on the use of the scenarios, the potential improvements to the scenarios, and the potential improvements or changes to the scenarios use.

The analysis of the participatory set-up points towards the need for an adequate equipment of the model and scenarios' narration. Such an equipment must include: (1) a scientific presentation of the model understandable by the participants; (2) an interactive interface; (3) knowledge of the history of the model, scenarios and the decisions embedded (from the scientific, economical and political points of view); and (4) a list of the impacts (economical, social) of the use of the model for a particular purpose. In order to respond to the above-mentioned recommendations, it is proposed to equip models and scenarios with two matrices that each includes the following columns: choices, reasons, concerns, evaluation, suggestions. The first matrix represents the *choices in models and scenarios* and their (economic, technical, political, ...) reasons (this is the 'deconstruction' part), the concerns of various actors related to particular choices (the 'impact' part), and suggestions to improve the models and scenarios by making different choices (the 'reconstruction' part). The second matrix represents the *the role in the regulatory framework* (here a pesticide reduction plan) and concerns, evaluations and suggestions related to these choices.

The proposed matrices are to be related to our concept of *scenarios narration* developed in chapter 1. PRIBEL scenarios' narrations existed before the participatory exercise, but they were not equipped with the necessary elements to understand their history, policy role and use. In our participatory assessment, as well as in our representativeness assessment, an interpretation is performed in order to produce a new interpretant which is the new scenario narration. However, this is done here through a participatory process, while this was not the case for our quality assessment of the FOCUS-GW scenarios.

A contradiction may exist between the need of interactivity with the model and a co-construction, since in many cases an interactive interface to complex environmental models is only feasible when the model and the scenarios are in an advanced stage of construction. The lack of interactivity is also a problem for the understanding of the model, since understanding implies also for the citizens to understand what the model does, through "what if" questions regarding the results of the outputs of the model. It is difficult to answer to such questions without actually being able to run interactively the model, although the modeller can compensate this by up to a certain extent by answering with his knowledge of the possible probable outputs of the model. This is however limited, since the model is precisely there because the modeller can not fully replace it.

One of the observations that was made was that for stakeholders and citizens appropriateness depends on the acceptability of the results and impacts. The impacts are impacts of *using this particular model and scenarios in a particular way in a particular policy context and on a particular geographical level. This acceptability depends, hence, on the context in which, and the objectives for which, the model/scenarios will be used.*

7.6 Chapter conclusion

Acknowledging the links between the political and the scientific dimensions of scenarios quality assessment through the interpretation choices that are involved there in, this chapter tested (step F of our framework: assess the quality of the assessment) the possibility to open those choices to a participatory assessment by citizens, using PRIBEL as a case study. PRIBEL articulation with a Belgian national pesticide reduction plan offered a good test case.

The method involves citizens in three main functional parts: presentation to the citizens of the scenarios by the modellers and by the regulators; presentations, comments, critics of the scenarios by stakeholders and other scientists; deliberations between citizens, and mapping of the variety of opinions and arguments. This is a contribution to step D.2 of the framework (take the links between quality components into account).

When making the balance of our results, we can say that the method demonstrated its ability to produce information regarding the influence of the scenarios quality on the use of the scenarios, potential improvements to the scenarios, and potential improvements or changes to the scenarios use, in a transparent way:

- The discourse related to simplifications in models and scenarios generates a request for confidence in order to be used. Confidence has a clear social character, all the more in a regulatory context with consequences for citizens. The results show that involving lay-people can lead to useful assessments of the acceptability or robustness of the models and scenarios, because for citizens, the quality of the model and scenarios depends on the acceptability of their results and impacts, and on the context in which, and the objectives for which, the model and scenarios are used. This point is a contribution to steps A, C.3 and C.4.3 of our framework (defining quality, assessed characteristics, and quality criteria).
- The application of the method has the potential to generate a co-production of knowledge regarding models and scenarios, and particularly an understanding and assessment of the social impacts of the model use, of the history and reasons of the decisions and choices embedded, of the policy context (this is a contribution to step C.2 define uses and purposes, step E.2, intended uses assessment). The application of the method may lead to proposals for changing the scenarios

use (step G.2). When the modeller and its institution are involved, the knowledge gained can be used to reconstruct the model and scenarios with different choices (step G.1, suggest improvements).

- Transparency and robustness of the scenarios can be improved through this method if the information generated is put into a form that is communicable. As a contribution to steps C.1 (interpret and document the interpretative choices) and H (include the quality assessment into narrations), we suggest the use of two matrices: one for the choices and the other for the role in the regulatory framework, both describing the choices, reasons, concerns, evaluation, and suggestions of improvements.

The method encountered some limits and difficulties. The multidisciplinary team had difficulties to understand the model, the scenarios and their consequences, the policy context, and the possibly conflicting objectives between stakeholders. This shows that more time is needed for the preparatory phase. Moreover, a recursive iterative re-construction of the model and scenarios with the citizens seems to be difficult if the time available is short, mainly for technical reasons: models and scenarios are difficult to manipulate in their construction phase (no graphical user interface, long time to implement the changes, big tables, raw data, etc). This opens possibilities for research into programming interfaces and languages that can allow such interactivity.

It must be mentioned here also that often regulatory models and scenarios' building is performed as "projects" financed in the short term, without giving the opportunity to include participation and interactivity, or the development of a graphical user interface. In the present case, participation was financed from "drawer bottoms" by a different source, complemented by funds from a social foundation interested in participation promotion. The financing, although essential to the project and very welcome, was not sufficient to develop this interactive aspects.

Our test results suggests that the success of the method is dependent on local conditions (the modeller's will, the institutions objectives, etc). If the method is seen as an innovation in Latour's sense, it means that it must find allies in order to become institutionalised. A method can not by itself generate changes into the way expertise and decision-making is made, the necessary allies for the presented method must still manifest themselves.

Our results shows also that citizens are not radicalized against the use of pesticides by the participatory deliberations despite the uncertainties in the scenarios; they rather remain coherent with their way of performing the assessment: they judge the scenarios also in the light of the possible advantages of pesticides for them. This can be seen as a balance between risks, uncertainties and benefits. A difference with the classical risk and benefits analysis is that usually the benefits are not necessarily balanced for the same persons that takes the risks, while here citizens tries to go from "acceptable" risks as defined by experts, to *accepted* risks, uncertainties and benefits for them.

While various studies addresses the use of models into a participatory decision making, they use the models as one of the tools rather than the object of the participation. The test performed here is quite original in that sense. However, the action of performing a participatory quality assessment does not improve by itself the scenarios: the present study shows that the quality assessment of models and scenarios in a participatory setting, does only have a meaning when co-construction of the scenarios is envisaged, otherwise the citizens incurs in the risk of being used as foils. This is a contribution to our steps B, define the assessment purpose, and step G, use the quality assessment.

It is necessary to stress that there are limits as to the inferences that this study supports. It is a test with heuristic and cognitive objectives, without a political or regulatory mandate for the citizens. It gives insights about the way they understand the scenarios, an about the criteria they use for assessing them. As it is presented, the method is useful if the scientists themselves are ready to take into account possible changes to their scenarios. However, the extrapolation of the method to situations that binds the politicians or the regulators to the results through a mandate is still to be studied. Extrapolations towards an implementation of the four phases of Latour's collective functioning (perplexity, consultation, hierarchisation, and institutionalisation, see chapter 3) can not directly be made, except that it gives cognitive insights for the perplexity and consultation phases.

General Conclusions

Decisions based on forecasting simulation scenarios having a bad quality for a regulatory use are likely to have unwanted consequences, depending on the product that is being regulated. Our initial problem was the definition of the meaning of "the quality of regulatory scenarios", and the methods for assessing it.

The objectives of this thesis are thus normative and pragmatic: to define scenarios quality, and to implement concepts, methods and means for a quality assessment of forecasting regulatory simulation scenarios, and to test their applicability, limits, and interests.

The approach we adopted for the demonstration is developed in three stages: (1) a proposal of a generic framework for scenarios quality assessment, that includes a definition of quality and generic steps for performing quality assessments; (2) the consolidation and implementation of the generic framework with conceptual developments and observations relevant for regulatory forecasting simulation scenarios; (3) the test of the framework with two scenarios used as case studies (the FOCUS groundwater scenarios and the PRIBEL bees exposure scenarios).

The definition of our quality assessment framework started with the definition of the meaning of quality and quality components. Scenario quality is defined as the adequacy of the characteristics of a scenario to given use(s) or purpose(s). A quality component is the set composed by a characteristic, a use, and an adequacy definition.

Generic steps for quality assessment were defined, they include the definition of quality, of the assessment purposes, of the quality components, of the links between quality components, the assessment in itself, as well as the assessment of the quality of the assessment, its use, and they specify the role of interpretation and narrations in quality assessment.

In the implementation stage of our demonstration, a definition of relevant quality components of regulatory simulation scenarios was provided, with their corresponding characteristics and adequacy criteria: scientific criteria for scenarios quality assessment (representativeness, theoretical consistency), as well as criteria for their political, juridical, and logical quality assessment. We showed that all criteria are based on an interpretation of the scenarios in a particular use context.

An anthropological observation of a scenarios quality assessment project concluded that regulatory scenarios are actants with a very fluctuating identity, the identity being their representing concept definition. The representing concept results from the translation. This leads us to describe scenarios in risk assessment as a border object. The scenario representing concept viewed as a border object, is able to link quality components through the process of interpretation that results from multiple translations. They are polysemous in order to play their role: to be a communication point between networks using different languages. The quality components are thus not independent, e.g. different scientific quality levels may be obtained, depending on initial interpretation choices that are equally scientifically justifiable, and thus have a political component since they may impact the regulatory decision process. In consequence, the link between quality components must be taken into account in the quality assessments.

The process essential to the work of interpretation is the social translation, defined according to the actor-network theory and the sociology of translation. This means in our case that the identity of a scenario, which is the definition of its representing concept, is negotiated between actants in successive steps until it stabilizes at least for the duration of the assessment project. The interpretation of the use of a scenario, and the use itself results also from translation. The identity of an actant should not be understood in our conception as "its being as such", but rather in relation to our research question: the identity that is of interest is ‘‘what it is’’ in an assessment context, i.e. its definition in that context. It is worth noting that the question of the being of the scenarios, as such outside the assessment process, may be interesting, but it is left in brackets here as considered metaphysical rather than pragmatic.

More generally, our framework involves interpretation in many of the steps of quality assessment because to those steps correspond questions that, when answered, involve interpretation. We advocate that scientists assessing scenarios should acknowledge the existence of interpretation and take it into account, otherwise they run the risk of being the plaything of political interests, or to mislead the public about the conclusions that can be drawn from their scientific work. We proposed a series of tools to take interpretation into account in quality assessment, based on a concept of representativeness, hermeneutics, criteria and norms of interpretation, intersubjectiviy conceptions, transparency, and scenarios narrations.

Our concept of scientific assessment of representativeness starts from an initial condition: the scenario interpretation. This concept allows room : (1) for a constructed interpretation of the scenarios (through translation involving actor-networks from the political, economical, and scientific fields) into a scenario representing concept; (2) for the establishment of a referential link between the scenario representing concept and the phenomena; (3) and, if needed, for the establishment of a link between the scenario representing concept and a mathematical structure that may provide access to the empirical adequacy of the representation, and from there to the phenomena.

Two quality assessment routes are identified and equipped with the tools of our framework: a quality assessment through the expert route, equipped with transparency through a justified documentation of the interpretation of the scenarios, and a quality assessment through the participatory route, that opens interpretation of the scenarios to citizens.

The next stage in our demonstration is a test of the assessment framework with the two identified routes, using case studies. It must be stated that the results of the tests are to be divided in two categories: the results about the quality of the scenarios themselves, and the results about our methods for quality assessment.

The expert route, supplemented with transparency in the interpretation choices, was tested for a representativeness assessment of a regulatory forecasting scenario, taking the FOCUS scenarios as a test case. The assessment results concerning the scenarios concluded that probably the representativeness of the FOCUS scenarios is low for five out of the height that were assessed. This is per se an important result due to the actual regulatory use that was made of those scenarios.

Our test case for the expertise route showed also that an experimental assessment of the quality of the representativeness of the FOCUS scenarios is not possible, that the simulations used in the assessment have on the one hand intrinsic uncertainties that are partly quantified, but on the other hand they have unquantified uncertainties due to the quality of the databases that are used, databases for which there is no better alternative. In consequence, the results of our assessment of the representativeness of the FOCUS scenarios is the best estimation available, but remains uncertain. It is important to stress here that the uncertainties regarding the representativeness of the FOCUS scenarios do not come from limitations in our quality assessment framework, but rather from the scale at which the assessed scenarios are used, that prevents an experimental approach that our framework can also take into account whenever it is possible.

The exercise showed also that the legal basis of the regulations leaves a margin of interpretation for the expert, between equally scientifically justifiable options that can lead to different assessment results. We believe that such a situation where the scientific quality of the scenarios is uncertain is common to many regulatory scenarios that relate to large scale environmental problems for which no experimental basis exists at the scale they are applied. The point is not to say that such scenarios are not useful: we needed them, as our best guess, if (in our example) the marketing of pesticides is accepted. Our opinion is that while FOCUS scenarios are useful as a way to extrapolate experimental results in view of providing some indications about the existence (or not) of safe uses on a significant area in the EU, the uncertainties about their representativeness should prevent their use in a regulatory tiered process to decide if lysimeter or field experiments are needed.

The test of our framework with the FOCUS case shows that the proposed concepts and methods for representativeness assessment are currently operational, and that they allow a high degree of transparency. The framework lead to an evaluation of the influence of the quality on the use of the scenarios, to suggestions for improvement of the scenarios, and also to suggestions for changes to the scenarios use.

We advocate however that the scientific expertise route must be supplemented with other approaches. The consideration of other dimensions than the scientific is needed both for assessing the interpretation choices, and for assessing acceptability of the uncertainties.

Various interpretations are possible when defining the assessment criteria, even for the scientific quality component. The expert might conceptualize some of the alternatives, but one should try to maximize the number of alternative interpretations before making a choice. Intersubjective consideration of an object is a possible way to vary the points of view, and from there to increase the number of potential interpretations. It is with that cognitive question in mind that the test we made of the potential of participatory methods must be evaluated.

A participatory assessment, opening scenarios interpretation choices to an assessment by citizens, was thus tested with the PRIBEL scenarios.

The test results give insights about the way citizens understand the scenarios, an about the criteria they use for assessing them. Citizens assess the quality of scenarios, including the scientific aspects like the representativeness, through the use of the scenarios: acceptability of their context of use, objectives, results, and impacts. The results show that involving lay-people can lead to useful assessments of the acceptability or robustness of the models and scenarios. The application of the method has the potential to generate a co-production of knowledge regarding models and scenarios. Transparency and robustness of the scenarios can be improved through the use of two matrices: one for the choices and the other for the role in the regulatory framework, both describing the choices, reasons, concerns, evaluation, and suggestions of improvements. The method showed, and this a result of the assessment, that there is a lack of knowledge about those aspects for PRIBEL, and thus that the impact of equally scientifically justifiable choices in the scenarios are little known.

The work presented here has the potential for generating fears about a weakening of the scientific method through participatory methods, or by introducing interpretation. Foremost is the fact that this work is a way to take into account, rather than to hide, the already existing links between the scientific and political fields when they meet around regulation, as shown in our anthropological observations. Then the reader may think that we mean that the content of scientific knowledge should be submitted to a democratic deliberation: this is not what we are proposing. The holistic view of quality that we developed is needed in order to understand how scientific methods and reasoning can be linked to a political or juridical interpretation of the scenarios, while remaining scientific. The case of the functional representativeness showed that a scientific assessment of the representativeness takes a scenario concept (i.e. a statement, an interpretation of what the scenarios represents) as a starting point, as an initial condition. From that point begins the scientific work of assessing the representativeness. But such an interpretation is, in the case of regulatory scenarios, intrinsically political, and linked to other dimensions of the scenarios (e.g. juridical, or social acceptability). By virtue of that link, that interpretation, that can not be questioned from the scientific point of view, can be questioned politically, juridically, democratically. In our method, thus, scientific knowledge is not submitted to participatory deliberations or to interpretations, but rather linked to them for specific purposes: investigate and make transparent a wider range of possible and equally scientifically justifiable interpretations of the scenarios. Transparency is also useful strictly from the scientific point of view: assessments of the same scenarios are only comparable and repeatable if the initial interpretation on which they are based is explicitly documented. Moreover, by doing assessments that considers a wide range of interpretations, the scientist is precisely in a position to avoid to be used as a political instrument through the choice of a unique interpretation. He will rather remain in its role by investigating the scientific implications of such choices. If however the regulatory context and the negotiations around it leads to the choice of a unique interpretation, the method at least requires the justification of such a choice.

Epistemologically speaking, the framework developed here is not a discovery in the sense of the discovery of a preexisting fact, but rather the construction and proposal of a method. As such, methods are judged by their effects, and thus their generalization can be conceived in the sense of an adoption by the practitioners because it places them in a better position to act. On the contrary, our anthropological and philosophical developments have a general scope in the sense that they respectively describe regularities found in many cases where science and politics meet each other, and they contribute in conceiving the use of scientific representations in general and the role of interpretation in science.

We conclude that we have settled strong methodological, anthropological, philosophical, and scientific bases for conceiving the quality assessment of regulatory scenarios. By the development of our framework, it is now possible to assess the multiple quality components of the scenarios, acknowledging and taking into account their links including the links between science and politics, in a cohabitation that preserves the specificity of each approach. The assessment framework and methods proposed are applicable to complex regulatory forecasting simulation scenarios. The framework is useful for indicating their adequacy to a given use, and renders explicit the embedded choices in scenarios, shows the interest of additional criteria, highlights the limits of the scientific knowledge when used in a regulatory context, highlights the uncompressible uncertainties remaining, and makes them transparent,and proposes ways for assessing the social acceptability of such uncertainties and of the use of the scenarios. The method gives insights about the influence of the scenarios quality on the use of the scenarios, the potential improvements to the scenarios, and the potential improvements or changes to the scenarios use.

We see various limits to the implementation of the framework as it is now. There is a difficulty of appropriation of a representativeness concept conceived in a symbolic objects mathematical formalism. Some of the quality dimensions were not explored in-depth, as the juridical, logical and theoretical consistency ones, or were not explored at all like the economical one. Citizens assess the scenarios according to their consequences, and thus the discussion comes rapidly to "what if" questions about the consequences of changes suggested to the scenarios when used with the model. The possibility to answer to those questions is limited by the difficulties to modify complex scenarios in "real time", i.e. during the participatory exercise itself which is often limited in time. Moreover, if they are not linked to a political mandate, both the transparency in the expertise choices and the participatory assessment remains mainly of interest for knowledge as their results may or may not be used by scientists or regulators. The possibility of the incorporation of the framework in formal democratic processes is still to be demonstrated.

Perspectives

While our results suggest a number of improvements that converge in the same direction than the one that is explored here, perspectives that go in new directions can be suggested. The definition of quality could be shifted from a normative one as presented here, to a definition in terms of the characteristics of scenarios (and models also) that explain their success in terms of adoption by scientists or regulators. The work presented here displays an important number of aspects of models and scenarios that are not covered in the academic courses, despite that students in science and engineering will potentially be confronted with requests for simulations in their future career. It is suggested that multidisciplinary approaches are integrated in systems theory and modelling courses. Models could be developed that explore the impacts, on the regulatory decisions, of uncertainties and alternative interpretation choices in environmental simulations, or at least sensitivity analysis of the simulations to alternative interpretation choices could be conducted.

We stated that the main conclusions of our participatory test are cognitive, for mapping possible interpretation choices in an assessment. The participatory test that we conducted had no political mandate, despite that regulators and politicians were present. Moreover, the period after the participatory test showed limited changes in the expertise of the scientific institution where it was conducted. The possibility, effects, added value and means for an integration of participatory scenarios assessments into regulatory, scientific and democratic processes is thus an open question. However, our work showed that quality assessment has a normative and prescriptive aspect, and as such, norms and prescriptions may be submitted to participatory democratic deliberations. A simple procedural justification can be given to this: deliberative democracy is a concrete expression of a democratic ideal. More interesting is the fact that participatory deliberative procedures can also be considered in an instrumental perspective, as tools, and this leads to the question of their empirical effects: it is in that sense that the extrapolation of the method to situations that binds politically to the results is thus still to be studied. The potential number of situations to which it can be applied is important since simulation scenarios are increasingly being used in various regulations in addition to the pesticides one, including the EU regulations of chemicals, biocides, the EU water framework directive, etc. Maybe a realistic mandate in the present regulatory situation could be that regulators and politicians engage themselves to at least give a justified reply to the participatory exercises stating why and how they took (or not) into account the results.

Amongst the potential obstacles for the adoption, we already mentioned the fears about relativism, that in fact is at odds of our proposal that searches for a cohabitation of various approaches including the scientific one with its own specificity. But there are other potential obstacles, like the fear to be confronted to the public, doubts, uncertainties or political positions about the benefits of transparency, the polarization of positions resulting from the so called "science wars" (the Bricmont and Sokal affair) that prevents a depassionate debate. We must also mention the potential displacements in power positions that could result from the application of the kind of methods we are proposing: Which will be the status, or the reason for the existence of scientists and expertise in a new configuration where their interpretations are discussed by lay-people? We advocate that the links we render explicit between science and other dimensions already exists, and that a win-win situation between science and democracy is possible if the method is made accessible with an appropriate communication, if the authority of knowledge is recognized, but not as the only argument. In other words, an accepted, negotiated place is to be found to both participation and science. We believe that such a win-win situation is achievable, at least in the case of regulatory scenarios, by stating that science can start its own work taking as an initial condition a participatory exploration of the interpretations of the scenarios between various equally scientifically justifiable options. Such a win-win situation requires also that the limits of science are recognized and made transparent when large uncertainties remains, and that some margin of independence is found in the actual convergence between funding bodies, industrial interests and science. Science is thus to be complemented and not replaced by other approaches. However, in a spirit of good resources allocation, a good balance between no participation and too much participation is also to be achieved.

Appendices

APPENDIX A

FOCUS scenarios legal basis, and FOCUS guidance

A.1 Directive 91/414/EEC, Annex II

The 91/414/EEC Directive states, in annex II (entitled *"Requirements for the dossier to be submitted for the inclusion of an active substance in Annex I"*):

> *§ 7.1.3.3*
> *Expert judgement will be necessary to decide whether lysimeter studies or field leaching studies should be carried out, taking into account the results of degradation and other mobility studies and the predicted environmental concentrations in groundwater (PEC_{GW}), calculated in accordance with the provisions of Annex III, Section 9. The type and conditions of the study to be conducted should be discussed with the competent authorities.*

A.2 Directive 91/414/EEC, Annex III, Requirements for the dossier to be submitted for the authorization

Annex III "Requirements for the dossier to be submitted for the authorization of a plant protection product", part A "Chemical preparations" states that:

> *§ 9 (v) Fate and behaviour in the environment.*
> *Justified estimates must be made of the expected concentrations of the active substance and relevant metabolites, degradation and reaction products, in soil, groundwater, surface water and air, following use as proposed or already occurring. In addition a realistic worst-case estimation must be made.[...] For the estimation of these concentrations all relevant information on the plant protection product and on the active substance must be taken into account.*
> *When models [sic] are used for estimation of predicted environmental concentrations they must:*
>
> – *make a best-possible estimation of all relevant processes involved taking into account realistic parameters and assumptions,*
> – *where possible be reliably validated with measurements carried out under circumstances relevant for the use of the model,*
> – *be relevant to the conditions in the area of use.*
>
> *§ 9.2.1 Estimation of concentrations in groundwater.*
> *The groundwater contamination routes have to be defined taking into account relevant agricultural, plant health, and environmental (including climatic) conditions. Suitable estimations (calculations) of predicted environmental concentration in groundwater PEC_{GW}, of active substance and relevant metabolites, degradation and reaction products, must be submitted. PEC estimations must relate to the maximum number and highest rates of application, for which authorization is sought. Expert judgment is required to decide if additional field tests could provide useful information. Before performing these studies the applicant shall seek the agreement of the competent authorities on the type of study to be performed.*

A.3 Directive 91/414/EEC, Annex IV

The 91/414/EEC Directive states, in annex VI that:

> *No authorization shall be granted if the concentration of the active substance or of relevant metabolites, degradation or reaction products in groundwater, may be expected to exceed, as a result of use of the plant protection product under the proposed conditions of use, the lower of the following limit values:*
> *(i) the maximum permissible concentration laid down by Council Directive 80/778/EEC (1) of 15 July 1980 relating to the quality of water intended for human consumption*[1], or
> *(ii) the maximum concentration laid down by the Commission when including the active substance in Annex I, on the basis of appropriate data, in particular toxicological data, or, where that concentration has not been laid down, the concentration corresponding to one tenth of the ADI laid down when the active substance was included in Annex I*
> *unless it is scientifically demonstrated that under relevant field conditions the lower concentration is not exceeded.*

A.4 Generic guidance for FOCUS groundwater scenarios, Version: 1.1

This section presents some excerpts of FOCUS (2002).

> *The remit of FOCUS was to develop consensus amongst the Member States, the European Commission, and industry on the role of modelling in the EU review process of active substances. The FOCUS organisation consists of a steering committee and working groups. The working groups consist of experts from regulatory authorities, from industry and from research institutes. Guidance was firstly developed for leaching to groundwater (FOCUS, 1995) and later for soil persistence and surface water (FOCUS 1996 & 1997). The guidance developed by the work groups included a description*

[1] 0.1 $\mu g/l$.

of the relevant models and their strengths and weaknesses. Any PEC model calculation assumes a scenario which is therefore an important element of the guidance.[...]

Standard scenarios are needed because they increase the consistency of the regulatory evaluation process by minimising the subjective influence of the person who performs the PEC calculation. Standard scenarios also make interpretation much easier, and <u>enable the adoption of a consistent scientific process for a Tier 1 evaluation of the leaching potential of substances at the EU level</u>. Therefore the FOCUS Work group for Groundwater Scenarios was charged in 1997 by the FOCUS Steering Committee with developing a set of standard scenarios which can be used to assess potential movement of active substances and metabolites of plant protection products to groundwater as part of the EU process for placing active substances on Annex I. Since this process proceeds at the community level, the standard scenarios have to apply to the whole EU. (FOCUS, 2002).

The intended framework within which the scenarios would be used [...]. All relevant scenarios (but not all models [...]) would be run for every active substance as a standardised Tier 1 assessment of leaching potential. In this context the relevant scenarios are defined by the intended use of the substance [...]. <u>The purpose of this assessment would be to establish if a "safe" scenario exists which is relevant for use of the substance</u>. If one or more of these relevant scenarios result in predicted groundwater concentrations less than 0.1 µg/l, then in principle the active substance could be included on Annex I (with restrictions on its use if necessary). The Member States would then further assess the leaching potential of the relevant plant protection products under their own conditions in the process of national authorisations. In addition to modelling, there is also a role for lysimeter or field studies and monitoring data at higher tiers when these data exist. (FOCUS, 2002)

What are the FOCUS-GW scenarios representing? According to FOCUS (FOCUS, 2002):

> *Vulnerability of ground water to contamination resulting from the use of an active substance is represented by nine realistic worst-case scenarios. Collectively, these represent agriculture across Europe, for the purposes of a Tier 1 EU-level assessment of leaching potential. The scenarios do not mimic specific fields, and nor are they necessarily representative of the agriculture at the location after which they are named or in the Member States where they are located. The purpose of the standard scenarios is to assist in establishing if "safe" scenarios exist which are relevant for use of a substance. Since they form Tier 1 of the assessment, they have been defined to represent a realistic worst case. [...] the nine scenarios should be viewed collectively as representing major agricultural areas in the EU. [...] FOCUS groundwater scenarios are virtual, in that each is a combination of data from various sources designed to be representative of a regional crop, climate and soil situation. As such, none can be experimentally validated. [...] The combinations of crop, soil, climate, and agronomic conditions should be realistic. The scenarios should describe an overall vulnerability approximating the 90th percentile of all possible situations (this percentile is often referred to as a realistic worst case). The vulnerability should be split evenly between soil properties and weather. [...] Locations were selected by an iterative procedure with the objective that they should: represent major agricultural regions (as much as possible); span the range of temperature and rainfall occurring in EU arable agriculture; be distributed across the EU with no more than one scenario per Member State. Some of these added scenarios are not located in major agricultural but they represent areas with a significant percentage of arable agriculture in the EU, albeit diffuse. [..] The selection of the soil was based on the properties of all soils present in the specific agricultural region represented by a location. Thus unrealistic combinations of climatic and soil properties were avoided. The intent was to chose a soil that was significantly more vulnerable than the median soil*

in the specific agricultural region, but not so extreme as to represent an unrealistic worst case. [...] target values for soil texture and organic matter were developed for each location to ensure that they were more vulnerable than the average. .

APPENDIX B

Quality Assurance Project Plans for Modeling

EPA (2002) is devoted to modelling in general, but gives indications about ‘‘inputs’’ and ‘‘parameters’’ which in fact are scenarios. They prescribe the following:

- Parametrisation documentation should be sufficiently detailed to understand the outputs of the model.
- Peer review of parametrisation should be done before model coding.
- Performance criteria should be defined for inputs before applying the model; the intended use of the data should be described; the acceptance criteria for the use of such data in the project and the specification of any limitations on the use of the data should be defined. Expression of data quality relatively to tolerable levels of potential decision errors should be done. The definition of performance or acceptance criteria is usually an iterative process that should at least involve modelers and users.
- Data quality indicators, which measure features of data quality such as precision (i.e., variability in data under prescribed similar conditions), bias (i.e., systematic error), accuracy, representativeness, completeness, and comparability should be used. An examination of how the sampling method can affect it should also be done.

- Examination of relevance and appropriateness should be done.
- Bias should be considered. Would any characteristics of the data set directly impact the model output (e.g., unduly high or low process rates)? For example, has bias in analysis results been documented? Is there sufficient information to estimate and correct bias? If using data to develop probabilistic distributions, are there adequate data in the upper and lower extremes of the tails to allow for unbiased probabilistic estimates?
- Precision should be considered. How is the spread in the results estimated? Is the estimate of variability sufficiently small to meet the uncertainty objectives of the modelling project.
- Qualifiers should be considered. Have the data been evaluated in a manner that permits logical decisions on the data applicability to the current project? Is the system of qualifying or flagging data adequately documented to allow data from different sources to be used on the same project (e.g., distinguish actual measurements from estimated values, note differences in detection limits)?
- Summarization should be considered. Is the data summarization process clear and sufficiently consistent with the goals of this project (e.g., distinguish averages or statistically transformed values from unaltered measurement values)? Ideally, processing and transformation equations will be made available so that their underlying assumptions can be evaluated against the objectives of the current project.
- The likely consequences (e.g., incorrect decision-making) for selecting data that do not satisfy one or more of these areas (e.g., are non-representative, are inaccurate) should be considered.
- Uncertainty (i.e., lack of knowledge) should be characterized.
- The model input should satisfy all assumptions (e.g., lognormality) needed for use in the model.

APPENDIX C

Rhetoric and pragmatic study of language

Our conception of the rhetoric and pragmatic approaches is based on (Doury, 1997). For her, rhetoric is a descriptive approach of argumentation as a linguistic phenomena (as opposed to thought phenomena) found in rhetorical situations. That approach is supplemented by elements from the pragmatic approach of language (Baylon and Mignot, 1999) that considers all statements as having a purpose, and by the linguistics of enunciation that takes into account the circumstances of enunciation.

A rhetorical situation (Doury, 1997) is characterised by a question (rhetorical) which arises from the confrontation of two conflicting theses, one of them generally accepted (doxa) and the other the antagonistic proposition[1]. In this view, each person supports a line of argument that suits his interests, and the rhetorical situation is born from this conflict of interests. There is therefore a stake, this is not a pure game of logic; it is a verbal struggle. In rhetorical situations, there is a third party to

[1] Often, a rhetorical question is defined in the sence of *erotema*: "to affirm or deny a point strongly by asking it as a question". For example, <<Why are you so beautiful?>> is likely to be a statement regarding one's opinion of the person addressed rather than a genuine request to know. See http://rhetoric.byu.edu We are following here the definition used by Doury (1997).

be constructed as interested, as allied. Rhetorical speech constitutes a search for allies, through the means of persuasion: we can see from this the link with sociology of translation via the idea of the actant. The idea of the ally can be extended to non-humans in the theory of translation, which does not deal with rhetoric. However, the theory of rhetoric bears in mind the possibility that the third party is not present in flesh and blood: this is the dialogic situation in which a third interlocutor is virtually presented in the course of the interaction. So it would be a matter, for example, of identifying what this third interlocutor is, or of how it is defined by those taking part in the discussion. Classically, rhetoric identifies four phases:

- exordium: introduces the subject and attempts to interest the listener;
- narratio: objectively lays out the facts (the objective character itself being constructed by argumentation);
- confirmatio: discussion and proofs; and
- peroratio: summary and conclusion.

This can be compared in part to the four phases of translation, according to (Callon, 1986): problematisation, interessement, enrolment, mobilisation (see appendix D for a description of these four stages).

Rhetoric, pragmatics and linguistic of enunciation offers a series of useful discourse descriptors, e.g.:

- Who are the interlocutors (including the speakers virtually presented by the interlocutors)? What are their conversational territories?
- What support (written, audiovisual, etc.) do the interlocutors have available to them? What is the situation of speech (i.e. the enunciatory context), and how does it influence the exchanges?
- What is the question (for or against X), and what are the accepted and opposing propositions? What doxa is being referred to?
- What are the boundaries of argument (the established limits to the debate) that the interlocutors attempt to set up? What are admissible proofs?
- What are the issues and interests (i.e. economic, symbolic, ideological, institutional...)?
- What legitimacy do the interlocutors lay claim to?
- What kinds of argument are used (i.e. arguments from scientific authority, from hierarchical authority, logical, from competence, metaphorical, from dissociation, connection, the feelings...)?

- What type of refutation do these arguments allow or exclude?
- Who has the responsibility of supplying proof?
- What is accepted as a fact at the end of the debate? What negotiated definitions may have been reached?

APPENDIX D

The sociology of translation and actor-network theory

The sociology of translation (also called the actor-network theory) was developed mainly by B.Latour, J.Law and M.Callon. The description we give here is essentially based on (Latour et al., 1992), (Latour, 2001), (Callon, 1986) and (Vinck, 1995). Other important references for the actor-network theory are Law (1992), Law and Mol (2002) and Callon and Law (1997). In accordance with these authors, we should use a common vocabulary to speak in the same terms of both social aspects and technical aspects, of humans and of non-humans. Callon proposes using the vocabulary of translation, where translation is the process by which a socio-natural world steadily becomes established and stable. ‘‘Actant’’ is a common term that designates both the human and the non-human in a network, who mutually define their identity through translation. A network is a combination of statements, knowledge, techniques, databases, humans, etc. When the networks so constructed behave as an actor, Callon calls them the actor-network. Callon describes four stages in this process:

- Problematisation: the actant presents itself as obligatory point of passage, attributing interests to other actants, interests which define these actants (attributes an identity to them). In order to do this,

the actant formulates “problems”.
- Interessement: actions, trials of strength, attempts at forging alliances by which an actant tries to obtain a definition of the others. according to the problematisation it has itself carried out.
- Enrolment: successful interessement, with alliances formed and identities accepted.
- Mobilisation: generation of representatives who act in the name of allies, who mobilise the allies.

The sociology of translation is extended with the idea of representatives and mobilisation: this is not simply a matter of translating the statements of the former into the latter, but of reducing a group of actants to a point (punctualisation) so as to be able to mobilise them. Puctualisation is the fact that a group of actants can be considered by others as a single actant and that they interact with them as if they were a single actant. A scientific publication, for example, mobilises in this way a group of objects, concepts, texts, and laboratory experiments. This is how, through the formation of alliances, sciences and techniques build heterogeneous combinations of statements, technical methods and knowledge.

The punctualisation of a network does not imply that it is internally homogeneous: certain elements make it possible to maintain this diversity. Often, in a controversial context, the internal heterogeneity is revealed. Border objects and mediators allow heterogeneous and disconnected networks to coexist. These particular actants are polysemic, multifunctional and complex. When the networks so constructed behave as an actor, Callon calls them the actor-network.

From this perspective, a large part of laboratory work consists in producing chains of inscriptions, chains made up of a series of jumps between matter and form. The chains of inscriptions are dealt with in section H.6. The laboratory is not closed off: it is connected to networks, which extend into society and not only via chains of inscription. The extent of these networks creates the strength of a laboratory.

APPENDIX E

Analysis of APECOP in terms of interests

E.1 Problematisation in a large regulatory and economic context around APECOP

In order to understand problematisation, it is necessary to know who is speaking on behalf of who, and what is said. In consequence, in the following paragraphs, the references are given together with the affiliation of the author as it appears in the cited paper. However, as always, the persons we are referring to are speaking on behalf of they institution if and only if this is explicitly stated in the cited reference.

Russell et al. (1994)[1] stated that historically, FOCUS started in a period when models were developed by various groups but were not readily available for use by others; that most of the developments in computer modelling have occurred relatively recently at that time and therefore there were few explicit guidelines established by regulatory agencies; that regulators should be able to duplicate the modelling results submitted by registrants to ensure a common understanding of the issues; and that the establishment of uniform modelling and reporting standards

[1]From DuPont Agrochemicals Products.

will help ensure that the modelling work is conducted appropriately and will facilitate the review of the results by regulatory agencies.

FOCUS is certainly a step in the direction shown by Russell et al. (1994). He explains also that *''appropriate resources should be used to maintain, support and further develop regulatory model codes. Because of this consideration, it is likely that models for regulatory purposes [...] will primarily be supported and/or developed by regulatory agencies.''*[2]. Russell et al. (1994) also explains that many users of simulation models in both industry and regulatory agencies are not code developers, and are not able to investigate and correct errors in models, and therefore *''it is imperative that regulatory codes have high quality assurance standards''*.

In various cases, it is mentioned that models and scenarios (or parameters) must be *agreed* between industry and regulators:

- *''It is important that regulators and industry agree on appropriate models and practices'' (Russell et al., 1994)*[3]*;*
- *''In recognition of the fact that there was no agreed methodology for PEC calculation, [...] DG-SANCO set up [...] FOCUS'' (Vanclooster et al., 2003)*[4].

Such an agreement can be seen as a form of co-regulation limited by the framework of the legislation (the increasing use of co-regulation in the EU is e.g. described in Roger (2009)).

It is often stated that it is important to have scientifically ''validated'' models for using them in a regulatory framework, and industry often sees political decisions as a potential problem and wants to rely

[2]FOCUS-GW models are indeed mostly financed by public research (research institutes and universities, although some were co-financed by industry (e.g. PELMO).

[3]Russell is affiliated to DuPont Agrochemicals Products.

[4]It is worth nothing also that in other texts, FOCUS is described as a join initiative from regulators and industry (e.g.Scientific Committee on Plants (2000): *''the European Commission and the European Crop Protection Association (ECPA) jointly established FOCUS in 1993 to provide guidance to the Member States, the European Commission and industry on the role of predictive modelling in the EU process''*; Scientific Committee on Plants (2000): *''the FOCUS forum was established as a joint initiative of the Commission and industry in order to develop guidance on the use of mathematical models in the review process under Council Directive 91/414/EEC''*).

on scientific-based decisions. Many actors sees advantages in having "valid" models and scenarios:

- *Due to the complexity of natural systems, the use of simulation models will continue to be an economical, reliable and efficient means of obtaining an integrated understanding of the environmental fate of pesticides [...] more and more regulatory decisions are being influenced by the results obtained from computer simulation models. Because the significant economic advantages and versatility possible with modelling, this trend will undoubtedly increase in the coming years.* (Russell et al., 1994)[5].
- *Standardised scenarios are needed because they increase the uniformity of the regulatory evaluation process by minimising the influence of the person that performs the PEC ground-water calculation and because they make PEC calculations and their interpretation much easier for administrators, regulators and industry.* (Boesten et al., 1999)[6].
- *Modelling is an economic way of assessing pesticide behaviour under field conditions; it is cheaper and faster than field experiments. [...However...] The EU drinking water limit of* $0.1 \mu g/l$ *implies leaching of less than 0.1% of a dose of 1 kg/ha-1. At such a low leaching level, the validation status of the models is still low [...].* (Boesten, 2000).
- *The FOCUS work group objective was to develop a scheme in which the initial (or earlier) tiers are quick, simple, and cheap to undertake and allow the compounds that clearly do not cause any concern to be passed. The later (or higher) tiers are more complex and expensive but should provide a more realistic result. Therefore, results of higher tier assessments supersede results from lower tier assessments.* (Vanclooster[7], personal communication, 2008).
- *Difficulties as industry would be unable to defend itself against non-scientific criteria and uncertainties of political decision-making.* (D.French[8] in 2009, cited by Rana (2009)).

[5] Russell is affiliated to DuPont Agrochemicals Products.

[6] Boesten is a member of the FOCUS Steering group, groundwater group and APECOP, as well as a researcher at Alterra Green World Research, Wageningen University and Research Centre, The Netherlands

[7] APECOP coordinator and member of the FOCUS groundwater group, professor at Université catholique de Louvain.

[8] From Syngenta.

Industry invested in FOCUS scenarios through the work of their scientists that participated to the FOCUS working groups, as well as through various projects supporting the implementation of FOCUS activities (e.g. the development of SPADE2 soil profiles database, and the map of groundwater depth, cited in APECOP (2002)).

Divergences in scientific opinions and model results are seen as a problem by industry and regulators:

- *[...] Many inconsistencies, especially in risk assessments. [...] Results vary significantly between the [five] models.* (M.Carroll[9], cited by Rana (2009)).
- *Scientific opinion is not as uniform as policy makers would like; you can only orchestrate the network to avoid unnecessary scientific divergence.* (R.J.Coleman[10], in Hicks and Raveney (2001)).

Standardization of the risk assessment is closely linked to "regulatory costs" for industry and regulators, as clearly stated by OECD (1998) which estimates the annual savings for new pesticides putted on the market from harmonization of industry dossiers, harmonization of regulators review reports and avoidance of repeated tests to 18 million euros at that time[11]. Although they do not estimated it, they mentioned also costs avoided for industry due to delays in registrations of a pesticide that might lead to missing sales (they however cited an industrial source stating that such savings are around 15 million euros[12]).

It worth noting here that FOCUS scenarios are part of a Tier 1 assessment, supposed to represent worst cases. Additional testing is required if the Tier 1 shows concerns. This is also a saving in experimental costs.

Industry and farmers associations often express themselves about their regulatory expectations:

- *The aim of the crop protection industry is to ensure a rapid and predictable European evaluation and authorisation of new active substances; [...] decisions should be based on a technical/scientific evaluation and not criteria that are purely politically driven;*

[9] From Dow AgroSciences.
[10] DG of DG-SANCO at EU Commission.
[11] 120 million FF in 1998.
[12] 102 million FF in 1998.

given the very high cost of developing a new active substance, it is essential that an early return on investment is possible. (ECPA, 2006).

- *[To] decrease the time needed for pesticides registration*[13]. (Rana, 2009).
- *[...] There is a need for robust regulations and predictability to make that investment. [One] key strength of 91/414/EEC [is] the fact that it is based on science and risk assessment and this should be properly recognised.* (D.Castle[14], cited in Anonymous (2003)).
- *It was reasonable to expect a regulatory framework which is timely, predictable, consistent, harmonised and pragmatic, allowing for continued innovation and business planning.* (B.Julin[15] in Anonymous (2002)).
- *It is essential that regulators be seen to regulate sensibly and with sound scientific reason. There should be no further political decisions, no further obsession with "greenness" of supermarket chains and more balanced views.* (C.Wise[16], in Anonymous (2002)).

Regarding the APECOP project, various actors expressed themselves about its objectives, its potential usefulness, the way it should communicate:

- *[...] end users [are] scientists, regulators, industry people, politicians.* (APECOP, 1999).
- *The Steering Committee expressed its interest in this work. Communication between APECOP and FOCUS-SC is encouraged, so that the appropriate regulatory perspective can be discussed. A meeting could be arranged between the APECOP group and people nominated by the FOCUS-SC to discuss the findings. The results of the APECOP project have the potential to improve the FOCUS-GW scenarios in the future, provided that due process is followed.* (FOCUS Steering Committee, 2001b).
- *[...] concern that there was rigorous checking that what was being done [in APECOP] was consistent with the FOCUS groundwater approach.* (FOCUS Steering Committee, 2002a).
- *[someone from APECOP] Results of the project might in future be used to update procedures within the registration process. [...a sug-*

[13] This is linked to patents and investments.
[14] From Syngenta.
[15] from DuPont, representing the European Crop Protection Association.
[16] From the National Farmers Union - UK

gestion is made about the consequences of modifying the scenarios/ models:] try to compare final decisions, i.e. (... registration accepted or not) [...] [industry] will continue the projects on the European groundwater map and the improved soil map. The results will be made available to registration authorities in the EU member states. [...] Some of the results of the APECOP project may form the starting point for harmonised higher tier assessments. [someone from APECOP] is writing a research proposal for this. Belgium, France and the Netherlands are interested. [...] thinks that UK will be interested also. [...industry] is also interested, but cannot make any promises at the moment. [...industry] involvement could be both financial and active. (APECOP, 2002).

- *[...] End users of the project products are regulatory bodies at the European level (e.g. Standing Committee on Plant Health, FOCUS steering committee) and industry (ECPA, EPPO). They have been indirectly involved in the project through the project partnership.*
  [...] poor predictive ability in absolute terms does not invalidate the [FOCUS] models for screening appropriately active ingredients for ground-water exposure or for ranking correctly the active gradients in terms of leaching risk. However, the predictive ability for relative ranking or screening was not tested in the project.[17] *If the results [of scenarios' evaluation] are confirmed, then some of the current scenarios, in particular those for Southern Europe, should be considered for revision. The positive results for other scenarios will increase the confidence of regulators and applicants in the appropriateness of the FOCUS ground-water scenarios for lower tier ground-water exposure assessments. [...] The spatially distributed modelling approach could therefore be at the core of a higher tier refined assessment procedure which considers variability of land properties in an explicit way. [...] This should be the basis for further study and the development of scientifically based, harmonised, detailed, refined higher tier exposure assessment procedure supporting the current EU registration process.* (Vanclooster et al., 2003).
- *[someone from DG-SANCO]: If this [the APECOP results] is not*

[17] Up to our knowledge, this situation has not been significantly improved since then. According to Vanclooster (2010, personal communication), the justification for continuing to use those models lies in the fact that there is no other alternative for products that are not already on the market [editor's note: i.e. monitoring is not available and field data can not be extrapolated without a model]. Russell et al. (1994) sees also modelling as complementary to field experiments in decision making.

presented carefully, the result is that everyone will get the impression that the FOCUS models are not good enough at the [Piacenza] meeting[18] *when in fact they are fine and at present there is no alternative. This should be avoided. [someone from APECOP] The project is not invalidating the FOCUS approach, but evaluating the quality of the work given the availability of new data bases and modelling techniques. The APECOP approach may also result in suggestions for developing new pan European higher tier approaches. FOCUS is fine as a first tier. [someone from DG-SANCO]: if the higher tier predicts higher concentrations then it has no value as a tiered approach. [someone from APECOP] At present, the [APECOP approach] has not been sufficiently evaluated to decide what could be used in practice for the harmonised registration of pesticides. [Chairman]: That statement should be made at the start of the presentation in Piacenza*[18]. (FOCUS Steering Committee, 2003).

– *Given that the sensitiveness of various hypothesis is not well understood, the validation level remains low. We thus must remain extremely careful and transparent in the interpretation of the results, and thus also in the results communication, specially towards people from registration.* (In a mail from someone from APECOP).
– *[DG-SANCO] was extremely provocative. He asks us to communicate extremely carefully the results.* (In a mail from someone from APECOP).
– *I would suggest to shorten the section on ''Multiple Aspects of Scenario's Evaluation'', which is a little ''boring'', and make the section on database development, how the metamodel works in combination with the deterministic model, etc. more to the point. I think this is also the aspect that the public in [...the congress] is most interested in - the remainder is rather philosophical.* (In a mail from someone from APECOP).
– *A validation of the FOCUS scenarios without the permanent crops serves no practical/political purpose : it serves only the academic purpose of pointing at inconsistencies in the text of the FOCUS report (which is perfectly right but can also be done in another way).* (In a mail from someone from APECOP and FOCUS).

[18] XII Symposium Pesticide Chemistry, Piacenza, Italy, 4-6 June 2003.

Regarding FOCUS, various documents can also be found[19]:

- *[...] It became clear in the development of the Annexes [of the 91/414/EEC Directive] ... that mathematical modelling [... was] going to play an important role in the decision-making process*[20] *[...] modelling is more generally applicable than monitoring because monitoring gives only information for the circumstances considered [...] modelling alongside practical studies is the only route to a scientific risk assessment, the only way to ensure we share the benefits of pesticide use without unacceptable pesticides residues in groundwater. Alternative is decision-making on the basis of inadequate and non-comparable data, leading to unnecessary groundwater contamination and unnecessary restriction of safe products of benefit to agriculture. Also slower decisions and higher costs to all. [...] Each climatic zone is represented by at least one typical climatic scenario [...] a map which shows the areas where the climatic scenarios should be used*[21]*. [...] it is recommended that a special project should be started to select European soil scenarios for the use in leaching models.* (FOCUS, 1995).[22]
- *FOCUS is an independent group of experts from regulatory authorities, registrants*[23] *and government institutes/universities, which operates by consensus. FOCUS aims at technical harmonisation of complex generic environmental issues in support of the 91/414/EEC process. [...] [FOCUS-SC] liaises with stakeholders to ensure that FOCUS is effective and useful. FOCUS Work groups [...] when necessary, raise up any non-technical issues to the Steering Committee for resolution.* (FOCUS Steering Committee, 2001a).
- *[...] reminded the [FOCUS Steering] committee that their broad remit was to simplify and clarify the process for approvals in the EU.* (FOCUS Steering Committee, 2002a).

[19] Although it is out of the scope of this work, an interesting feature of the FOCUS models is that they can clearly be linked to a particular researcher through their history.

[20] See the text of the relevant annexes in section 1.2.

[21] It is interesting to see that persons that were members of FOCUS in 1995 (and wrote the FOCUS (1995) report) had different views about the scenarios at that time and when it came to validating it in APECOP (six years later), particularly about the fact that the FOCUS scenarios represent or not worst cases inside scenarios' areas

[22] It is a constant feature through this rehearsal that at the end of a research the next steps are proposed, and many of the actors are constant from 1991 to 2010.

[23] I.e. industry.

- *The FOCUS [...] developed an EU-wide representative and scientifically valid scheme for the assessment and quantification of potential groundwater contamination. It started in 1995 and was finally adopted in December 2000. The Member States mainly funded the cost of about 1.5 Mio Euro. The (unanimous) adoption of the 10 FOCUS scenarios is a big step towards harmonised risk assessment in the EU. It will lead to significant gains in time and effort in the review process and provide finally the necessary orientation for notifiers to submit uniformly formatted groundwater sections in their dossiers. This work is now being extended and built upon in a research project funded by the Commission under the Framework Programme (QLK4-2000-1238*[24]*). The full benefit of this achievement will only be felt in future reviews in coming phases of the evaluation programme.* (DG SANCO, 2001).
- *Don't get lost in the scientific detail, focus on aspects for which concrete proposals can be developed.* (FOCUS Steering Committee, 2002b).
- *[about the results of APECOP]: The Steering Committee is impressed that 6 out of 9 GW scenarios are well-represented. [...] This is good news. The [APECOP metamodel] is interesting from the scientific perspective. However, it may be used to undermine the work to date if is is over-interpreted (considering the limitations of the APECOP analyses), and this should be avoided. We should consider to update the southern EU scenarios with more recently available data. [someone from DG-SANCO]: Need to be very cautious because we need the FOCUS GW scenarios for decision making. We must not remove them without a viable replacement. We should not give up the flexibility to make 'political' decisions on the acceptability of individual substances. Not all should rest on technical guidance. The [FOCUS Steering Committee] should have the ultimate decision about how many scenarios should be passed.* (FOCUS Steering Committee, 2003).
- *The FOCUS scenarios should be used to assess potential leaching for existing substances reviewed under the 2nd and 3rd stages of the programme.* (DG SANCO, 2005).
- *The steering committee accepted to create a new 'Groundwater' work-*

[24] Author's note: i.e. APECOP.

> *ing group*[25] *in which the results of APECOP could be considered to revise some of the FOCUS procedures.* (In a mail from someone from APECOP).

The fact that NGOs are not represented in FOCUS is in contrast to similar committees in other legislations (e.g. REACH, the main and recent EU chemicals legislation) where their are invited and invest large amounts of resources to comment or even produce technical dossiers. However, NGO's argued for greater participation of public interest groups in the pesticide regulatory process and greater transparency (see e.g. Hicks and Raveney (2001)). The finding of an explanation to that absence can be a subject in itself, and the research should start by comparing the way REACH and the pesticide legislations were negotiated at the EU level and the date at which they entered into force: it can indicate an evolution of regulatory practices and NGO's methods during the last two decades. Presently in EFSA[26] the development of guidances includes public consultations (see e.g. PPR Unit (2009), where both regulators and industry asks for the development of models similar to FOCUS models).

It is also interesting to follow the debates about the 91/414/EEC Directive, its advantages and shortcomings, and the revision that was proposed by the Commission and the new text that fell into place in 2009 and entered into force in 2010. An important feature of the 91/414/EEC Directive was the review program for active substances that were on the market before 25th July 1993 (article 8 of 91/414/EEC Directive, see e.g. Chemicals Regulation Directorate (2010)). The European Commission started the review program in 1992, and the four stages (by priority) defined for the program were completed in 2009, with a significant delay (the original deadline for completion was 2003 (Howes, 2001b), and the European Commission send reminders to 12 MS in 1998 and was preparing infraction proceedings for the first stage but finally the situation was resolved (DG SANCO, 2005)). DG SANCO (2006) indicated groundwater as a major issue in the review program. A process of requiring notification of support followed by evidence of support substantially reduced the number of active substances proceeding to the detailed review programme. Over 500 substances were withdrawn as a result of these

[25] Many of the group members were also members of the previous FOCUS-GW group and APECOP.

[26] The new agency responsible for pesticide.

requirements. Once the review program was finished in 2009, many national authorisations were still to be completed (Rana, 2009).

In July 2001 a conference organised by ECPA[27] brought together industry, regulators, and some NGO's. Hicks and Raveney (2001) reported about the conference. Industry stated that:

- some of the original objectives of the Directive were *"reducing costs, speeding up approvals and improving the internal market"*, but *"the Directive could not be proclaimed a total success"*;
- from 1991 to 2001 *"ECPA members have spent some EUR 1 billion on the review programme alone"*;
- *"regulatory costs [in 2001] exceed 10% of sales turnover in the EU for [ECPA] member companies. By 2002, for every EUR 10 spent on R&D, an additional EUR 3 will be spent on regulatory costs."*;
- *"any additional registration or regulatory requirements could affect industry's competitiveness."* Jérôme Péribère, VP at Dow AgroSciences for Europe, reinforced this view, and asked for an *"acceleration of the review process"*.

Those statements were made despite the fact that improvements took place in 2000 and 2001 with a significant increase of the number of decisions, in comparison to the 1994-1999 period (Howes, 2001b).

During the discussions of the 91/414/EEC Directive revision, the subject of "zonal authorisations" was mentioned many times. It was finally included in the new (EC) No 1107/2009 regulation, and justified therein by the following: *"The principle of mutual recognition is one of the means of ensuring the free movement of goods within the Community. To avoid any duplication of work, to reduce the administrative burden for industry and for Member States and to provide for more harmonised availability of plant protection products, authorisations granted by one Member State should be accepted by other Member States where agricultural, plant health and environmental (including climatic) conditions are comparable. Therefore, the Community should be divided into zones[28] with such comparable conditions in order to facilitate such mutual recognition."*. Industry sees it as a contribution to increase the "speed to market" (Rana, 2009). There is possibly a link between this zones and the

[27] The conference was entitled: "Review for the Future – Getting the Best out of Directive 91/414/EEC".

[28] North, Central, South.

discussions that took place in FOCUS as suggested by the discussions in the FOCUS Steering Committee (FOCUS Steering Committee, 2003): *"National registration of products ends up as a higher tier. Then divergence starts. Parameter selection is different from country to country. So a new modelling study is required for each country. This creates huge workloads without giving much clarity. Harmonization of approaches (including higher tier) would be very welcome"* and also to the discussions with industry about the fact that FOCUS scenarios do or not represent areas(APECOP, 2002): *"There is no consensus on the interpretation of the FOCUS scenarios. (This was also noted at the EU modelling workshop in Lisbon*[29]*). Persons have different views (even between former members of the FOCUS Working Group). The definition of FOCUS areas - as used in the APECOP project is not subscribed by several people; i.e. they deny that the FOCUS scenarios are representative of the realistic worst case within the areas defined by the climatic limits and claim the term is misleading."*. However, the new FOCUS groundwater working group took the same approach as in APECOP: *"To revise the scenarios, the current work group had to decide the vulnerability criteria for revision of these scenarios. The criterion selected was the 90th percentile leaching vulnerability for the climatic zone represented by the respective scenarios."* (FOCUS, 2009).

The link between the APECOP and FOCUS discussions and the zonal authorization deserves more investigation. We traced back the discussions up to 2002, where a workshop organized by the EU Commission about the amendment of Council Directive 91/414/EEC (DG SANCO, 2002)[30]. At that time it was mentioned that *"zones should be defined in a scientifically sound way"* and that *"coherence should be established between the Water Framework Directive and Directive 91/414/EEC concerning evaluation, for example zonal authorisations may not be possible if decisions on PPP use are made at the level of the catchment area"* (DG SANCO, 2002). However, four years later, DG SANCO presented the zonal approach as having an objective towards improving the internal market and innovation[31] (Coggi, 2006). The zonal approach taken by the new EU regulation and the APECOP scenarios' areas are based on two

[29] The Lisbon Workshop is a scientific conference about modelling and scenarios. Industry and universities representatives, amongst others, gather there.

[30] According to Commission of the European Communities (2006), this was the first stakeholders consultation.

[31] *"Zonal mutual recognition accelerates access to bigger market"*.

very different logics. The zonal approach includes national boundaries because they are linked to a regulatory process (national authorizations) with market implications, while APECOP took not that aspect into account and went to a scientific approach in terms of environmental stratification and typology[32]. It must be said here that the zonal approach in the new regulation is used for mutual recognition, while FOCUS groundwater scenarios were conceived in 2000 for EU wide authorisation (inclusion in annex I of 91/414/EEC Directive). However, the new FOCUS groundwater work group is developing scenarios intended to be used both for national and EU registrations[33], but the proposal seems not to include specific provisions for the zonal authorizations. Voluntary cooperation between member States has taken place before the new zonal approach, and includes considerations about groundwater scenarios[34]. Regarding how exactly the zonal approach will be implemented, the preferred solution for industry is the harmonisation of data requirements and risk assessment procedures. They explicitly expressed themselves about the scenarios (ECPA, 2009): *''agreement should be reached at least at the zonal level on models and input parameters to be used for calculations''*.

[32] Such kind of approaches were recently reviewed by Hazeu et al. (2010). The approaches have the common objective of dividing European environmental gradients into convenient units, within which objects and variables of interest have relatively homogeneous characteristics. The SEAM agri-environmental zones, developed in 2006, and reviewed by Hazeu et al. (2010) takes into account e.g. the NUTS regions and climate: the objective was to develop a spatial framework that could be used to integrate agricultural sector modelling, bio-economic farm modelling and crop modelling. This shows that a complete different approach could have been taken for the Directive. However, the exercise resulted in 3287 zones, which probably is not politically tractable.

[33] The remit of the work group included developing guidelines for assessing potential movement to groundwater under both the EU and member state registration processes. The revised scenarios are directly applicable to EU registration. Some of member states also use or may also want to use these scenarios in their national registration process. [...] In the member state evaluations, a subset of the FOCUS standard scenarios or national scenarios are used. (Vanclooster, 2008, personal communication).

[34] See e.g. Heidemann (2006): With the revision of the Directive, the plan is to split up the EU into three zones for the authorization of plant protection products: North (Zone A), Center (Zone B) and South (Zone C). [...] Pilot projects for risk assessments are also currently underway in Zone A. The Project ''North'' (Denmark, Sweden, Finland, Estonia, Latvia, Lithuania, Norway) started in May 2005. Five products were selected, and rapporteur and co-rapporteur member states were chosen for each product. During this project, proposed agreements on operator exposure, groundwater scenarios, residues, ecotoxicology and efficacy were discussed. One year later, decisions on one product in Lithuania and Sweden have been achieved.

The delays encountered in the review programme are to be related to the reaction of DG-SANCO towards APECOP:

> *You know, if it comes to FOCUS Steering Committee, I tend to behave like a lion's mother in the sense that I am very alert about anything which might invalidate the current system*[35]. *[...] I want FOCUS to be THE standard for at least 5 more years until the review programme for existing substances is finished. We must not move the target any more.* (Personnal communication from the DG-SANCO coordinator of FOCUS, 2002).

E.2 Interpretation: interests

Mainly basing ourselves on Vinck (1995), we cite here below several interests that are attributed to scientific researchers by various sociologists studying the scientific content and describes how they link it to the whole society. We mention:

- For Collins (1981a), there is no hard social determinism of the scientific knowledge, only a global influence, that structures the knowledge by the interests of the different groups.
- The links between the scientific content and the whole society passes through relevant social groups which are the social and conceptual networks of each researcher (Collins (1974), Collins (1981b), Collins (1981a), Collins (1982), Collins (1985), Collins (1991), Collins and Pinch (1993),Edge (1979), Mulkay et al. (1975), Mulkay (1991)).
- The scientific production is predetermined by the interests of the present groups (Barnes (1974), Barnes (1977)).
- There is an interest in pursuing activities in which researchers already invested (laboratory apparatus, professional and social investments), in making choices that allows a continuity of their activities, in showing the importance of their methods.
- There are positions adopted in relation to the scientific's particular relation to the public, and also attempts to avoid to lose face, keep credibility and social recognition.
- For Barnes (1977), the link between knowledge and society is weak and does not concern "hidden interests" but rather a global influence

[35] In this phrase, the "current system" is essentially constituted by the scenarios.

of interests.

- Mackenzie (1981) links cognitive interests to social interests and from there to social classes interests.
- Habermas (1973), Habermas (1976) rejects causal relations in the relation between society and scientific knowledge: sociology should be interpretative.
- As mentioned earlier in this work, Woolgar (1981) considers that interests can not be tools for analysis but rather objects that must be explained, while Barnes (1977) advocates that interests can be supposed by the sociologists (as a methodological postulate). Woolgar (1981) considers the interests as a result of the scientific activity. He considers also that many sociologists are falling into the research of ''hidden causes'', and sees it as a ''denunciation sociology''. Moreover, he highlights that everyone claims that he knows the interests of his opponent.
- Garfinkel (1967) and Lynch (1985) advocates that agreements can only be explained by the procedures to obtain them, which are local.
- Vinck (1995) mentions that the concept of interest is linked to a vision of scientists as rational actors, although their rationality and their possibilities of maximising their interests are limited by their adherence to their networks and institutions (e.g. by norms).
- Chalmers (1991) considers that the social influence on the scientific content cannot be seen, thus it is not permanent, while sociologists of the scientific knowledge considers that the signs of that influence are erased (the credibility of a scientific claim being linked to the apparent lack of social motivations).
- When talking about interests, Bourdieu's concept of *scientific credit* must be mentioned (Bourdieu (1975), Bourdieu (1976)). Bourdieu considers that scientists are involved in a competitive market for the accumulation of scientific credit, which is a symbolic capital, and they exchange scientific knowledge against credit. Scientific credit itself can be re-exchanged against scientific knowledge, in order to gain more credit. The value of the scientific knowledge is determined by the importance that others give to the specific knowledge that is being exchanged.

We gave the opportunity to industry to answer to direct questions after having presented the objectives of the present work, but the results were deceptive (we got answers but not to our questions). We already said previously in the introduction of this chapter that interests are not observable facts. Our observations of a rhetorical question indicates that in on-going negotiations, arguments are chosen in view of an objective (or to obtain a given effect) rather than being a direct expression of intentions or interests. Moreover, in many cases, it is interesting not to express its interests but rather try to convince the others through their own interests! When negotiating, to try to understand which are the interests of the people that are involved is essential; however it is always through hypothesis and when those hypothesis are expressed openly, risks exists of being accused of raising a trial by speculation against someone.

In consequence, we believe that asking directly to APECOP participants or to industry ‘‘why do you do this or that’’ or ‘‘what are your interest in doing this’’ does not give an indication of the interests because the answer always will depend on the status of the person asking, and the rhetorical question. Does any ‘‘out of context expression’’ exist? Participatory observations always gives a status to the observant.

From opinions / problematisations collected from written sources about APECOP, one can induce hypothesis about the interests of the various actants around APECOP, through a method inspired from the historical method for ‘‘making the history’’. We performed the first steps of such a method: an investigation through the selection of sources, the diversification of those sources in relation to the question answered (‘‘Which are the interests?’’), the collection of the material in an chronological order. However, the next steps are missing here for that: a critical review of the sources (e.g. incompleteness, limitations, origin, addressee); identify facts and their relations; search for factors and causes; and finally make a rehearsal that includes all those elements.

However, without taking many risks, it can be said from our observations that pesticide is a major chemical product in terms of market volumes. Development time (for industry) is long and costly, and industry wants to make profitable its huge investments before the patents falls in the public domain. At the same time, those products are under particular scrutiny from the public and the NGO’s. Problems of pesticide pollution in groundwater are observed in aquifers, that have lead in some cases to restrictions to products that were already in the market. Agriculture is becoming increasingly dependent on pesticide. The 91/414/EEC Directive, and the FOCUS scenarios are then playing in a field with very important tensions. Industry is demanding for a regulation at the European level to allow them to make strategic planing in a science based stable and predictable regulatory environment[36].

It does not seem that some of the ''validation'' methodologies proposed were in line with those needs (e.g. a juridical or political ''validation'' vs a ''science based approach'', an agroecological stratification of Europe vs the ''zonal approach'' and the market size, the risk of new delays resulting from criticisms towards the simulation scenarios vs ''time to market'', etc). A political approach rather than the ''science based'' one is seen as a sources of uncertainties, and diminishes the needed predictability of the regulation. As well, divergent scientific results are also seen as a source of problems. The FOCUS scenarios are also probably playing an important role in avoiding unnecessary costs and more in-depth testing when industry develops new products that reveal themselves to have a too high potential to leach to groundwater. This is because modelling costs less than experimentation, so it is valuable as a screening tool (which is also a positive factor for the ''time to market'' objective)[37]; but also because standards avoids by themselves costly discussions about the right method to use, about the interpretation, allows the use of models by non-modelers and opens possibilities of savings in regulatory costs in the framework of the mutual recognition between Member States. Standards does have a certain level of arbitrary decisions and can contain scientific uncertainties (as stated by FOCUS-GW itself). Those uncertainties does not prevent to use them, precisely because they are the result of an agreement and plays important functions despite its scientific uncertainties. Any argument against the ''validity'' of the scenarios thus has consequences outside the scientific field (e.g. on the review program in the 91/414/EEC Directive, and thus on the ''time to market'' and the benefits expected, as well as on potential fine to MS imposed by the European Commission for taking too long to review the products).

It must be recalled that APECOP end-users are FOCUS-SC and ECPA, its results must be useful for them. FOCUS itself was conceived to be useful into the 91/414/EEC Directive, with the aim of promoting dialogue and provide guidance for industry, the member states and the

[36]It seems to us that major pesticide producers are at the same time demanding for a regulation of pesticide because it is also a way to rise the barrier before entering the market, which can be seen as positive for major industrial players against small concurrent companies, as well as way to regulate the market competition.

[37]It seems however that molecules are first developed and tested for efficiency, and then several candidates are tested against the scenarios.

Commission. No NGO are involved into FOCUS, while a democratic assessment criterion probably will criticize that. Moreover, FOCUS-SC defines itself as the place where the political and juridical questions that arise in the working groups should be discussed and solved (as opposed to a participatory and juridical assessment). APECOP is expected by FOCUS to have a practical purpose, and to contribute to the political objectives of FOCUS; APECOP accepted to play that role (as stated in its own initial objectives). There is also a problem of audience: various of the proposed "validation" criteria came from scientific disciplines that are very far from the groundwater modelling, and thus judged as "philosophical". Those things were translated in the representativeness as conceived by APECOP: scenarios represents for being used in the regulatory context.

Finally it must be said that there is a chain of projects since 1993, with the results of one being used in a following project with many of the same scientists. Along the same line, some scientists involved in APECOP were also members of FOCUS and could be seen their own work criticised.

APPENDIX F

A chronology of APECOP arguments related to FOCUS scenarios "validity"

Chronology of the arguments proposed by the APECOP participants to answer to the rhetorical question "Are the scenarios valid?", classified according to the received (doxa) and antagonist propositions. For doxa and antagonist propositions definition, see appendix C. From the recorded material, two propositions can be distinguished as follows:

- Proposition 1 (doxa). The FOCUS scenarios are valid, but they have some defects that we must correct; our classical modelling method is the good one to find the defects and improve the scenarios.
- Proposition 2 (antagonist). The scenarios may have more defects than those you expect; other methods are needed to find those defects.

The chronology is as follows:

DATE AND ENUNCIATORY CONTEXT	ARGUMENTS IN FAVOR OF THE RECEIVED PROPOSITION (DOXA)	ARGUMENTS IN FAVOR OF THE ANTAGONIST PROPOSITION
•1999, APECOP project technical annex		•the scenarios and models have defaults, they must be improved (through new approaches compatible with what already exists) and not be replaced or destroyed
•April 2000, APECOP plenary meeting	•a member of FOCUS remind to APECOP, in detail, of ‘‘what the scenarios are’’	•validation must be done using spatialized climatic and soil variables
•October 2000, APECOP plenary meeting	•take into account only parameters taken into account by FOCUS-GW (potential agricultural areas)	•to take into account socio-economic criteria for the scenarios’ validation
•March 2001, APECOP subgroup meeting (not during APECOP plenary meeting)	•to take into account only soil and climate criteria, no agricultural areas: rebuild homogeneous leaching areas	•take into account potential agricultural areas
	•validation by agricultural areas is not scientifically interesting because the answer is known: the scenarios will be invalidated •the project coordinator requires not to validate using agricultural areas	
•May 2001, APECOP plenary meeting	•only the experts from FOCUS-GW can tell which areas are represented by the scenarios, they must be consulted	•only text matter, not the intention of the FOCUS-GW reports authors
	•The scenarios should describe an overall vulnerability approximating the 90th percentile of all possible situations	•there logically exist situations where vulnerability is not well described by the scenarios
	•the APECOP validation protocol involves too many assumptions, we will not be able to conclude	•assumptions are used in any modelling

... continuation of the table

DATE AND ENUNCIATORY CONTEXT	ARGUMENTS IN FAVOR OF THE RECEIVED PROPOSITION (DOXA)	ARGUMENTS IN FAVOR OF THE ANTAGONIST PROPOSITION
•May 2001, email conversations	•the expert opinion by FOCUS-GW is better than the APECOP databases (that are not reliable) and therefore is APECOP itself is not better than expert opinions	
•September 2001		•the FOCUS-GW specialists did not have clearly defined the areas represented by the scenarios. APECOP will then base itself on the FOCUS-GW texts. FOCUS area = a subdivision of arable land into climatic areas
•November 2001, meeting of a subgroup during a plenary APECOP meeting	•this is not science	•scenarios represent the European Union, however the overseas territories were not considered by FOCUS-GW. Thus the scenarios are invalid. Scenarios represent arable land, however grasslands are included in the scenarios; the scenarios are therefore invalid.
	•that story is a pain in the ass, it's boring	
	•It is a matter of vocabulary (not a scientific question)	•as FOCUS reports are legal texts, vocabulary is important
•December 2001, FOCUS Steering Committee meeting (an APECOP member is present)		•the APECOP work can improve the scenarios if the correct procedure is followed

... continuation of the table

DATE AND ENUNCIATORY CONTEXT	ARGUMENTS IN FAVOR OF THE RECEIVED PROPOSITION (DOXA)	ARGUMENTS IN FAVOR OF THE ANTAGONIST PROPOSITION
•December 2001, email from a member of FOCUS (and APECOP) to the APECOP coordinator	scenarios' validation should serve a policy and practical purpose, the academic aspect is secondary	
•June 2002, plenary APECOP meeting	•APECOP lacks good data bases	•good databases (that industry builds) might be available for APECOP
•July 2002, meeting between APECOP and Industry	•some people think that APECOP's interpretation is wrong when it says that the scenarios represent areas (whatever those areas are), they represent the whole Europe •the APECOP databases are imprecise	•APECOP's work is useful because it contributes to industry's efforts regarding spatialised leaching modelling
•November 2002, plenary APECOP meeting		•statistical demonstration that the scenarios do not necessarily represent the 90th percentile (p90), but it is possible by chance. This contradicts FOCUS to some extent •the different possible scenarios' interpretations must be tested, to see the consequences on validity
•April 2003, FOCUS Steering Committee meeting (with some APECOP members present)	•APECOP and FOCUS approaches are not more realistic one than the other (same algorithm); APECOP is based on assumptions of which we do not see the consequences	•APECOP is based on more precise databases but the latter also have problems

... continuation of the table

DATE AND ENUNCIATORY CONTEXT	ARGUMENTS IN FAVOR OF THE RECEIVED PROPOSITION (DOXA)	ARGUMENTS IN FAVOR OF THE ANTAGONIST PROPOSITION
	•the scenarios are good and there is no alternative; APECOP is not operational as an alternative	•APECOP does not want to invalidate FOCUS-GW, it only wants to assess the quality for improvement purposes; any changes to scenarios should be made within FOCUS
	•FOCUS must keep a margin for political decision	
		•FOCUS should be open to scientific progress
•April 2003, APECOP subgroup meeting (not during APECOP plenary meeting)	•the statistical method (in its 2003 state of development) has too many uncertainties, so the conclusion that the scenarios are invalid can not be drawn •the method using a numerical model indicates that 3 of the 9 scenarios have problems because they are not restrictive enough, but 6 are good, so overall, scenarios are not so bad	
•June 2003, APECOP plenary meeting	•the scenarios and models as they are, bring more transparency in contrast to previously used expert judgements	

... continuation of the table

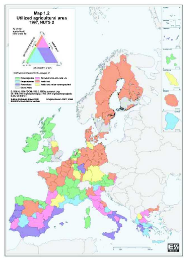
DATE AND ENUNCIATORY CONTEXT	ARGUMENTS IN FAVOR OF THE RECEIVED PROPOSITION (DOXA)	ARGUMENTS IN FAVOR OF THE ANTAGONIST PROPOSITION
•June 2003, meeting of a subgroup of APECOP	•one must help FOCUS, not demolish it; a graph potentially showing a low representativeness of the scenarios was intensively discussed; a member of both APECOP and FOCUS said that *"the conclusion is to be written in order to avoid any possibility to conclude that the scenarios are invalid"*; one must present the results in a diplomatic manner; one must present the shortcomings of the statistical method, otherwise they will be apparent and one will be scientifically discredited; it was decided to suppress the graph because there were too much uncertainties and showing it could lead to problems (with the registration process, and also could lead to scientific discredit).	
•June 2003, a scientific conference about pesticide, members of administrations and industry are present	•the statistical method does not allow to conclude on the validity of scenarios, the method using a numerical model shows that three of the nine scenarios have problems that should be corrected	•APECOP shows that methods for scenarios' validation are under development and cites as an example the fact that overseas territories are not taken into account. It is said that these methods can be based on law, text analysis, logic

APPENDIX G

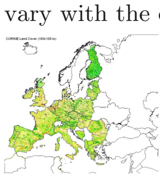
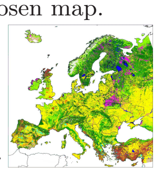
Fluctuation of the definition of the FOCUS scenarios in APECOP

In appendix F, a series of arguments related to the scenarios are presented, including those related to ‘‘what they represent’’. If we take the same received and antagonist propositions, and the same rhetorical question (‘‘Are the scenarios valid?’’) as those in appendix F, but keep only arguments related to ‘‘what the scenarios represent’’, we obtain the following table showing the fluctuation of the definition of the scenario’s representing concept according to the rhetorical question.

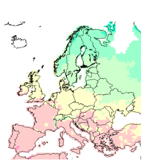
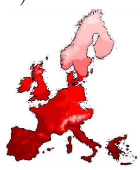
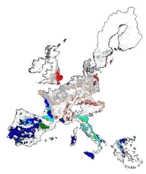
The chronology is as follows:

Year, enunciatory context	Definition of the scenario's representing concept, according to proponents of the received proposition	Definition of the scenario's representing concept , according to proponents of the antagonist proposition
•1999, technical annex	•scenario = realistic worst case input [for a model]	
•April 2000, plenary APECOP meeting	•The scenarios must represent (according to a member of FOCUS and APECOP) at the same time – realistic cases – cases of highest vulnerability to pollution by pesticide (*''realistic worst cases, 90th percentile of leaching''*) – the main agricultural areas of Europe – all climatic conditions	•the scenarios are a combination of a crop, a location and a pattern of pesticide application. Moreover, the scenarios must represent the 80th percentile in regard to the spatial component
		•no precise map of the area represented by the scenarios exists (even from FOCUS), only a sketch of climatic areas 
		•main agricultural areas are defined geographically by e.g. EUROSTAT maps of agriculture. 

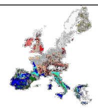
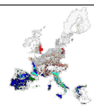
continuation of the table...

Year, enunciatory context	Definition of the scenario's representing concept, according to proponents of the received proposition	Definition of the scenario's representing concept, according to proponents of the antagonist proposition
•April-October 2000	•scenarios were created by FOCUS and we can not take the liberty of interpreting them, we can just use them as they are	•interpretation can not be avoided, FOCUS reports and texts must be used for that
•October 2000, APECOP plenary meeting	•scenario = potential agricultural areas. Potential = defined without any anthropogenic factor like socio-economics (thus from climate, soil, slope, etc only). This is opposed to EUROSTAT map.	•a scenario is designed to represent a part of reality. Scenarios represent: 1. the main agricultural areas of Europe 2. the 90th percentile of leaching within those agricultural areas
•March 2001, meeting of an APECOP subgroup (not during a plenary meeting)	•scenarios = potential leaching areas, called since then "FOCUS areas"; scenarios represent the p90 within the FOCUS areas, one must rebuild those areas from available maps	
•March-May 2001		•several land cover maps exist: CORINE landcover or PELCOM; the geographical extension of the scenarios can vary with the chosen map.  or  •permanent crops as well as grass are present in the FOCUS scenarios parameters but excluded from the reconstructed FOCUS areas. In consequence, the representativeness of some of the scenarios can be questioned if they represent only arable land.

continuation of the table...

Year, enunciatory context	Definition of the scenario's representing concept, according to proponents of the received proposition	Definition of the scenario's representing concept, according to proponents of the antagonist proposition
		•various climate maps exist, the geographical extension of the scenarios can vary with the chosen map (GISCO or FRAUNHAUFER).  or 
•May 2001, plenary APECOP meeting	•the scenarios represent the vulnerability in all situations inside FOCUS areas •FOCUS areas = climate and arable land areas; climatic areas are defined by FOCUS rainfall and temperature ranges	•the scenarios represent the vulnerability in some situations within these areas •FOCUS areas =the climate and soil of all cropland
•May 2001, email conversations	•the scenarios are standards that should not be changed too often	
•May to September 2001	•the FOCUS areas are reconstructed using only climate maps and arable lands maps	•significant differences exists between the FOCUS sketch map of climatic zones and the reconstructed climatic zones  vs  • due to the interpretation that the scenarios represent only arable lands, Finland and Ireland are more or less not included into any FOCUS area •FOCUS areas are significantly different with or without permanent crops and grasslands

continuation of the table...

Year, enunciatory context	Definition of the scenario's representing concept, according to proponents of the received proposition	Definition of the scenario's representing concept, according to proponents of the antagonist proposition
		 vs 
•September 2001,email conversations		•people from FOCUS-GW do not agree on what the scenarios are, one must base itself on a justified interpretation of their texts
•November 2001, meeting of a sub-group during a plenary APECOP meeting	•the scenarios represent only the European continent area of the EU 	•the scenarios represent the soils and climates throughout the whole European Union, including overseas tropical areas  •the texts accompanying the parameters of the scenarios are part of the scenarios
	•the scenarios are exclusively a matter of science	•the scenarios are legal texts and therefore they must be compatible with existing legislation and must meet a minimum of logic and vocabulary rules
•December 2001, email from a member of FOCUS and APECOP	•the scenarios may include logic errors but what is most important is their practical and political usefulness	
•July 2002, meeting of APECOP and industry	•the scenarios are objects to interpret, one must take the interpretation agreed "by the majority", the scenarios represent the vulnerability of the whole EU	•the scenarios represent vulnerability inside climatic areas

continuation of the table...

Year, enunciatory context	Definition of the scenario's representing concept, according to proponents of the received proposition	Definition of the scenario's representing concept, according to proponents of the antagonist proposition
•November 2002, plenary APECOP meeting	•the scenarios must be consistent with the "simple political thought", the scenarios are 50% "empirical"[1] , there may be several ways of interpreting the scenarios	•statistically the scenarios do not necessarily represent the 90th percentile (p90), but it is possible by chance and expert advice
•April 2003, meeting of the FOCUS Steering Committee, attended by some members of APECOP		•the scenarios can not be valid or invalid, they simply have a degree of quality
•April 2003, meeting of an APECOP subgroup (not during APECOP plenary meeting)	•the scenarios are important objects for many people, thus one must criticize them diplomatically	
•June 2003, plenary meeting APECOP	•the scenarios can not be changed outside institutional-political spaces; outside of these spaces, the criticisms must be very diplomatic; scenarios are objects that introduce a transparency that lets you know how decisions are taken and to discuss , test, criticize	

[1] In the sense of non-scientific, pragmatic choice.

continuation of the table...

Year, enunciatory context	Definition of the scenario's representing concept, according to proponents of the received proposition	Definition of the scenario's representing concept, according to proponents of the antagonist proposition
	•"lively" discussions, about how to present the results of the scenarios' validation at the following scientific conference. The main important think for many were the consequences of the presentation on the pesticide registration. The question of knowing if scenarios represents vulnerability inside climatic areas or for the whole Europe was judged secondary: it was decided to say that it is "just" a question of interpretation (i.e. not a scientific question) and present the two options to avoid that discussion at the conference.	
•June 2003, a scientific conference about pesticide, members of administrations and industry are present	•the scenarios represent the vulnerability of either: areas (climatic-arable) or of the whole of Europe, according to the interpretations •there are many criteria for scenarios' validation, but it is an academic question	•there are many criteria for validation of scenarios, corresponding to different aspects of the scenarios

APPENDIX H

Representation in philosophy

Chapter 5 develops a concept of representation using scenarios that allow us to propose a representativeness assessment criterion starting from a transparent and justified philosophical base. At the same time, our representation concept must be compatible with the anthropological observations of chapter 4, and be scientifically useful.

In order to achieve our objective, various sources are used and they are presented here for those not necessarily familiar with some of the philosophical background that we are referring to. First, a selection of representation concepts from the philosophical tradition up to the 19th century. Then, a closer look to some proposals made in the 20th century: the phenomenology and its representation concept (mainly based on the work of Husserl), the semiotics of representation, hermeneutics, radical critics to the representation paradigm, Latour's circulating reference, and finally the empirical structuralism of B. van Fraassen.

H.1 Representation from ancient Greece up to the 19th century

One of the philosophical problems of representation is that it is often conceived of as the relation of correspondence between the actual nature of the empirical object and the appearance conferred on it by the perception that is attempting to seize it. As a result, in that context, perception is the underlying problem of representation. The Pythagoreans held perception to be dependent on the subject alone (e.g. according to Dumont (1996), Alcmaeon attributes the possibility of the vision to a light coming from the eyes). The Epicureans, on the other hand, conceived of the perception by the senses as being passively affected by the external presence of the object, whose outline, shape or image imprinted themselves on the soul (Epicure, 1910).

In fact[1], in the Greek era, the most widely held theory was that which presents the thing perceived as a mixed *physical* entity designated by the term "phenomenon", determined both by the senses and the nature of the object. The epistemological consequence of this is that the physical nature of the object is never known in itself, but is apprehended by a physical intermediary: the ‘‘phenomenon’’. A kind of relativism then developed in Greece. Protagoras limited all reality to the sensory and changing universe of phenomena, and the human logos[2] was the measure of all things. Aristotle and Plato believed that to limit reality to the universe as perceived by the senses held true only for the sensory world, but that there existed intelligible and universal realities (respectively, separate forms and immanent types or species) that the understanding and the work of reason might grasp or see in a speculative or theoretical way.

The debate continued much later. We will mention on the one hand, Descartes and Spinoza, for example, who thought that the basis of all human knowledge was located in consciousness; and, on the other hand, Hume, who believed that all knowledge came from perception. But unlike in Antiquity, the phenomenon was no longer a material body, and no longer gave rise to a relativism of experiential objects.

---

[1] This paragraph about the Greek era is mainly inspired by the reading of Dumont (1996).

[2] Logos is the reason expressed in words, thinks, arguments or justifications.

In Kant's time, the debate may be summarised as follows (Ferry, 1987):

	Representation		
subject in itself (outside representation)	subject for us (self-consciousness)	object for us (phenomenon)	object in itself (outside representation)

In principle, self-consciousness and consciousness of the object may suffice to define representation. However, we may think that we exist outside consciousness, and that the object exists outside of the look we bring to bear on it. Access to the object in itself poses the problem of truth (or of objectivity). For Descartes (according to Kant (1987-1781)) to ask if our representations of objects are "true" is to try to discover if they are appropriate to the object as it exists in itself, outside representation.

	Representation		
		Objectivity	
subject in itself (outside representation)	subject for us (self-consciousness)	object for us (phenomenon)	object in itself (outside representation)

Formulated in these terms, such a problem is a priori insoluble; we have no means of getting outside our representation in order to compare the two objects. Descartes offers the solution of a divine guarantee.

	Representation		
		Divine guarantee	
subject in itself (outside representation)	subject for us ((self-consciousness))	object for us (phenomenon)	object in itself (outside representation)

For Descartes, a correct analysis of the concept of God shows that He could not be a deceiver. If it were the case that the object in itself was different from the object for us, this would mean that God had wished to deceive us, which is impossible, so therefore the object in itself and the object for us correspond to one another. For Kant, this is a wild argument and leads to a vicious circle: we need to be certain of our knowledge of God; what guarantees that "God for us" corresponds to

"God in Himself"? In the "Critique of Pure Reason" (Kant, 1987-1781), Kant puts forward a different solution to the problem of objectivity:

OUT OF REACH	OBJECTIVE	SUBJECTIVE	OBJECTIVE	SUBJECTIVE
	a priori intuition: - space/time		a priori schemes: - self-consciousness - categories - causality laws	
object in itself (=object for God)	------------------> intuition (sensations)	phenomenon matter/form	-----------------------> understanding	object for me
	representation of the singular		representation of the general	

Phenomena and objects for us have in common the quality of being **products**, of the **understanding** and of **intuition** respectively. It is no longer a matter of passive contemplation: the subject intervenes in their **creation**. But this creation is not entirely dependent on the subject. There are conditions for their production that are universal to humankind, and that thereby transcend them: what is objective is not what is produced but **how** they are produced, the **a priori** of their production. The differences between the products derive from experience, which is the exercise of the understanding and of sensation. What is a priori for sensation is the *a priori intuition*: the fact of perceiving in terms of space and time. For the understanding, these are the a priori schemata: self-consciousness, causality, categories. In phenomena, what comes from a priori intuitions is their "form"; what comes from our sensations is their "matter". Concepts, as schemata, present two characteristics; comprehension and extension. What we grasp by a concept is what is common to the particular elements, independent of the differences. We have partial access to real existence by means of intuition: we access the extension of the concept, in its singularity. But this access is constrained by a priori intuition. In our schematic presentation of Kant's thought, it seems to us that what is considered objective is the process of the intuition and of the understanding, and not its products. We have been guided in this interpretation by the following text of Kant:

> *Pure understanding is not, however, in a position, through mere categories, to prescribe to appearances any a priori laws other than those which are involved in a nature in general, that is, in the conformity to law of all appearances in space and time. Special laws, as concerning those appearances which are empirically determined, cannot in their specific character be derived from the categories, although they are one and all subject to them. To obtain any knowledge whatsoever of these special laws, we must resort to experience; but it is the a priori laws that alone can instruct us in regard to experience in general, and as to what it is that can be known as an object of experience.* (Kant (1929-1781) part II, 1st division, chapter II, § 26).

The object in itself is considered as the result of the exercise of divine understanding, which is inaccessible to us. However, for Kant, the idea of an omniscient understanding, of a complete scientific understanding of the universe, remains useful for scientific progress and activity: it demands of science that it should be organised in a systematic whole and should ceaselessly seek further. It is transcendence that, in its universality, allows us to escape from scepticism, and from pure relativism.

Kant gives us a theory of meaning: a concept has meaning if it is applied to intuition via schemata. This implies that concepts empty of meaning can exist: those that can find no application in intuition. These concepts are "non-representable".

H.2 Phenomenology and representation

Phenomenology took up the problem of objectivity on the basis, notably, of a criticism of the way in which Descartes conceived of the body. Descartes' *cogito*, understood as self-reflexive and thinking substance, seems to owe nothing to bodily existence. It is only by a double diversion, that of divine veracity and that of the clear and distinct idea, that Descartes manages to some extent to validate the feeling of having a body (Canivet, 2002). As for perception, this is essentially the judgement of the mind that uses and interprets the sensations of the body. The body appears as a tool, which is "my being in the world", and I perceive the world through its mediation, that of the senses.

Phenomenology (in the sense given to it from the XX$^{\text{th}}$ century) gives a quite different status to consciousness. For phenomenology, we are not first of all conscious and then conscious of something: consciousness is always consciousness *of*. There is therefore no pure and isolated subjectivity. Consciousness is always consciousness of something, and this something is the product of a process, that of intentionality. Our experience (Husserl, 1976) is directed toward, intends, represents things through its content, its meaning towards an object.

According to Smith (2009), phenomenology is the study of structures of consciousness as experienced from the first-person point of view. Some classical texts in phenomenology are: Husserl (1976), Heidegger et al. (2010-1927), Merleau-Ponty (2010), Sartre (2001-1943). The central structure of an experience is its intentionality, its being directed toward something, as it is an experience of or about some object. An experience is directed toward an object by virtue of its content or meaning (which represents the object) together with appropriate conditions of possibility (including embodiment, cultural context, language, and social practices).

Husserl contends that intentional objects are not an internal representation, a kind of "copy in the mind"; likeness does not make an object the image of another: it is the intentionality that allows to use an object to point at another (Romano, 2010).

Phenomenology addresses the meaning things have in our experience, notably objects, events, tools, time, and the self. Experience includes emotion, perception, thought, etc. Phenomenology aims at describing things (the phenomena) as they appear, and thus describes also the *way*, the *mode* by which they appear (Montenot, 2002). Husserl (1976) is proposing to define the phenomenology as a comeback to the *things themselves* (the phenomena). However, he does not oppose an appearance to an hypothetical thing-in-itself: he sees phenomena as the way the reality of the consciousness manifests itself. His phenomenology tries to catch the pure forms, essence or idea (*êidos*) of the phenomena, by the way of a "reduction".

He distinguishes the natural self, the phenomenological self and the transcendental self, all the outcome of two successive reductions (Mutelesi, 1998):

> *[P]henomenological reduction consists, for the self, in the suspension of any judgement or any belief about the world, so that, above the self of the natural attitude, which expresses in a naïve way judgements about the being of the objective word, the pure self is placed, which simply observes the life of the consciousness of the "natural" self. [...] [F]ollowing the phenomenological reduction that suspends any judgement of being or any act of belief comes transcendental reduction which leads to the transcendental ego. The transcendental ego is an instance that analyses the relation of the pure phenomenological self to the natural self and that describes the pure lived experiences of the phenomenological self.*

The phenomenological reduction allows to let emerge the essence of phenomena, while the transcendental reduction allows to reach the purity of the consciousnessMontenot (2002). Consciousness is seen as the origin of all the possible meanings of the world. Husserl distinguishes also between noesis and noema: the intentional process of consciousness is called noesis, while its ideal content is called noema(Smith, 2009). What interests him is the meaning consciousness gives to its noema. There is the perceived object, the imagined object, the disturbing object, etc., which are correlative to the perceiving, practical, moved, etc. consciousness.

Phenomenology in its modern sense, after and following Husserl, arises at the point where being (the object in itself) is placed in parentheses in the analysis (Barbaras and Greisch, 1996) and when one deals in an autonomous manner with the way things appear, that is to say, phenomena. Furthermore, appearing is considered to be an essential dimension of being. The very essence of being is to appear: the phenomenon becomes the absolute and no longer, as in Kant, an imperfect and finite point of view, on what would be the being itself.

The idealistic turn of Husserl, as shown by the central rôle played by consciousness, is rejected by the majority of its disciples.

Consciousness is first of all, primordially, pre-reflexive: consciousness, but without self-consciousness; there is a lack of distinctness between subject and object and therefore also anonymity in pre-reflexive consciousness (for consciousness as towards the Other, which is not therefore perceived as the Other). Reflexive consciousness distinguishes the object from itself. The originating experience is that of a lived body. It is reflexive intentionality that is able to conceive the body as an object. More fundamentally, the body itself is not a thing since it is through it that we have access to things.

It is useful to distinguish, in the process of intentionality, four variants that together participate in the process of consciousness:

– experiential intentionality: development of data, sensations, heat, light...
– intelligent intentionality: development of concepts, by selection and assembly of data
– rational intentionality: the judgement affirms the reality of what has just been understood, concepts become acts
– responsible intentionality: all knowledge is a knowing-how or a power; the end point is decision

All our objective and scientific considerations are subordinate to a lived experience of one's own body. The same holds true of our relations to others. It is therefore necessary to add an intersubjective element, which is present in Husserl's work. This intersubjectivity starts from the experience of my body towards the experience of the other's body and ends in the possibility of the existence of the world of the other. According to Mutelesi (1998), empathy in Husserl is a way of gaining access to the other: it is "the experience of the Other (and above all of the other body) – starting from my own experience of myself – as the original body of the Other; in effect, a primordial experience of the Other within an analogising 'representation' starting from the perspective of my primordial experience of myself." It becomes conceivable that based on the reciprocal constitutions of bodies there may be crosschecking between my world and the world of the other. Objectivity appears, therefore, only at the point where the other with its world is formed in an intersubjective experience. In this way, objectivity becomes relative to intersubjectivity, while subjectivity is the only absolute objectivity that is possible; I can only describe as objective those assertions about phenomena that are particular to me. But we may note intersubjectively that our experiences confirm each other. This observation is supported by communication. What is objective is therefore what is apprehended as being jointly understood, and as a result identified as being the same. This interpretation of objectivity, grounded in Husserl on intersubjectivity, is shared by both Mutelesi (1998) and by Dumont (1996).

Having introduced this dimension of objective intersubjectivity, it seems to us possible to synthesise its phenomenological location in the way presented in fig. H.1. This way of seeing intersubjectivity is complementary to our sociological analysis in chapter 4.

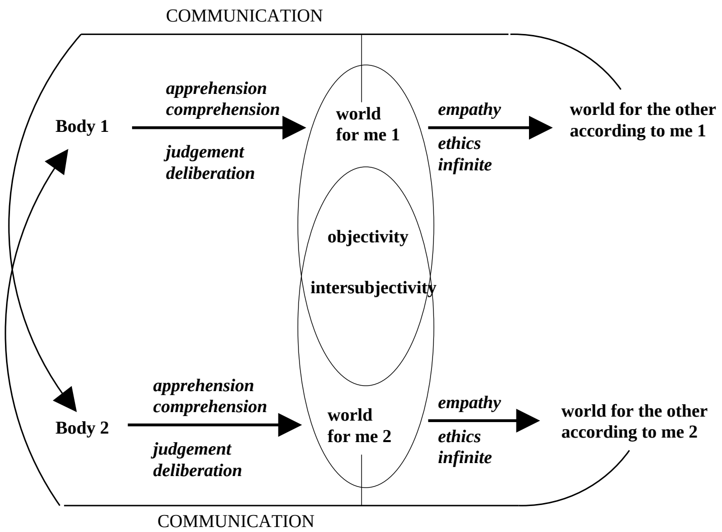


Figure H.1: *Intersubjective objectivity in phenomenology*

H.3 Semiotics of representation

Semiotics is the study of the nature of the signs, of their production, transmission and interpretation(Montenot, 2002).

According to Barsotti (2001), the semiotic approach to representation explores signs from their necessary structure to their representativeness power, but takes as solved the question of the ontological access to the represented being.

The representor links us to the represented through a relation of correspondence (conventional, felt, passive, admitted) but not necessarily through likeness. Goodman (1976-1968) contends that *''A picture that represents - like a passage that describes - an object refers to and, more particularly,* denotes *it. Denotation is the core of representation and is independent of resemblance.* Ultimately, any representation is based on an arbitrary representativeness as with the mathematical symbols.

H.4 Hermeneutics, interpretation and representation

Hermeneutics was originally the interpretation of hold texts, and particularly the interpretation of the Bible (Montenot, 2002), and latter all kinds of interpreting activities. In the modern sense, hermeneutics is the philosophy of interpretation.

According to Barsotti (2001), the critics to the representation paradigm can be classified in two groups: the conciliatory (to complete representation with other access to the being) and the radicals (deny representation). Gadamer and Ricœur, two hermeneuts, are classified amongst the conciliatory.

Gadamer (1996-1960) emphasizes the unavoidable subjectivity of hermeneutics, and includes in it the relations between consciousness and the world, from the historical and existential points of view. For Gadamer, the science is a linguistic possibility, it is through the language that we approach the facts, the language preforms our interpretation of our observations (Barsotti, 2001). Science is seen as imposing a particular type of language: the representative language, but it is not the only linguistic possibility.

According to Kaplan (2006), Ricœur contends that if language is the medium for thought and experience, the unmediated knowledge that Husserl is looking for is not possible. The symbolic language requires an interpretation in terms of rules and regularities to understand how it mediates experience, but Ricœur does not reject the intentionality of consciousness. Ricœur (1986) conceives the task of hermeneutics of texts as autonomous with respect to the intentions of the author of a text. The matter of the text (Kaplan, 2006) *''discloses a proposed world (of real and imaginary references) one could possibly experience, inhabit, verify, criticize [...].*

H.5 Radical critics to the representation paradigm

According to Barsotti (2001), philosophers like Merleau-Ponty, Heidegger, Derrida, Deleuze, Foucault are to be classified amongst those who repudiates the representation. Unlike in the hermeneutics, language is not before the being, it is always the language of something, a reality that it can only show, an ontological reality which is the non-representable of the representation. As an example of this, Merleau-Ponty (2010-1945) states that space can not be observed, because it is supposed in any observation.

H.6 Latour and the circulating reference

In (Latour, 2001), Latour describes the path followed by an interpretative report, from field research in soil science in Amazonia to its publication. He shows how the reference to the soil passes, through a series of inscriptions using devices (e.g. the results of the readings of the Munsell soil color code), to the graphs of the report then to the report itself. Each of these inscriptions constitutes an element of representation in which form and matter come together, as represented in fig. H.2.

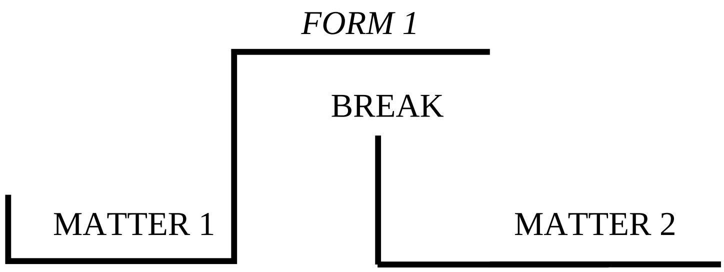


Figure H.2: *Inscription as an element of representation. Form and matter meet in the process of inscription to create new matter.*

The example of inscription Latour gives is one of the tools used by the pedologists, but in order to clarify matters, let us give an example that directly relates to the problem of scenarios that is the subject of analysis in this thesis. In order to "validate" scenarios, we have seen that scientists seek to delimit on maps the areas represented by the scenarios (which has given rise to much discussion). But where do these maps come from? One of the classic stages in cartography is image-capture by remote sensing. In this case, the form is supplied by the camera's matrix of pixels or by the digital sensor. The matter is the portion of the Earth that is photographed.

In Latour's concept, the (circulating) reference is the chain formed by a set of inscriptions as represented in fig. H.3.

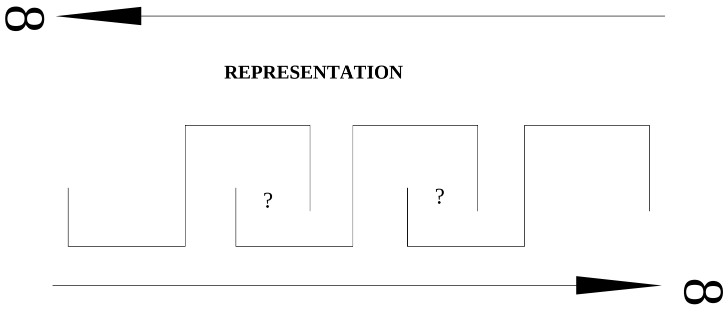


Figure H.3: *The circulatory notion of the reference follows the trace of a series of matter/form transformations. This chain develops in two directions in an infinite manner. The quality of the transformation determines the reference.*

The translation implies that there are uncertainties and risks linked to the break in the inscription process. Those uncertainties are at the center of the models and scenarios' quality assessment.

To return to our cartographic example, once the remote-sensing image has been obtained it is necessary to interpret it, to pass it through a series of other inscriptions before reaching the map.

H.7 van Fraassen: representation and empirical adequacy

van Fraassen is not offering a philosophy about what nature is, but rather a point of view about what science is, with a particular attention to practice.

van Fraassen published in 2008 a book which is both a synthesis of its previous work and the promotion of representation as a central concept in science (van Fraassen, 2008). van Fraassen presents an analysis of past and present scientific practice, and concludes that quantum physics presents a major shift in the criteria that prevails in practice. He argues that philosophy of science should take note of it. According to him, a criterion called ‘‘the appearance from reality criterion’’ (AFRC) prevailed before the advent of quantum physics. That criterion supposed 3 levels in the scientific representation of the world/nature: (1) the underlying reality: forces, fields,... (2) observable phenomena: objects, motions,... (3) the appearances: measurements outcomes, i.e. how the phenomena look in/from measurement set-up. The AFRC implies that science must explain how appearances are *produced* in reality. van Fraassen argues that the AFRC is not satisfied in quantum physics (in the Copenhagen interpretation[3]) because there is no causal *mechanism that produces them.* Quantum physics can be used to predict (mainly on a statistical basis), without a causal explanation. The assumption that a parameter measured has already a definite value before the measurement is, according to van Fraassen, generally rejected in foundational studies of quantum mechanics.

Starting from that view, it is easy to see how van Fraassen is advocating that scientific theories are not representing ‘‘the reality as it is’’. ‘‘Reality as it is’’ is for him a metaphysical postulate, that an empiricist must avoid. For him, scientific theories (or more precisely their models as we shall see) are representing the phenomena through empirical adequacy. The *purpose* of the representation determines the criteria for an adequate/accurate representation and the means for representing.

‘‘There is no such thing as representation apart or independent of our practice’’ says van Fraassen. The adequateness of the representation can not be seen independently of the use and practice. Here the ‘‘use’’ is seen as the fact that representing is an act made by someone (Z1): it is to *use* something(X) to represent something else (Y). Moreover, the ‘‘use’’ also means that the result (the representor X) has also uses (for someone Z2). In consequence, representation can not be conceived as independently specifiable relations between a representor(X) and a represented(Y): the user (Z1,Z2) is part of the act of representation. The full concept of representation can be summarized as following:

$$\textit{Z1 uses X to depict Y as F for purpose P to Z2} \tag{H.1}$$

where F is a predicate[4].

[3]See Faye (2008) for a description of the Copenhagen interpretation.

[4]van Fraassen (2008) gives the following example: *as* F = ‘‘as draconian’’ in the sentence ‘‘the caricature depicts Mrs Thatcher *as draconian*’’.

Not all representations relies on likeness, and if it is the case, they are selective likeness, and representation involves necessary distortions, and surplus structures. Moreover, representation can be also at the level of the structure of the properties. This is the case with the FOCUS modelling results (described in section 1.2) when they are said to represent the 90^{th} percentile of pesticide leaching: what is being represented is the statistical distribution which is a structure of the property "leaching". When representation relies on resemblance, what matters is accurate/appropriate resemblance in *relevant aspects* for the *purpose P*.

Another property of representation according to van Fraassen is that representation is indexical: it requires a user who locates himself with respect to the representation (e.g. a map can not be used without someone saying "I am here now"). An indexical term is (Makins, 1991) in linguistics and logics a term whose reference depends on the context of utterance, such as I, here, now, you, tomorrow. van Fraassen (2008) states that a representation needs an essentially indexical operation, that can not be done without a user intervention:

> *Is there something that I could know to be the case, and which is not expressed by a proposition that could be part of some scientific theory? The answer is YES: something expressed only by an indexical proposition. For to use a theory or model, to base predictions on it, we have to locate ourselves with respect to it.*

The subject is thus introduced in a triangle with the representor and the represented. van Fraassen (2008) gives an example: in a sentence like *"The exponential function represents [...] the growth of the investigated bacterial colony"* receives its reference from the context in which it is used.

Appearances are what appears to us, we never do see beyond them: there is no direct access to phenomena (and so no access to parameters without measurement). There is a co-construction, through a historical process, of measurements and theory. Measurement is a physical interaction made by agents, with the purpose of obtaining information, it has the function of a selective representation of a phenomenon by means of parameters according to a theory and a context. Measurement is relational: it is relevant with respect to a question. It is intensional (gives meaning to the measurement outcomes into a theory), intentional (with a purpose). Representation through measurement is perspectival (shows how phenomena look like from the measurement setup), and when measurements outcomes are used, is also the act of locating the measurement outcome into the space of possible states as depicted by a model. Measurement is indexical, because it requires a user to locate ''what is measured'' with respect to a theory. Measurement needs a physical interaction, and presents a series of invariants and coherence constraints. Regularities, when repeated and done by multiple users, qualifies representation by measurement as ''public representation'' rather than ''private''.

van Fraassen (2008) mentions his vision about predictions by scientific modelling, for the case of meteorological predictions:

> *The calculation provides models of the region for the next five or so days and [...] requires that those accurately represent conditions on those days - with the model's success gauged by some measure of accuracy.*

He also emphasises the indexicality for such a model, when it must be used:

> *The model presents a space of possible states and their evolution over time - the input* ***locates*** *the weather forecaster in that space, at the outset of the forecasting process. [...] user, as to provide [...] something to the effect of ''I am now here in the space of possible states''*

van Fraassen seems not to distinguish much here between a model in the mathematical sense and in the scientific sense. This was criticized and solved by the introduction of the use and indexicality in his conceptual framework. If we extrapolate from the preceding citation, the outputs of the prediction by a (scientific) model are a (mathematical) model, that must be subject to empirical adequacy, against experimental models.

We depict in fig. H.4 the conceptual framework developed by van Fraassen to support empirical adequacy. In that figure, a theory is a family of models.

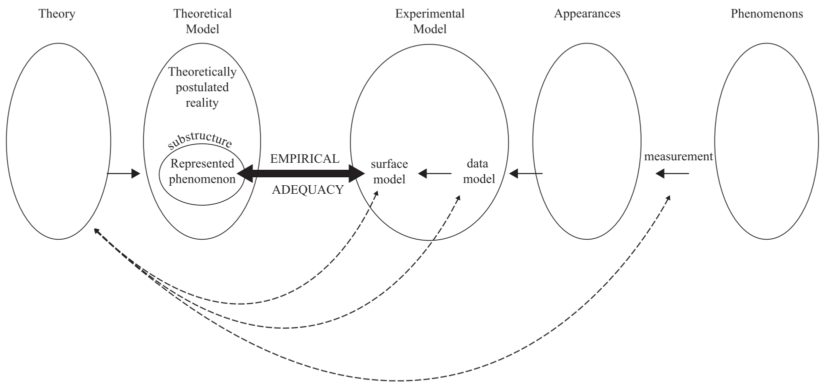


Figure H.4: *van Fraassen's empirical adequacy.* In place of the traditional empiricist dichotomy, between theory and observation, a tripartite distinction is introduced: (1) Phenomena will be observable entities (objects,events, processes) (2) Appearances will be the contents of observation or measurement outcomes (3) Theory is a family of models. Acceptance of a scientific theory only implies that the theory is empirically adequate (as opposed to true). Theory represents phenomena through the adequacy between surface models and empirical mathematical substructures of the theoretical models. *Things* in science are temporarily stabilized representations of phenomena through models. Theory, theoretical models, data/surface models, measurements are all conceived as representations of the phenomena.

As stated in van Fraassen (2004), the claim that the theory is adequate to the phenomena is indeed not the same as the claim that it is adequate to the phenomena as represented by someone.

van Fraassen subdivides experimental models EM into data models and surface models. The data models express the frequencies of measurement outputs, for a given measurement choice. The surface models (SM) are composed by: (1) a set of measurement choices (MC) (2) a set of possible outcomes (PO) (3) a function P (called surface state) that gives probabilities of outputs conditional to measurement choices

$P(PO|mc)$. So, surface models are an abstraction, an idealization of data models. Similarly to SM, theoretical models TM are composed by (1) a set of states S (2) a family of observables O (physical magnitudes) each with a range of possible values $PV(o)$ (3) a probability measure on $PV(o)$ conditional to s in S $(P(PV(o)|s))$.

The fitting between a theoretical model (TM) and an experimental model (EM) is reduced to the fact that data or surface models must be embeddable in theoretical models. We interpret van Fraassen fitting condition as follows:

> A TM fits an EM just in case TM has some s such that $P(PV(o)|s)$ contains $P(PO|mc)$, relative to the given (indexical!) identification of mc as measurements of the physical magnitude o.

The emphasis on use and indexicality places van Fraassen concepts in the field of pragmatics [5] rather than syntax or semantics. Use implies consciousness and agency. Use is an intentional activity and a practice. The *"as"* term in the fitting condition involves also contextual aspects like audience, coding conventions, function in preexisting representational systems. Representation is both an artifact and a social fact. The concepts of use and indexicality, in conjunction with empirical adequacy is supposed to avoid the introduction of a metaphysical underlying reality. However, van Fraassen recognises that a realist metaphysic can be used to stimulate the scientific imagination.

According to Kockelmans (1997), van Fraassen position is amongst others that philosophical positions can not be captured in dogmas that asks for beliefs about what the world is like, that a philosophical position may be an attitude or an approach, that he rejects any form of transcendent dimension[6] as well as metaphysical arguments, a criticism of scientism in the name of science and through a skeptical attitude.

[5]The study of language where the relation between thing-word is seen as an abstraction of a word-thing-user relation.

[6]Understood by Kockelmans (1997) as the conditions of the possibility of all finite knowledge, including HUsserl's transcendental phenomenology and Heidegger's ontology.

APPENDIX I

Symbolic objects formalism

(According to Diday and Esposito (2003)). Let Ω be a set of individuals, D a set containing descriptions of individuals $d\omega$ or of a class of individuals d_C, "ψ" a mapping from Ω into D which associates, to each $\omega \in \Omega$, a description $d\omega \in D$ from a given symbolic data table. We denote by R a relation defined on $D \times D$. uRv takes its value in a set L. We can have $L = \{true, false\}$, in this case $[d'Rd] = true$ means that there is a connection between d and d'. We can also have $L = [0, 1]$ if d can have various degrees of connection to d'. In this case, $[d'Rd]$ can be interpreted as the "true value" of uRv or "the degree to which d' is in relation R with d'". For instance, $R \in \{=, \equiv, \leq, \subseteq\}$ or R is an implication, a kind of matching (canonical, partial or flexible) which considers missing values, etc. R can also use a logical combination of such operators. A *symbolic object* is defined by a description d, a relation R for comparing d to the description $d\omega$ of an individual and a mapping a called "membership function". More formally: a symbolic object is a triple $s = (a, R, d)$ where R is a relation between descriptions, d is a description and a is a mapping defined from Ω in L depending on R and d.

APPENDIX J

Latour's power, collective, and state concepts

Bruno Latour (Latour, 1999) starts from a concept that is radically opposed to extroverted empiricism. For Latour, the distinction between facts and values, between the scientifically proven and the opinion, is to be reconsidered, to be reformulated. He proposes to put an end to the idea of nature as objective reality confronting subjects, who are human beings. For him, it is necessary to think of our relation to the world in terms of a collective of humans and of non-humans. The collective would be a continuous historical process of the redefinition of the links between humans and non-humans. He thus offers the diagram presented in Figure J.1.

In this diagram, the collective is delimited by the ellipse. Inside, the entities that are accepted by the collective, that in this sense have been integrated into the world, that have been made to enter the world, are called institutionalised essences. The collective world is called the cosmos. These essences constitute the world built by the collective process; here we truly find ourselves in an epistemology that rejects the distinction between subject and object. The institutionalised essences are the product of the collective process. Latour describes this process as a system with four main functions: perplexity and consultation, which are

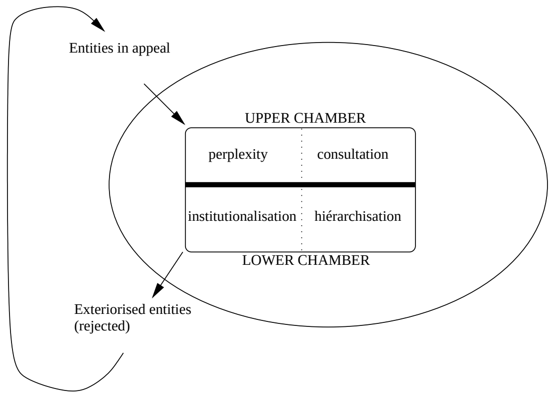


Figure J.1: Latour's collective of humans and of non-humans

the prerogative of the upper chamber; hierarchisation and institutionalisation, which are the functions devolved to the lower chamber. Power in this system is expressed through the functions of the two chambers so there must be, for the good functioning of the system, a clear separation of the two powers. Here is a review of the functions. (1) Perplexity: entities, before they become institutionalised essences, exist in the form of propositions. These propositions are articulated well or badly in the shared world of the collective, and are neither true nor false. Perplexity is the function through which the collective makes itself sensitive to and attentive to the propositions that want to become part of the cosmos. (2) Consulation: the search for trustworthy witnesses, who are the representatives of propositions. Representatives are, for example, a scientific experiment that ‘‘serves as speech’’ to a salmon for which one is trying to construct a ladder that will allow it to bypass a dam. An appropriate jury for each proposition is also formed. The ethics of discussion demand that nobody can be led to apply the results of a decision in whose discussion they have not participated (Edgar, 2006). (3) Hierarchisation: the search for compromise, negotiation, remodelling propositions and essences in order to find a way or organising them, making them compatible, hierarchising them. The end point is the definition of the new cosmos. (4) Institutionalisation: rejection of propositions with which one has not been able to come to terms, incorporation of accepted propositions into lasting practices.

The power of the upper chamber is said to be ‘‘power of taking into account’’; that of the lower chamber is said to be ‘‘power of putting in order (arrangement)". The effective functioning of this collective process is guaranteed by the State, which Bruno Latour redefines according to this new ordering of powers. Who contributes to these functions? Scientists as well as politicians, economists as well as moralists, and others! But they each have their own expertise.

APPENDIX K

A pesticide leaching metamodel at the European scale, using a radial basis functions neural network

K.1 Introduction

In the present study, a metamodel of the pesticide leaching in the former EU-15 is presented. The purpose of constructing such a metamodel is to use it for the assessment of the FOCUS-GW scenarios representativeness, according to the assessment test developed in chapter 3.

.

Tiktak et al. (2004) presented the results of simulations performed by means of the EuroPEARL spatially distributed deterministic model, for 1062 unique combinations of soil type, climate and country, sampled throughout the EU-15. Results of the EuroPEARL model were presented in maps with a resolution of 10x10 km2. Those results indicated that the leaching concentration generally increased with precipitation and irrigation and decreased with increasing organic matter content. Because of the strong sensitivity of the leaching concentration to soil properties, there was a strong variability of the calculated leaching concentration between map cells. Results further indicated that due to large irrigation amounts combined with large temporal variation of rainfall in the Southern European countries, the trend in the calculated leaching risks from North to South was less extreme than expected. This implies that areas of high leaching risk ('hotspots') as assessed by means of the EuroPEARL model occurred in all countries of the European Union, including the Southern European countries.

The metamodel here presented is based on the inputs and outputs of the PEARL deterministic deterministic model of Tiktak et al. (2004). The inputs of the metamodel are the parameters available in EU-wide soil and climate databases, while the outputs are the inter-annual mean and standard deviation of the annual leaching concentration at 1m depth, for 1 pesticide (substance D defined by FOCUS (2000), classified as a low persistent and low adsorbing compound, with Koc = 60, soil DT50 = 20 d), and 1 crop (winter wheat).

K.2 Methods

The terminology used here is presented in Piñeros Garcet et al. (2006c). The metamodel used is based on a radial basis functions neural network (RBF NN). A MKM can not be used because, as learned during preliminary trials not shown here, some input parameters vectors are close to each other, leading to singular matrices during MKM calibration (this is also stated by Jin et al. (2001)). Moreover, the MKM can not be used in cases where the same inputs lead to different outputs, as it is the case here with the pesticide calibration data, due to the inputs dimensionality reduction from PEARL to MM inputs. Very recently, kriging metamodels are beginning to be adapted to stochastic simulation outputs (see proceedings of the Winter Simulation Conference: Staum (2009)).

K.3 The deterministic model

The deterministic model used to produce the calibration and assessment data sets is the PEARL model Tiktak et al. (2004). As an example of the outputs given by the PEARL model, fig. K.1 shows the simulations for nine FOCUS-GW scenarios.

FOCUS use 20 years simulations in order to estimate percentiles. However, Chen and Kelton (2006) and Chen and Kelton (2008) described methods for determining a workable finite simulation run length that will produce sufficiently precise quantile estimates. Such a method can

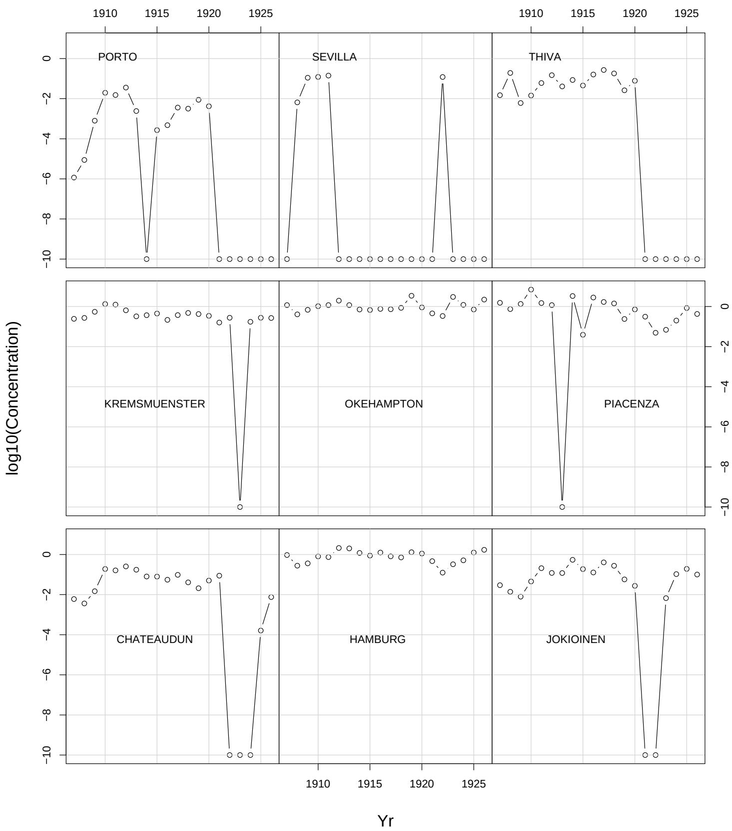


Figure K.1: *20 years simulation of leaching concentration at 1m depth, using PEARL deterministic model with FOCUS-GW scenarios and dummy substance D and winter wheat.* In the graphs, leaching of 0 $\mu g/l$ is replaced by $10^{-10}\mu g/l$

be useful for improving the FOCUS methodology. Folland and Anderson (2002) studied the performances of empirical ranking methods for percentiles estimation in case of serial correlation, which can also be useful for evaluating the FOCUS procedure for quantile estimation. However since our metamodel is designed to reproduce as close as possible the deterministic model simulations, we will consider the 20 years simulation results as a population, according to what was stated in section 6.5.1

K.3.1 The radial basis function neural networks (RBFNN)

The RBFNN used are of the same type as those presented in Piñeros Garcet et al. (2006c).

The number $\ddot{t}$ of basis functions must be determined. In this study, after some trial-error exploration, the following values for $\ddot{t}$ were tested: $\ddot{t} = 160, 165, 170, 175, 180, 185, 190, 195, 200$.

The value chosen for $\ddot{t}$ is ultimately determined based on the network which gives the best model efficiency both in calibration and quality assessment:

$$modelling\, efficiency: \; EF = \frac{\left(\sum\left(\left(\ddot{O}_i - \overline{\ddot{O}}\right)^2\right) - \sum\left(\left(\ddot{P}_i - \ddot{O}_i\right)^2\right)\right)}{\sum\left(\left(\ddot{O}_i - \overline{\ddot{O}}\right)^2\right)} \tag{K.1}$$

where (in $\mu g/l$) $\ddot{O}_i$ are the ''observed'' output values (i.e. calculated with the deterministic model) and $\ddot{P}_i$ are the output values predicted by the metamodel.

K.3.2 The metamodel

We consider here below that the MM is the function with having input in their original units, before being transformed for standardization, and having as outputs the μ_j and σ_j as needed for the upscaling procedure (in $\mu g/l$). Thus, the RBFNN is only a part of the MM, and comprises also the standardization shown here below.

The inputs for the metamodel are described in table K.1.

Table K.1: *Inputs of the metamodel*

R_y	mean yearly rainfall	mm
R_s	mean spring rainfall	mm
R_a	mean autumn rainfall	mm
T_y	mean yearly temperature average	$^\circ C$
PD_t	packing density of topsoil	g/cm^3
PD_s	packing density of subsoil	g/cm^3
AW_t	available water content in topsoil	m^3/m^3
AW_s	available water content in subsoil	m^3/m^3
S_t	sand fraction in topsoil	g/g
S_s	sand fraction in subsoil	g/g
C_t	clay fraction in topsoil	g/g
C_s	clay fraction in subsoil	g/g
OM_t	organic matter content in topsoil	g/g
OM_s	organic matter content in subsoil	g/g
pH_t	pH of topsoil	$pH\ KCl\ 1N$
pH_s	pH of subsoil	$pH\ KCl\ 1N$

As stated by Bishop (1995), in order to facilitate a good spreading of the radial basis functions centers it is advised to standardise the inputs of the RBFNN prior to calibration. Standardisation of the calibration inputs in order to obtain standardised inputs for the NN was performed by subtracting to each MM input its mean and dividing it by its standard deviation. This is described in the following equations.

The matrix of MM calibration inputs is:

$$\mathcal{X}_{calMM} = \begin{pmatrix} x^{1,1}_{calMM} & \dots & \dots \\ \dots & x^{i,j}_{calMM} & \dots \\ \dots & \dots & x^{m,n}_{calMM} \end{pmatrix} \tag{K.2}$$

with m the number of calibration vectors and n the number of MM inputs ($j = 16$, see table K.1).

The matrix of NN calibration inputs is:

$$\mathcal{X}_{calNN} = \begin{pmatrix} x^{1,1}_{calNN} & \dots & \dots \\ \dots & x^{i,j}_{calNN} & \dots \\ \dots & \dots & x^{m,n}_{calNN} \end{pmatrix} \tag{K.3}$$

and

$$x^{i,j}_{calNN} = \frac{x^{i,j}_{calMM} - mean(x^{\cdot,j}_{calMM})}{stddev(x^{\cdot,j}_{calMM})} \tag{K.4}$$

where $x^{\cdot,j}_{calMM}$ is the j^{th} column of $\mathcal{X}_{calMM}$.

The output of the metamodel is an estimation of mean and standard deviation of the yearly leaching concentration (out of a 20 years series) for the "D" type pesticide defined by the FOCUS group, and for winter wheat. Two RBFNN (see Piñeros Garcet et al. (2006c) for the equations of a RBFNN) were trained, one for the mean μ_j and one for the standard deviation σ_j. Thus $\overrightarrow{Y_{MM}} = (\mu_j, \sigma_j) = MM(\overrightarrow{X_{MM}})$.

Concentrations are positive real numbers and leaching concentrations do have a detection limit due to the laboratory available measurement methodologies. Those detection limits are variable between pesticides, methods, and laboratories. In consequence small values, calculated by the deterministic model, are physically meaningless, while they perturb the calibration process of the RBFNN by constraining the algorithm optimize in meaningless areas. $10^{-10}\,\mu g/l$ is thus chosen as the MM minimal output. In order to take this into account, a variable transformation is done. RBFNN outputs calculated from Gaussian basis functions like in Piñeros Garcet et al. (2006c) are not constrained and can have values in $[-\infty, \infty]$. In consequence, a transformation of the outputs of the deterministic model (based on a artanh-like function, eq.K.5) was implemented and the RBFNN is calibrated on the transformed values (eq.K.6).

$$y_{transformed} = \ddot{t}(y) = s * \ln[\frac{(y - y_{min})}{((c - y_{min})/(y_{max} - c)) * (y_{max} - y)}] \quad \text{(K.5)}$$

where s is a slope factor, y_{min} and y_{max} are the minimal and maximal output values that will be allowed for the metamodel, c is the point where y_{calNN} equals 0.

$$y_{calNN} = \ddot{t}(y_{NM}) \quad \text{(K.6)}$$

where, y_{NM} is the output of the deterministic model, y_{calNN} is the calibration set for the RBF neural network model.

In order to obtain the set of the MM outputs, $\mathcal{Y}_{MM}$, the outputs of the RBFNN are back-transformed with the correspondent inverse function (eq.K.7) to generate values that are constrained between a minimum and a maximum (eq.K.8).

$$y = \ddot{t}^{-1}(y_{transformed}) = \frac{\exp[\frac{y}{s}] * y_{max} * (c - y_{min})/(y_{max} - c) + y_{min}}{1 + (c - y_{min})/(y_{max} - c) * \exp[\frac{y}{s}]} \quad \text{(K.7)}$$

$$y_{MM} = \ddot{t}^{-1}(y_{NN}) \quad \text{(K.8)}$$

where y_{MM} is the output of the metamodel (either μ_j or σ_j), and y_{NN} is the output of the RBF neural network,

Figure K.2 shows the form of the transformation function (from eq.K.5) with the chosen parameters. s and c were chosen rather empirically (see section K.4: results and discussion below). $y_{min} = 1e{-}10\ \mu g/l$ as explained before and $y_{max} = 50\ \mu g/l$ because it is an integer above but very close to the maximal observed value in the outputs of the deterministic model.

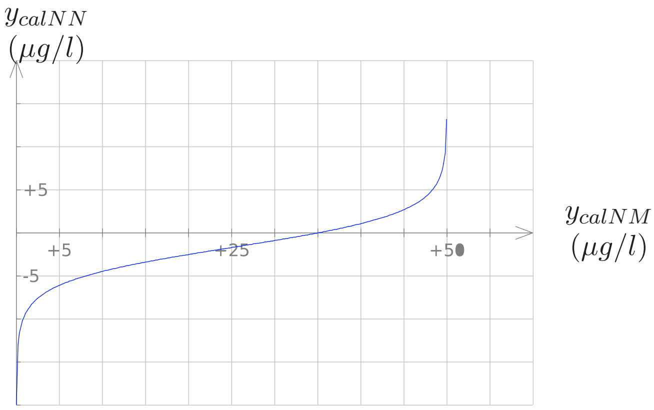


Figure K.2: *Transformation of the outputs of the deterministic model (i.e. the inter-annual mean and stddev of the annual leaching concentration), based on a argth-like function, prior to calibration of the neural network.* The equation is described in eq.K.5, with parameters $s = 2$, $y_{min} = 1e^{-10}\mu g/l$, $y_{max} = 50\ \mu g/l$, $c = 35\ \mu g/l$

K.3.3 Calibrating, assessing and running the metamodel

For the nitrate metamodel presented in Piñeros Garcet et al. (2006c), a sampling design was used to choose the calibration data (nitrogen fertilisation). Such an approach is not possible here because the possible parameter combinations for soil and weather are not known in advance, and must be derived from soil and weather databases. The sampling design used here is thus quite different from the one used in Piñeros Garcet et al. (2006c).

In order to obtain such a sample, simulations were performed by means of the spatially distributed model EuroPEARL, for 1072 “plots” (unique combinations of soil type, climate and country, sampled throughout the EU-15). Plots were defined only for the dominant STU that could be linked to the soil profiles database SPADE-1[1]. Daily weather data were obtained from 9 points of the MARS database, and scaled to cover the rest of EU-15. Other data, such as crop data and pesticide properties, have been compiled from various sources, such as inventories, field-studies and the literature (see Tiktak et al. (2004)). The 1072 unique

[1] SPADE-1: see Hiederer et al. (2006), JRC (2009).

combinations together represent at maximum 75% of the total agricultural area of the European Union. (52% if the area of the STU inside the SMU is taken into account). Austria, Sweden and Finland could not be included in the simulations, because there was insufficient soil profile information for these countries.

The simulations performed for 1072 ‘‘plots’’ allows to obtain the calibration and assessment data for the metamodel. 20 annual leaching concentration values obtained for each plot were used to calculate the inter-annual mean and standard deviations of annual leaching concentration for those plots.

Once the metamodel is calibrated and assessed, it can be run on the European scale soil and climate databases. The main databases used in the present work as inputs for the metamodel are: (1) for climate, Knoche et al. (1998) provides $5x5$ km^2 raster maps for mean annual temperatures, mean annual precipitation, mean rainfall in spring, and mean rainfall in autumn (spring means: March, April, May, and autumn means: Sep, Oct., Nov.); (2) for soil parameters, we used the European soil database (JRC, 2011), which includes a 1:1.000.000 soil map (Daroussin et al. (2006), Jamagne et al. (2002), JRC (2001)) which was provided to us as a $10x10$ km^2 raster map. As an example, fig. K.3 show a texture raster map for the dominant surface textural class of the STUs in the European soil database. According to JRC (2001), the spatial variability of soils is very important and is difficult to express at global levels of precision, thus the 1:000.000 map is a simplified representation of the diversity and spatial variability of the soil coverage. The database contains a list of Soil Typological Units (STU). Besides the soil names they represent, these units are described by variables (attributes) specifying the nature and properties of the soils. At that scale, it is not feasible to delineate the STUs. Therefore they are grouped into Soil Mapping Units (SMU) to form soil associations and to illustrate the functioning of pedological systems within the landscapes. The variables were estimated over large areas by expert judgement rather than measured on local soil samples. This expertise results from synthesis and generalisation tasks of national or regional maps published at more detailed scales. Delineation of the Soil Mapping Units is also the result of expertise and experience.

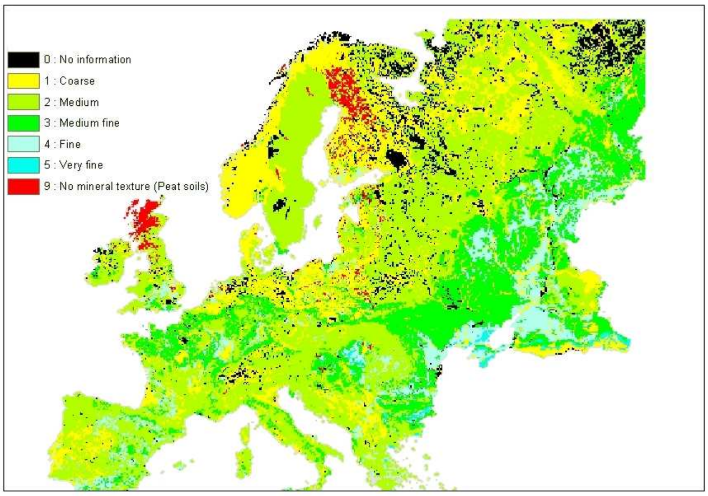


Figure K.3: *Dominant surface textural class of the STUs in the European soil database* (JRC, 2011).

Figure K.4 presents the data handling in order to calibrate, assess and run the metamodel on the pan-European soil and climate maps. The following steps were followed:

1. The dominant soil typological units (DomSTU) from the soil map are linked to soil profiles, in order to generate $\mathcal{X}_{NM}$, the set of inputs for the deterministic model. Then the deterministic model is run and mean and stddev is estimated ($\mathcal{Y}_{NM}$).
2. Soil deterministic model inputs are transformed by a soil horizons weighted mean into $\mathcal{X}_{calMM}$, the calibration inputs for the metamodel (top/sub soil values for the same parameters as in the map).
3. Using $\mathcal{X}_{calMM}$ and $\mathcal{Y}_{NM}$, a calibration set $(\mathcal{X}_{calNN}, \mathcal{Y}_{calNN})$ and a assessment set $(\mathcal{X}_{valNN}, \mathcal{Y}_{valNN})$ are constituted for calibrating a neural network (NN).
4. In order to estimate the mean and standard deviation $\mathcal{Y}_{MM}$, pedotransfer functions (ptf) are applied for some soil parameters, the metamodel inputs $\mathcal{X}_{MM}$ and then the metamodel is run over the soil and climate maps (including, for the soil map, all the typological units and not only the dominant STU).

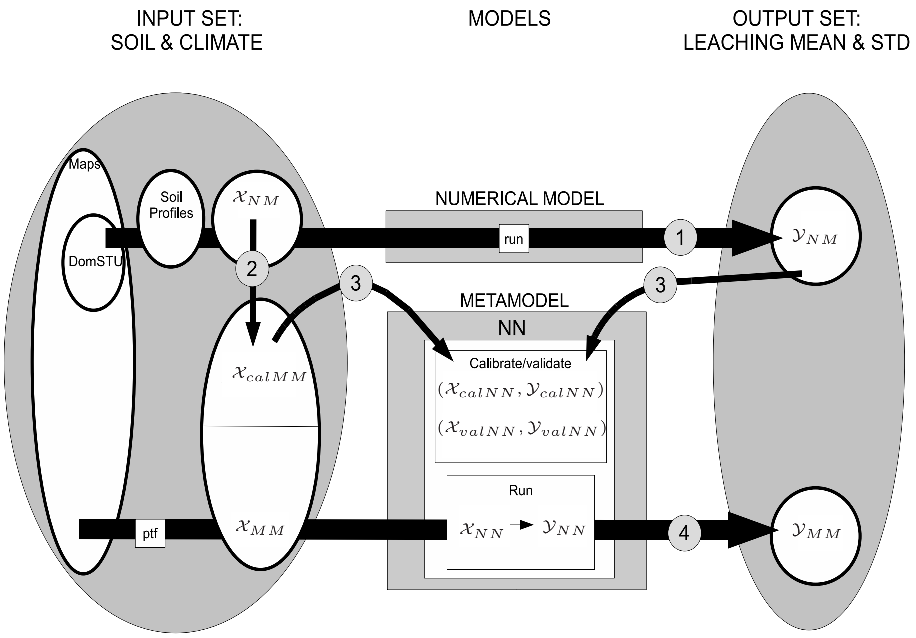


Figure K.4: *Overview of the chosen method to calibrate the metamodel, assess and run it over the soil and climate maps in order to obtain an estimation of the mean and standard deviation of the pesticide leaching.*

K.3.4 Novelty detection and applicability domain

As opposed to the work done with nitrate in Piñeros Garcet et al. (2006c), the calibration data for pesticide with the present RBFNN does not come from a well defined design region and are not sampled through a classical statistical experimental design. Instead, data choice is dependent on the availability of a link between the soil map and the SPADE-1 profiles database. In consequence, the relationships between the inputs and the outputs as captured by e.g. a simple linear statistical model are likely to be biased by the sampling and by possible correlations between the inputs.

The RBFNN are based on localized functions, thus highly adaptable to complex input spaces with correlations between the inputs, their drawback is that the prediction error rapidly becomes high when the inputs are located far from the original calibration inputs. As a consequence, a method for assessing the applicability domain of the model is needed.

Novelty detection is the identification of new or unknown data or signal that a machine learning system is not aware of during training (Markou and Singh, 2003). For calculating the novelty, the input vectors are considered as being sampled from a multidimensional probability function, and the degree of novelty of an input vector $\overrightarrow{x_{NN}}$ is estimated by the unconditional probability density $\hat{p}(\overrightarrow{x_{NN}})$, corresponding to the calibration data set. According to Bishop (1994), $\hat{p}(\overrightarrow{x_{NN}})$ for a new value of $\overrightarrow{x_{NN}}$ is called a likelihood. The likelihood function f is calculated as described in Bishop (1994) by using a Gaussian mixture model:

$$f(\overrightarrow{x_{NN}}) = \sum_j \alpha_j \Phi_j(\overrightarrow{x_{NN}}) \tag{K.9}$$

$$with \sum \alpha_j = 1 \tag{K.10}$$

The parameters α_j and the Gaussians Φ_j are known because they are estimated when calibrating the NN (see Piñeros Garcet et al. (2006c)).

Once the function to calculate f is established, it is used to define an applicability domain in the input space of the RBFNN. f is evaluated over the assessment data-set, and (as in Bishop (1994)) a threshold likelihood f_{min} was chosen in order that all the assessment inputs have a greater likelihood than f_{min}.

The applicability domain is defined here as follows:

$$\overrightarrow{x_{RBFNN}} \in applicability\, domain\ if: \ f(\overrightarrow{x_{RBFNN}}) > f_{min} \tag{K.11}$$

The applicability threshold and applicability domain are used when the metamodel is run over the soil and climate maps in order to discard inputs that are novelties and so are likely to lead to high prediction errors.

K.3.5 Metamodel's prediction intervals estimation

When describing confidence intervals on linear models, Milton and Arnold (1990) distinguishes between inferences made about the estimated mean and inferences made about a single predicted value. This distinction is used here to construct a confidence interval on the single prediction values made by the metamodel. However, many classical methods are not applicable here since, as it will be shown in the results, the residuals shows both that there is heteroskedacity and that the errors are not normally distributed.

In consequence, a non-parametric and rather empirical method was used to estimate confidence intervals around the metamodel estimations. The method is based on the residuals distribution (known from the assessment step): a graph of y_{NN} versus y_{valNN} is drawn, with $y_{NN} = RBFNN(\overrightarrow{x_{valNN}})$, and a moving window is moved along y_{NN}. From the corresponding sub-sets (windows) of y_{valNN}, the percentiles p_5 and p_{95} are calculated and back-transformed to y_{MM} by eq. (K.7).

K.4 Results and discussion

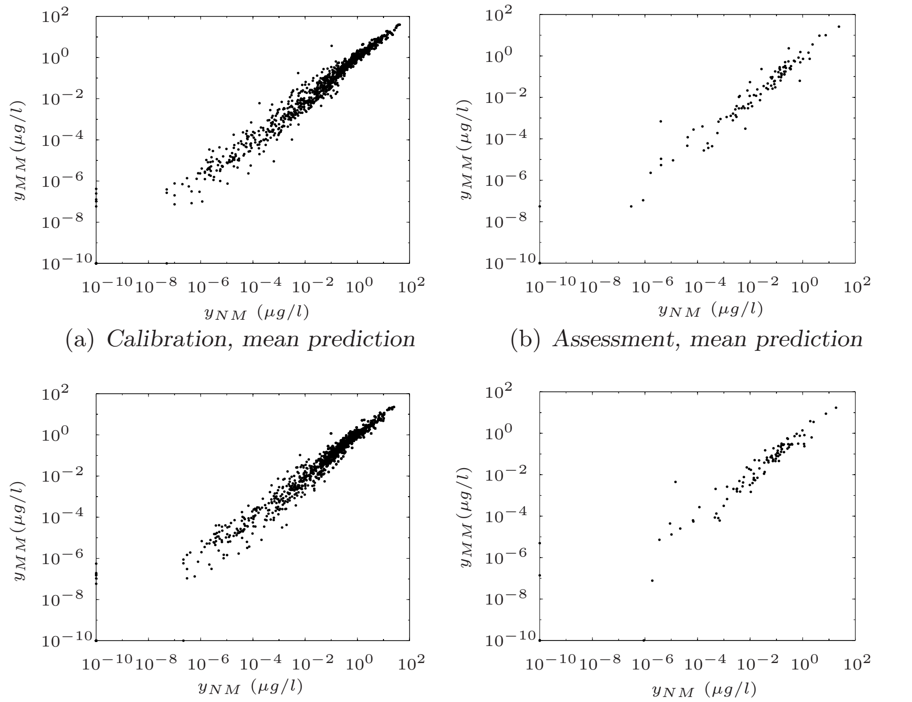
Two RBFNN were trained, in order to obtain the two outputs of the metamodel $y_{MM(mean)}$ and $y_{MM(stddev)}$, respectively for the inter-annual mean and standard deviation of the annual leaching concentration.

Based on the comparison between the EF values for calibrated neural networks using various values of t (the number of basis functions, p. 308), two networks were chosen. These networks have the following characteristics:

metamodel	t	EF_{val}
$y_{MM(mean)}$	185	0.931
$y_{MM(stddev)}$	195	0.969

This results show very good model efficiency in the assessment step (these model efficiencies are calculated from a sample that is not used for the calibration step, so those values are not due to overfitting). Model efficiency (from the assessment step) were calculated with an independent data set of inputs/outputs generated from a random choice amongst the original dataset of deterministic model simulations.

Figure K.5 presents a comparison of the calibration and assessment results of the metamodel, using scatterplots. The scatterplots shows a good correspondence between the deterministic model calculations and the metamodels both for the mean and the standard deviation. Slope is 1 in those log-log graphs, indicating no systematic bias in the predictions. The variability in the predictions by the metamodels remains within the order of magnitude of the deterministic model calculations, allowing to use the model to distinguish clearly between situations where leaching is



(a) *Calibration, mean prediction*

(b) *Assessment, mean prediction*

(c) *Calibration, standard deviation prediction*

(d) *Assessment, standard deviation prediction*

Figure K.5: *Calibration and assessment scatter plots for prediction of the inter-annual mean and standard deviation of annual leaching.* Abscissa: deterministic model outputs (y_{NN}); ordinate: metamodel outputs (y_{MM}).

higher or lower than $0.1\mu g/l$ (i.e. the legal limit[2]).

Another informative way of comparing the deterministic model with the metamodel is looking to the results of applying both to the same geographical data. Since EuroPEARL is run only on dominant STU, the same must be done with the metamodel for having a meaningful comparison. Such a comparison is presented in fig. K.6 that shows the maps of the inter-annual mean of annual leaching for both models. There is a good correspondence between the two. This means that the metamodel is able to reproduce the deterministic model behaviour when applied on

[2]The maximum permissible concentration laid down by Council Directive 80/778/EEC (1) of 15 July 1980 relating to the quality of water intended for human consumption.

the same input data. In order to construct that figure, the input data for the deterministic model are profiles derived from SPADE-1 soil profile database, and the metamodel inputs are derived from the same profiles by depth weighted means. Climatic data for the metamodel are also derived from the deterministic model inputs.

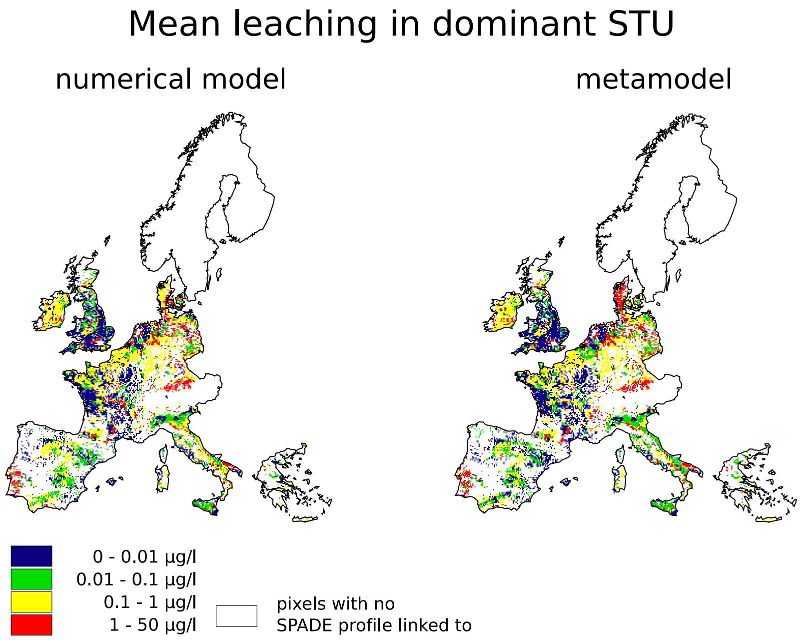


Figure K.6: *The deterministic model and the metamodel applied on the same dominant soil typological units (STU). The map displays the inter-annual mean of the annual leaching.*

Using the residuals calculated from the difference between the RBFNN predictions and the assessment data, a qqplot of the distribution of the residuals can be calculated (fig. K.7 and K.8). The qqplots shows that the residuals can not be adequately represented using a normal distribution. The non-normality of the distribution of the residuals is not the result of the the transformation made before calibration of the RBFNN (eq. (K.5)) because the qqplots are calculated from the transformed values (i.e. directly from the RBFNN results and not from the metamodel outputs $\mathcal{Y}_{MM}$).

In the transformation equation (eq.K.5), s and c parameters can be chosen rather empirically since the more important parameters are y_{min}

and y_{max}. However, high slopes will result in points that are more separate in the transformed values than in the original dataset. This is useful for allowing a good calibration of the RBFNN when parts of the dataset presents values very close to each other, since the first step of calibration is a clustering in the input space. Our calibration data presents (regarding outputs) a few high values of leaching and many very low values very close to each other. We chose then parameters c and s for having the majority of the calibration points below c. s is chosen to keep the very high and very low values of leaching in a sloppy part of the curve in order to obtain high distances between them in the transformed values. In consequence the transformation allows to maintain low residuals even with very small values of concentration, as shown in fig. K.9. In those figures, a moving window is used in order to plot percentiles p_5 and p_{95} (continuous line) based on the residuals on the corresponding window. Residuals decreases with decreasing concentrations.

Using the method described in section K.3.5 (calculation of a p_5 and p_{95}), a predictions confidence interval is calculated (fig. K.10 and K.11) from the assessment step residuals, both for the mean and standard deviation by the metamodel. Those confidence intervals confirm that the metamodel predictions remain within the order of magnitude of the deterministic model calculations.

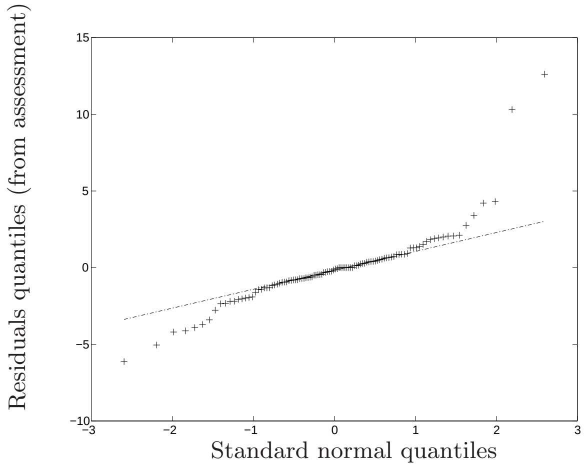


Figure K.7: *Qqplot from the assessment step residuals of the RBFNN for the transformed mean leaching.*

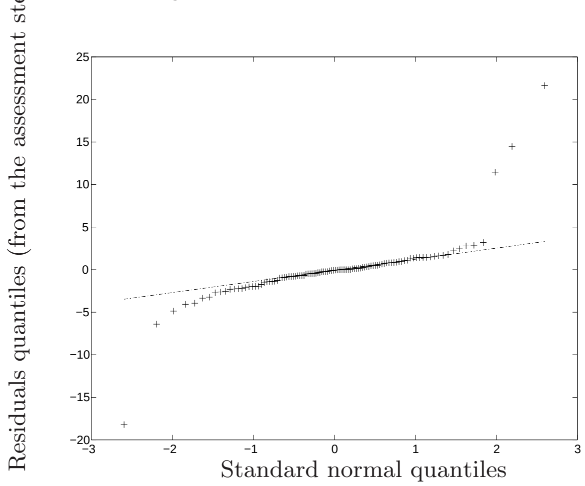
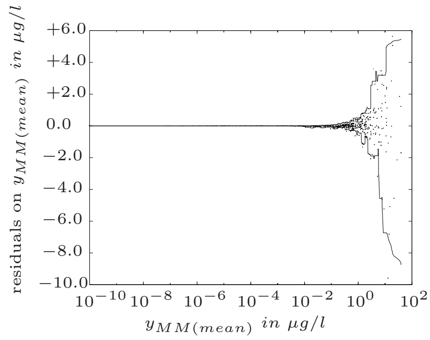
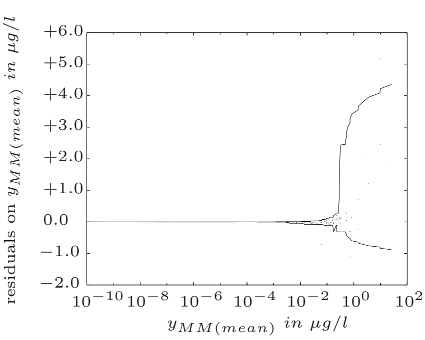


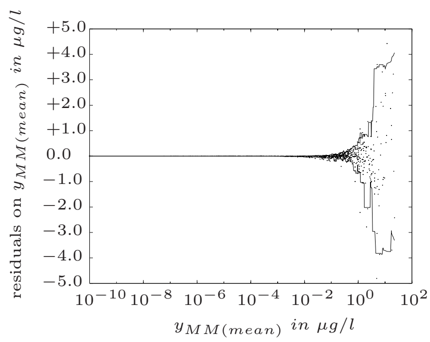
Figure K.8: *Qqplot from the assessment residuals of the RBFNN for the transformed standard deviation of leaching.*



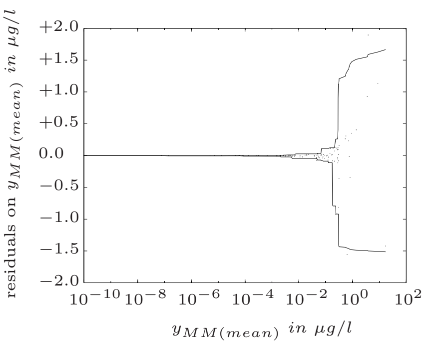
(a) *Calibration: residuals on the mean*



(b) *Assessment step: residuals on the mean*



(c) *Calibration: residuals on the standard deviation*



(d) *Assessment step: residuals on the standard deviation*

Figure K.9: *Residuals on the prediction of the mean and standard deviation by the metamodel, for the calibration and assessment points. Continuous line: residuals p_5 and p_{95} percentiles*

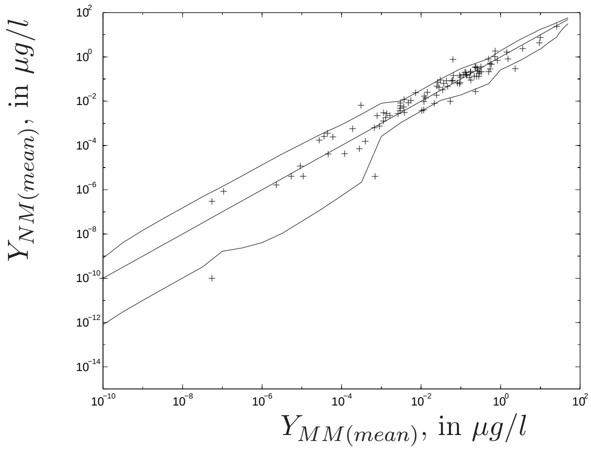


Figure K.10: *Prediction interval on the predicted mean $Y_{MM(mean)}$ calculated using the assessment step residuals*
______ $\equiv p_5$ *and* p_{95} of $Y_{NM(mean)}$

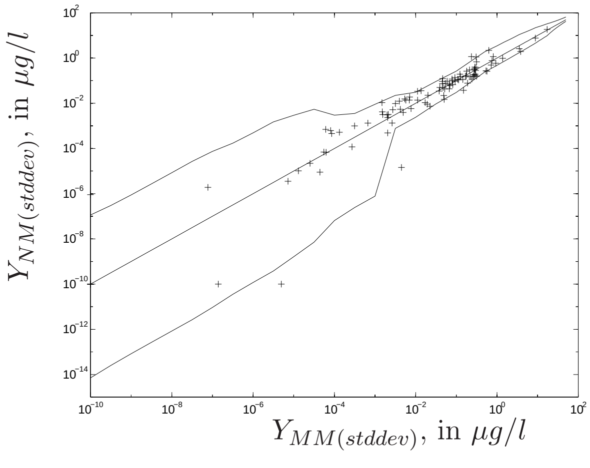


Figure K.11: *Prediction interval on the predicted standard deviation $Y_{MM(stddev)}$ calculated using the assessment step residuals*
______ $\equiv p_5$ and p_{95} *of* $Y_{NM(stddev)}$

In our metamodel, $y_{MM(mean)}$ represents a prediction of the inter-annual mean of the annual leaching, while $y_{MM(stddev)}$ represents a prediction of the inter-annual standard deviation of the annual leaching. $y_{MM(stddev)}$ is, in hydrological terms, representing the variability of annual leaching due to the weather (time). It must be noted that there is in our method a reduction in the number of parameters from the deterministic model to the metamodel. The consequence of this is that our methodology does not allow to distinguish in figures K.7, K.8, and K.9 between variability of the residuals coming from the metamodel fitting from the variability resulting from the reduction in the number of input parameters.

The issue of serial correlation in the deterministic outputs used for MM calibration has been discussed in section 6.5.1, and will not be recalled here.

Figure K.12 presents an approximation to the distribution of the novelty values of the climate and soil maps. As state in section K.3.4, the novelty of the inputs is estimated by their likelihood f. $\mathcal{X}_{MM}$ is the set $\{\overrightarrow{x_{MM,1}}, ..., \overrightarrow{x_{MM,n}}\}$ of the n input vectors coming from the climate and soil maps. In $\mathcal{X}_{MM}$, a vector $\overrightarrow{x_{MM,i}}$ can be repeated various times. From $\mathcal{X}_{MM}$, the set $\mathcal{UX}_{MM}$ of unique inputs $\{\overrightarrow{x_{MM,1}}, ..., \overrightarrow{x_{MM,m}}\}$ is constructed by eliminating the repeated vectors. The graph is representing the distribution of the values of $f(\mathcal{UX}_{MM})$.

All inputs $\overrightarrow{x_{MM}}$ having a likelihood of less than 10^{-20} in fig. K.12 are discarded from the calculation of the mean and standard deviation by the metamodel. This procedure is used in order to avoid having unrealistic values as output of the metamodel in regions where the metamodel predictions are of poor quality.

However, discarding some values is acceptable if they represent only a small area, and thus having a low weight in the final calculation of the cumulative probability by the upscaling procedure. Figure K.12 is constructed from unique values, so it does not allow to estimate those weights. Figure K.13 gives an approximation to the area, inside the FOCUS areas, that is novel for the MM. Figure K.13 shows effectively that the area with poor quality metamodel predictions is small. This figure is to be linked to the area coverage of the deterministic model. As stated in section 6.5.1, when intracell variability is considered, the deterministic model covers only 52% of the area of the EU-15 agricul-

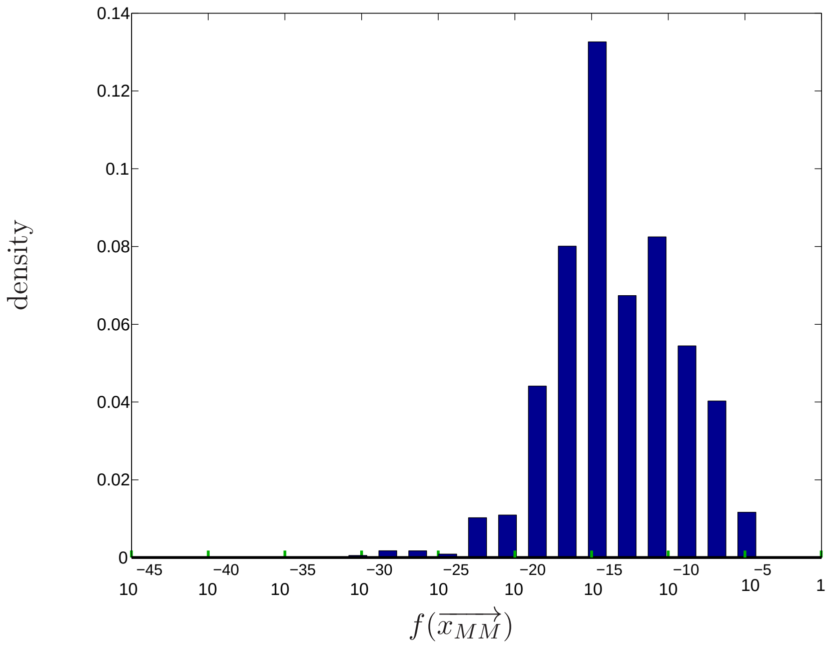


Figure K.12: *Distribution of the likelihood values of $\mathcal{X}_{MM}$, the set of map inputs.*

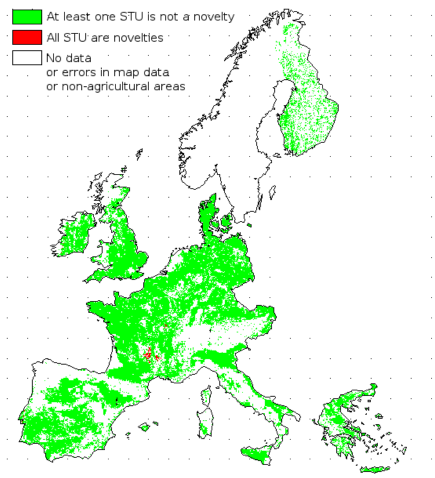


Figure K.13: *Pixels having at least one STU with data that are not novel (green) and pixels fow which all the STU's are novel (red). The threshold is a likelihood of* 10^{-20}

tural area. Figure K.12 and K.13 do not take into account the weight of the STU's inside the SMU's, and in consequence, the proportion of the STU's that is novel in the map. Using a database, the percentage of each STU inside the SMU's allows to give an estimation of the geographical coverage of the metamodel inside the agricultural area. The results are that the metamodel covers 71% of the agricultural area, there are 3% of the area that could not be modelled because it is novel, and 26% of the area have either missing data for some soil parameters or database errors (e.g. the sum of the proportion of the STU's is greater than 100% for various SMU's). Still, 71% is a significant improvement with respect to the 52% of EuroPEARL.

Figure K.6 presented a comparison of the results of the metamodel and the deterministic model when applied to the same input data. However, significative differences exists between the SPADE-1 database and the EU-15 soil map . This can be viewed e.g. when a comparison between the two is made regarding organic matter content (see fig. K.14).

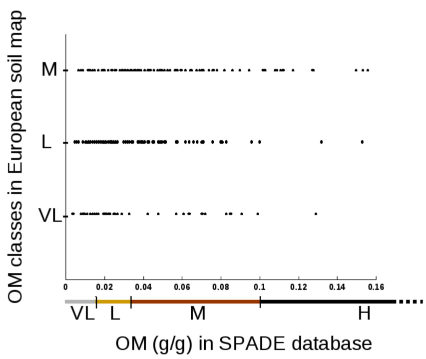


Figure K.14: *Relation between the topsoil organic matter content in the SPADE-1 database and in the European soil map. Organic matter in the European soil map is given in classes (VL = very low, L = Low, M = Medium and H = high).*

Such differences between SPADE-1 and the soil map obviously lead to differences between the leaching calculated using one or the other of the databases (fig. K.15). As a consequence, using the deterministic model with SPADE-1 and the metamodel with the soil map for assessing the FOCUS-GW scenarios will not lead to the same conclusion regarding their quality.

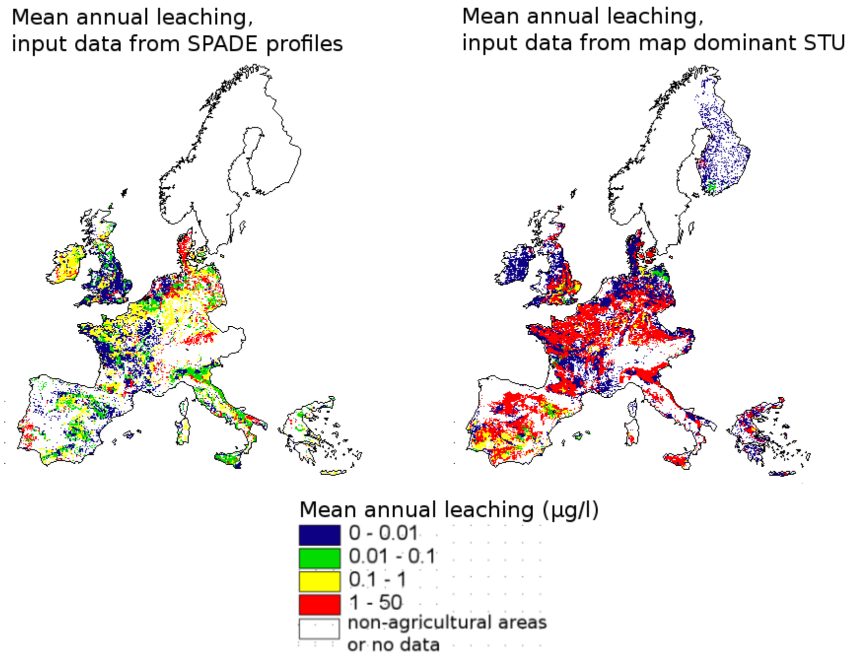


Figure K.15: *Comparison of the metamodel applied on the dominant STU, with: (1) data derived from the SPADE-1 profiles database, and (2) data derived from the soil map. The map displays the inter-annual mean of the annual leaching.*

Besides the greater geographical coverage, one of the arguments for using the metamodel is that it can take into account intra-cell variability by using all the STU information and not only the dominant STU as EuroPEARL does. Figure K.16 shows a comparison of the mean leaching per pixel, calculated with our without intracell variability. The SMU mean (i.e. the weighted mean of the STU's mean annual leaching) can be very different from the mean of the corresponding dominant STU. Avoiding to take intracell/intrapolygons variability in maps, when modelling leaching, can thus lead to significant bias (e.g. in EuroPEARL calculations). Another way of looking to the influence of intracell variability is a scatter plot of the SMU mean (with intracell variability) versus dominant STU mean leaching (see fig. K.17).

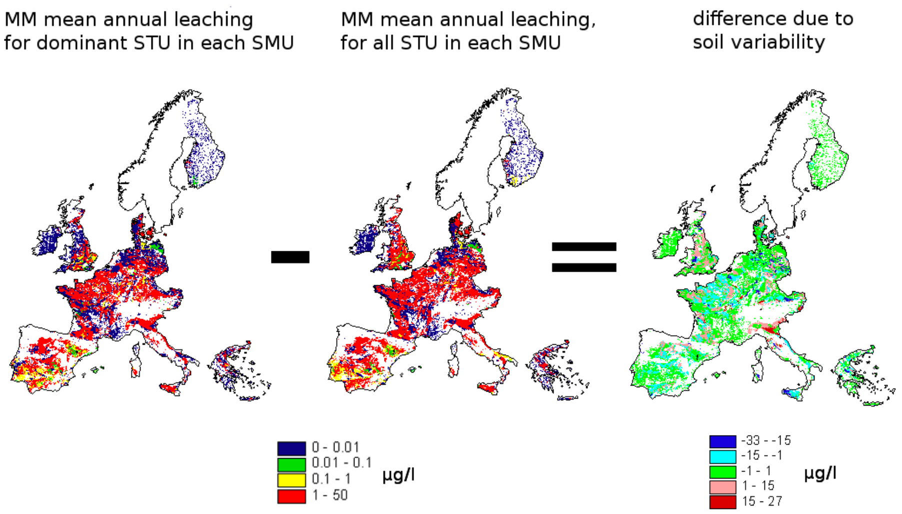


Figure K.16: *Maps showing the influence of the variablility of the soil typological units (STU), inside the SMU, on the inter-annual mean of the annual leaching concentration.*

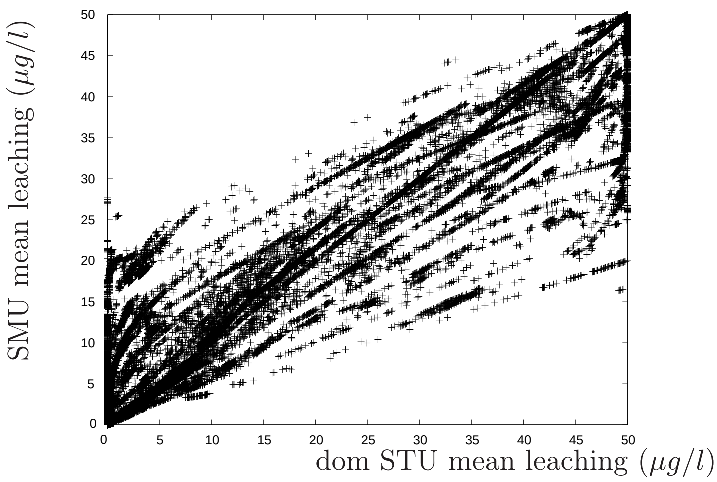


Figure K.17: *Scatterplot showing the influence of the variability of the soil typological units (STU) inside the soil mapping units (SMU).*

K.5 Conclusion

The present metamodel builds on the previous experience of a nitrate metamodel presented in Piñeros Garcet et al. (2006c) by improving the following aspects:

- It allows for the estimation of the inter-annual leaching variability, due to the estimation of both the standard deviation and the leaching mean. This inter-annual variability estimation is strictly necessary for FOCUS-GW scenarios quality assessment, and is thus an essential element of the upscaling and representativeness assessment procedure described in section 6.5.1.
- An artanh based function (eq.K.5) is used in order to constrain the outputs values of the MM to positive values above a chosen minimum.
- A statistical novelty detection procedure (as defined above) allows to avoid using the MM outside a well defined "applicability domain".
- A prediction interval is calculated for the predictions made by the MM.

The metamodel has the following advantages over the deterministic model presented by Tiktak et al. (2004):

- it can make a direct use of the EU soil map database, without relying on a problematic link between the soil map and the SPADE-1 profiles database;
- the error introduced by the link between SPADE-1 and the soil map is not quantifiable using the deterministic model (a qualitative estimation of the link between the two is possible however), while the metamodel gives an explicit estimation of the prediction interval;
- it can estimate leaching over a greater area, even in areas where a link between the map and the SPADE-1 profiles database can not be established;
- it allows to consider intracell raster map variability for soils; and
- the RBFMM can use directly pan-European climate databases, without time-series interpolation, as opposed to the EuroPEARL parametrization scheme.

Tiktak et al. (2006) presented a metamodel which is a robust linear regression fitting (RLR-MM) with the independent regression variables being based on process-based equations. The work presented here (RBFNN-MM metamodel development and application to the EU-15 maps) can be seen as an improvement over the work done in Tiktak et al. (2006) because:

- The RBFNN-MM takes into account vertical variation of the soil parameters, through the topsoil and subsoil parameters.
- The model efficiency EF of RLR-MM for substance D, autumn application is 0.934. That EF is to be interpreted as being calculated from the calibration data, because Tiktak et al. (2006) stated that they ‘‘applied the metamodel to the same dataset as used for the parameterization of EuroPEARL’’. RBFNN-MM EF_{cal} is 0.952 which is in consequence higher.
- Tiktak et al. (2006) gave no estimation of EF for the assessment step (‘‘validation’’): model efficiency of RLR-MM calculated with an independent data set (EF_{val}) was not calculated, while RBFNN-MM has a EF_{val}=0.931.
- The RBFNN-MM incorporates a novelty detection method, allowing the user to avoid applying the model outside the ‘‘applicability domain’’.
- The intracell variability is taken into account in RBFNN-MM. The area covered by the RLR-MM is 75% of the agricultural area only if each pixel is supposed to have it's entire area covered by the dominant STU. However, the non dominant STU's cover up to 48% of the SMU area. The RLR-MM is thus in fact covering 52% of the area. In consequence, the area covered by RBFNN-MM is much higher than the area covered by RLR-MM if this is taken into account: 71% instead of 52%.
- The entire area is covered with a single model, while Tiktak et al. (2006) had to split the area in 4 climatic zones in order to obtain satisfactory modelling results.

However the metamodel presented in Tiktak et al. (2006) has the following advantages:

- It needs less input parameters.
- It is simpler and hence easier for communication and transparency in a regulation context.
- It is an elegant combination of a process based approach with a statistical model.

- It has a much larger calibration dataset.
- It includes pesticide parameters. This was not considered feasible with the small dataset available for the present study.

Significant differences exists between the EuroPEARL and RBFMM EU-15 leaching estimations. The RBFMM model efficiency being high (0.93 for the mean, and 0.97 for the standard deviation), those differences are not to be attributed to metamodel fitting errors. Those differences can not be attributed to extrapolation errors, because the metamodel incorporates a novelty detection of the inputs. However, significant uncertainties and errors exists in the soil databases used both by EuroPEARL and the RBFMM, and this leads to unquantified errors in both approaches when estimating leaching at the point scale. EuroPEARL uses the SPADE-1 database to estimate soil properties, while the metamodel directly uses the European Soil Map 1:1.000.000. One would expect a good correlation between soil properties at the European Soil Map and SPADE-1, but unfortunately this is not true. This is particularly the case for the extremely poor relationship between organic matter derived from the soil map and organic matter derived from SPADE-1. We conclude that improvement of the pan-European databases should be a main research task for the upcoming years. Differences between the RBFMM and EuroPEARL are also due to the fact that EuroPEARL uses dominant STUs, while the metamodel uses all STUs. The incorporation of variability at lower spatial scales usually increases the mean leaching concentration (der Zee and Boesten, 1991). This effect is also verified in the present work with the metamodel, as shown by the comparison of a leaching map with all the STUs versus a map with only the dominant STUs.

List of Figures

List of Tables

Symbols

Greek symbols

β	constant used in a MKM definition
γ	a set of outputs (of a model or metamodel)
δ_M	design region for caibration and assessment of a metamodel
ε	error term
Θ_i	a parameter of the correlation component of a MKM, to be optimized for calibrating it
μ	mean of a distribution
$\overrightarrow{\ddot{\mu}_j}$	vector of means of the j^{th} gaussian basis function of a RBF NN
μ_j	point scale mean of the annual leaching concentration at 1 m depth
μ_X	mean of the mixture distribution describing the leaching into a scenario area
$\mu_{X,i}$	μ_X for scenario area i
σ	standard deviation of a distribution
$\ddot{\sigma}_\mathbf{j}$	standard deviation of the j^{th} gaussian basis function of a RBF NN
σ^2	variance of a distribution
σ_j	point scale standard deviation of the annual leaching concentration at 1 m depth

σ_X	standard deviation of a mixture distribution describing the leaching into scenario area
$\sigma_{X,i}$	σ_X for scenario area i
ϕ	description domain: the space of all the possible levels, alternatives, numbers of the symbolic variables
ϕ_g	geographical component of a description domain
ϕ_{cdf}	probability component of a description domain
Φ_j	the j^{th} basis function of a RBF NN
χ_N	the set of all the possible inputs of a deterministic model
χ_M	the set of all the possible inputs of a metamodel model
$\psi(\omega_i)$	the vector describing ω_i: $\psi(\omega_i) = (Y_1, .., Y_n)$
ω_i	an individual: the symbol used to idexically locate (represent) a phenomenon in the logical space of the scenarios
Ω	extension universe: the set of all individuals (e.g. a population, in contrast to a sample)

Numeric and latin symbols

$\mathbf{1}_n$	$\mathbf{1}_n = (1, ..., 1)'$
a	member ship function: allows either to decide which individuals ω are members of ESE or to estimate the degree of membership of a given ω_i to ESE
AW_s	available water content in subsoil (m^3/m^3)
AW_t	available water content in topsoil (m^3/m^3)
b	a number below μ_X
bf	number of basis functions of a RBF NN
C_t	clay fraction in topsoil (g/g)
C_s	clay fraction in subsoil (g/g)
d	description vector: e.g. $\{25°C, 100\, mm\, mean\, annual\, rainfall\}$
$d\omega_i$	description of ω_i: the description of ω_i using a vector $\psi(\omega_i)$ of symbolic variables Y_i
DRE	descriptive reference extension
$DRE_{max,i}$	upper bound estimation of DRE for scenario i
DSE	descriptive scenario's extension: the array of the description vectors of individuals in ESE
DSE_{cdf}	cumulative probability component of the descriptive extension
DSE_g	geographical component of the descriptive extension
DSE_i	the DSE for the i^{th} scenario

EF	model efficiency
EF_{cal}	model efficiency calculated from the calibration data
EF_{val}	model efficiency calculated with an independent data set
ESE	enumerative scenario's extension: is the set of the individuals that are reprented by SI and is termed "enumerative" to highlight that it enumerates the individuals and do not refer directly to the phenomena (which in fact are only accessible through indexical location which is not a mathematical operation)
$EXT(SI)$	extension mapping of SI: extension mapping of SI into Φ
$EXTR$	reference extension mapping: extension mapping used as reference for quality assessment
EXT_g	extension mapping for defining the geographical component of an extension
EXT_{cdf}	extension mapping for defining the cumulative probability component of an extention
F	cumulative distribution function (cdf)
f	likelihood function
F_R	reference cumulative probability function (cdf)
F_{Rmax}	a maximal reference estimation of the cumulative probability of the concentration
f_K	a MKM metamodel
f_M	a metamodel
f_{min}	a minimal threshold chosen for f
f_N	a deterministic model
f_{RBF}	a RBF metamodel
g	transformation function of $\overrightarrow{x_{N,i}}$ into $\overrightarrow{x_{M,i}}$
$\mathbf{I}$	identity matrix
I_n	identity matrix of size nxn: diagonal matrix with elements equal to 1 in the diagonal, and size being nxn
LU	land use
$MATCH$	matching function: a function giving the degree of matching between two extensions (i.e. between two sets)
MKM	multidimensional kriging metamodel
MM	metamodel
NM	deterministic model
NN	neural network
$\ddot{O}_i$	"observed" output values (i.e. calculated with the deterministic model)

OM_t	organic matter content in topsoil (g/g)
OM_s	organic matter content in subsoil (g/g)
P_i	parameters of the i^{th} FOCUS-GW scenario
p_i	a parameter of the correlation component of a MKM, to be optimized for calibrating it
$\ddot{P}_i$	output values predicted by the metamodel
p_j	weight of a distribution into a mixture distribution: calculated here from the proportion of the STUs in the pixels of the map, and the proportion of each pixel in the scenario area i
p_5	the 5th percentile of a probability distribution: if $x = p_5(X)$ then the probability that a random variable X will be less than x is at most 5/100
p_{90}	the 90th percentile of a probability distribution: if $x = p_{90}(X)$ then the probability that a random variable X will be less than x is at most 90/100
$\hat{p}(\overrightarrow{x_{NN}})$	unconditional probability density of $\overrightarrow{x_{NN}}$
pH_t	pH of topsoil $(pH\ KCl\ 1N)$
pH_s	pH of subsoil $(pH\ KCl\ 1N)$
PEC_{GW}	predicted environmental concentration in groundwater: the concentration of a given compound in groundwater, as calculated using a model (at 1 m depth in the FOCUS case).
PD_s	packing density of subsoil (g/cm^3)
PD_t	packing density of topsoil (g/cm^3)
R	a mathematical relation: e.g. $\{<, >\}$
$\mathbf{R}$	correlation matrix between all input calibration points in a MKM
R_a	mean autumn rainfall (mm)
R_s	mean spring rainfall (mm)
R_y	mean annual rainfall (mm)
$\mathbf{r}_x$	correlation between the input vector $\overrightarrow{x_M}$ and the calibration points in a MKM
RBFNN	radial basis functions neural networks metamodel
RE	reference extension for a scenario
S_t	sand fraction in topsoil (g/g)
S_s	sand fraction in subsoil (g/g)
S_i	the i^{th} scenario
S^2_{pop}	population variance
S^2_{sample}	sample variance

SC	scenario's representing concept: the concept that is used to represent ESE
SE	extension of a scenario
SI	scenario's intension: the scenario's intension is a symbolic object
SI_i	the i^{th} SI
SMU	Soil Mapping Unit
STU	soil typological unit
T_y	mean annual temperature ($°C$)
t	time (year)
$\ddot{t}$	transformation function for the outputs of the deterministic model
$\mathcal{UX}_{MM}$	set of of unique inputs of a MM
w	weight in a RBF NN
$\underline{X}$	the symbolic data array: the database available for the descriptions
$\mathcal{X}_{calMM}$	matrix of MM calibration inputs
$\mathcal{X}_{calNN}$	matrix of NN calibration inputs
$\overrightarrow{x_M}$	metamodel inputs vector
$\overrightarrow{x_N}$	deterministic model inputs vector
$\mathcal{X}_{MM}$	the set of metamodel inputs
$\overrightarrow{x_{MM}}$	an element of $\mathcal{X}_{MM}$
$\overrightarrow{x_{MM,i}}$	the ith element of $\mathcal{X}_{MM}$
$\mathcal{X}_{NM}$	the set of inputs for the deterministic model
$\mathcal{X}_{NN}$	the set of NN inputs
$\overrightarrow{x_{NN}}$	input vector for a NN
$\mathcal{X}_{valNN}$	inputs assessment set for a NN
$\overrightarrow{x_{M,i}}$	the i^{th} vector of metamodel inputs
$\overrightarrow{x_{N,i}}$	the i^{th} deterministic model input vector (scenario)
$x^{i,j}_{calMM}$	the element (i,j) of $\mathcal{X}_{calMM}$
$x^{.,j}_{calMM}$	the j^{th} column of $\mathcal{X}_{calMM}$
$x^{i,j}_{calNN}$	the element (i,j) of $\mathcal{X}_{calNN}$
x^i_M	the ith component of $\overrightarrow{x_M}$
x^i_N	the ith component of $\overrightarrow{x_N}$
(x,y)	the geographical coordinates of a location
Y_i	a symbolic variable: a mathematical variable that may have values such as subsets of categories, intervals, or frequency distributions
y_{calNM}	the outputs calibration set from the deterministic model

y_{calNN}	the outputs calibration set for the RBF neural network model
$\mathcal{Y}_{calNN}$	outputs calibration set for a NN
y_{MM}	output of the metamodel
$\mathcal{Y}_{MM}$	the set of the MM outputs
$\overrightarrow{y_N}$	deterministic model output
$\mathcal{Y}_{NM}$	the set of NM outputs
$\mathcal{Y}_{NN}$	the set of NN outputs
y_{NN}	output of a neural network
$\mathcal{Y}_{valNN}$	outputs assessment set for a NN
y_{valNN}	assessment output for a NN
$\overrightarrow{y_{M,i}}$	the i^{th} metamodel output vector
$\overrightarrow{y_{N,i}}$	the i^{th} deterministic output vector
$Y_{NM,t}$	output of a deterministic model for year t
Z	random function in a MKM

Other symbols

$'$	A' is the transpose of a matrix A

Bibliography

Advisory Comittee on pesticides, UK. An overview of the 102nd environmental panel meeting held on tuesday 19th february 2008, 2008. URL http://www.pesticides.gov.uk/acp.asp?id=2625. Date of access: 2009/01/08.

L. Albarello. *Apprendre à chercher. L'acteur social et la recherche scientifique*. De Boeck, 2d edition, 2002.

L. Albarello, F. Digneffe, J.-P. Hiernaux, C. Maroy, D. Ruquoy, and P. de Saint-Georges. *Pratiques et méthodes de recherche en sciences sociales*. Armand Colin, Paris, 1995.

S. Alsop. Understanding understanding: a model for the public learning of radioactivity. *Public Understanding of Science*, (8):267--284, 1999.

P. Andrews. Churchś type theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2009. URL http://plato.stanford.edu/archives/spr2009/entries/type-theory-church/. Spring 2009.

Anonymous. Council directive 91/414/eec of 15 july 1991 concerning the placing of plant protection products on the market. official journal of the european union. oj l 230, 19.8.1991, p. 1–32, 1991.

Anonymous. Registration of agrochemicals in europe. *Crop protection monthly*, July 2002. Issue 152. Report of the Conference on Registration of Agrochemicals in Europe, 25-26 June 2002 in London.

Anonymous. 28 mars 2003 - Loi modifiant la loi du 21 décembre 1998 relative aux normes de produits ayant pour but la promotion de modes de production et de consommation durables et la protection de l'environnement et de la santé. Belgian Official Journal: M.B. 29.04.2003, 2003.

Anonymous. Registration of agrochemicals in europe. *Crop protection monthly*, February 2003. Issue 159. Report of the 10th Annual Conference on Registration of Agrochemicals in Europe, 5-6 February 2003 in London.

Anonymous. 22 fÉvrier 2005 - arrêté royal relatif au premier programme de réduction des pesticides à usage agricole et des biocides. Belgian Official Journal: M.B. 11.03.2005, 2005.

F. Ansay, L.-M. Delescaille, F. Goor, and M.-C. Godin. Le dépérissement des ruchers : un symptôme de problèmes environnementaux ? In Cellule Etat de l'Environnement Wallon, editor, *Rapport analytique sur l'état de l'environnement wallon 2006-2007*, chapter FFH1 Les milieux agricoles, page 736pp. Ministère de la Région Wallonne - DGRNE, Namur, 2007. URL http://environnement.wallonie.be/eew/files/rapport2006/publication/FFH_01.pdf.

APECOP. Technical annex, apecop:effective approaches for assessing the predicted environmental concentrations of pesticides: A proposal supporting the harmonised registration of pesticides in europe. Technical report, European Commission, 1999.

APECOP. Minutes of the meeting of apecop with industry, at rivm, bilthoven, the netherlands, July 2002.

C. Atias. *Épistémologie du droit*. Que sais-je? PUF, Paris, 1994.

M. Augé. *Symbole, fonction, historie. Les interrogations de l'anthropologie*. Hachette, Paris, 1979.

R. Barbaras and J. Greisch. Phénoménologie. In *Encyclopaedia Universalis*. Encyclopaedia Universalis France, 1996.

B. Barnes. *Scientific Knowledge and Sociological Theory.* Routledge & Kegan Paul., London, 1974.

B. Barnes. *Interests and the Growth of Knowledge.* Routledge & Kegan Paul., London, 1977.

B. Barsotti. *La représentation dans la philosophie contemporaine.* Philo. ellipses, Paris, 2001.

C. Baylon and X. Mignot. *La communication.* linguistique. Nathan Université, 1999.

H. Bergson. *La pensée et le mouvant : Essais et conférences*, chapter Sur le pragmatisme de William James. Vérité et réalité, page 294pp. Collection : Bibliothèque de philosophie contemporaine. Éditions des Presses universitaires de France, 79th edition, 1969.

C. Bishop. Novelty detection and neural network validation. In *Proceedings of IEE Conference on Vision and Image Signal Processing*, pages 217--222, 1994.

C.M. Bishop. *Neural networks for pattern recognition.* Oxford University Press, 1995.

H.H. Bock and E. Diday, editors. *Analysis of symbolic data: exploratory methods for extracting statistical information from complex data.* Springer-Verlag, 2000.

J. Boesten, R.L. Jones, M. Businelli, A.B. Delmas, B. Gottesbüren, K. Hanze, T. Jarvis, M. Klein, A.M.A. van der Linden, W.M. Maier, S. Rekolainen, H. Resseler, M. Styczen, K. Travis, and M. Vanclooster. The development of focus scenarios for assessing pesticide leaching to ground-water in eu registration. In *The 1999 Brighton conference - Weeds*, volume 2 of *Conference Proceedings*, pages 537--542, Farnham (UK), 1999. British Crop Protection Council.

J.J.T.I. Boesten. From laboratory to field: uses and limitations of pesticide behaviour models for the soil/plant system. *Weed Research*, 40 (1):123--138, 2000. URL `http://dx.doi.org/10.1046/j.1365-3180.2000.00158.x`.

P. Bourdieu. La spécificité du champ scientifique et les conditions sociales du progrès de la raison. *Science et structure sociale*, 7(1): 91--118, May 1975.

P. Bourdieu. Le champ scientifique. *Actes de la recherche en sciences sociales*, 2(3):88--104, 1976.

P. Bourdieu. *Science de la science et réflexivité*. Cours et travaux. Raisons d'agir, Paris, 2001.

M. Callon. Éléments pour une sociologie de la traduction. la domestication des coquilles saint-jacques et des marins-pêcheurs dans la baie de saint-brieuc. *L'année sociologique*, pages 169--208, 1986.

M. Callon. Des formes de démocratie technique. *Annales des mines, responsabilite & environnement*, (9):63--73, January 1998.

M. Callon and J. Law. L'irruption des non-humains dans les sciences humaines: quelques leçons tirées de la sociologie des sciences et des techniques. In B. Reynaud, editor, *Les limites de la rationalité. 2. Les figures du collectif*, pages 99--118. La Découverte, 1997.

M. Canivet. Philosophie, 2002. URL http://www.icampus.ucl.ac.be/philo/. Date of access: 20/11/2002.

R. Carnap. *La construction logique du monde*. Vrin, 2002-1928. french tr. by T. Rivain.

H. Caswell. Theory and models in ecology: a different perspective. *Ecol. Model.*, (43):33--44, 1988.

A. Chalmers. *La Fabrication de la science*. La Découverte., Paris, 1991.

Chemicals Regulation Directorate. The ec review programme for existing active substances, 2010. URL http://www.pesticides.gov.uk/approvals.asp?id=2317. Date of access: 2010/10/08.

E.J. Chen and W.D. Kelton. Quantile and tolerance-interval estimation in simulation. *European Journal of Operational Research*, 168: 520--540, 2006.

E.J Chen and W.D Kelton. Estimating steady-state distributions via simulation-generated histograms. *COMPUTERS & OPERATIONS RESEARCH*, 35(4):1003--1016, APR 2008. ISSN 0305-0548. doi: {10.1016/j.cor.2006.05.015}.

P.T. Coggi. Proposal for ep and council regulation on placing plant protection products on the market. eu regulatory developments, 2006. URL http://www.ecpa.be/files/ecpa/documentslive/24/15740_1%20-%20Paola%20Testori%20Coggi.pdf. Paola Testori Coggi, Acting Deputy Director General, European Commission, DG SANCO / ECPA Annual Conference, Limassol, Cyprus 19-21 October 2006. Date of access: 2007/01/08.

H.M. Collins. The tea set : tacit knowledge and scientific networks. *Science Studies*, 4:165--186, 1974.

H.M. Collins. Introduction: Stages in the empirical programme of relativism. *Social Studies of Science*, 11(1):3--10, 1981a. ISSN 03063127. URL http://www.jstor.org/stable/284733.

H.M. Collins. The place of the « core-set» in modern science: Social contingency with methodological propriety in science. *History of Science*, 19:6--19, 1981b.

H.M. Collins. Tacit knowledge and scientific networks. In B. BARNES and D. EDGE, editors, *Science in context*, pages 44--64. The Open University Press, Milton Keynes, 1982.

H.M. Collins. *Changing order. Replication and induction in scientific practice*. Sage Publ., London, 1985.

H.M. Collins. Les sept sexes: étude sociologique de la détection des ondes gravitationnelles. In Callon et Latour, editor, *La science telle qu'elle se fait*. La Découverte, 1991.

H.M. Collins and T. Pinch. *The Golem*. Cambridge University Press, 1993. Trad. fr. Tout ce que vour devriez savoir sur la science, Paris Le Seuil, 1994.

Commission of the European Communities. Communication of 1 july 2002 from the commission to the council, the european parliament and the economic and social committee - towards a thematic strategy on the sustainable use of pesticides. com(2002) 349 final., July the 1st 2002.

Commission of the European Communities. Commission staff working document report on the impact assessment for a regulation replacing directive 91/414/eec on plant protection products. sec(2006) 931, 2006.

G. Cook, E. Pieri, and P.T. Robbins. 'the scientists think and the public feels': Expert perceptions of the discourse of GM food. *Discourse Society*, (15):433--449, 2004.

J. Daroussin, D. King, C. Le Bas, B. Vrscaj, E. Dobos, and L. Montanarella. Chapter 4 the soil geographical database of eurasia at scale 1:1,000,000: History and perspective in digital soil mapping. In A.B. McBratney P. Lagacherie and M. Voltz, editors, *Digital Soil Mapping - An Introductory Perspective*, volume 31 of *Developments in Soil Science*, pages 55 -- 65, 602. Elsevier, 2006. doi: DOI:10.1016/S0166-2481(06)31004-5. URL http://www.sciencedirect.com/science/article/pii/S0166248106310045.

S.E.A.T.M. Van der Zee and J.J.T.I. Boesten. Effects of soil heterogeneity on pesticide leaching to ground-water. *Water Resour. Res.*, 27:3051--3063, 1991.

DG SANCO. Technical annex to report from the commission to the european parliament and the council on the evaluation of the active substances of plant protection products. Technical Report SANCO/2692/2001, European Commission, 2001.

DG SANCO. Workshop on the amendment of council directive 91/414/eec held on 10 - 12 july 2002 in corfu. conclusions of the workshop. Technical Report SANCO/10351/2002, European Commission, 2002. URL http://www.kemi.se/upload/Medier/Nyhetsarkiv/2002/Rapport_Korfu2002.pdf. Date of access: 2006/02/08.

DG SANCO. Overview of the state of main works in dg health and consumer protection d.3 with regard to the implementation of directive 91/414/eec. Technical Report SANCO/629/00 rev. 76, European Commission, January 2005.

DG SANCO. Progress with the EU review programme. L. Smeets, in Annual European Crop Protection Association Conference, Limassol., 19-21 October 2006. URL http://www.ecpa.be/files/ecpa/documentslive/24/15751_1%20-%20Louis%20Smeets.pdf. Date of access: 2007/03/08.

E. Diday and F. Esposito. An introduction to symbolic data analysis and the sodas software. *Intelligent Data Analysis*, 7:583--601, 2003.

M. Doury. *Le débat immobile, l'argumentation dans le débat médiatique sur les parasciences*. Éditions Kimé, Paris, 1997.

J.-P. Dumont. Phénomène. In *Encyclopaedia Universalis*. Encyclopaedia Universalis France, 1996.

ECPA. Commission proposal for a regulation of the european parliament and of the council concerning the placing of plant protection products on the market. [com(2006) 388 final] - ecpa position. Technical Report PP/06/EJ/15829 (Rev3), European Crop Protection Association, 2 October 2006.

ECPA. Ecpa's position on the implementation of the zonal authorisation and mutual recognition systems included in the new regulation on the authorisation of plant protection products, October 2009. URL http://www.ecpa.be/files/ecpa/documentslive/9/19085__ECPA%20position%20-%20Implementation%20of%20zonal%20process%20FINAL.pdf. Date of access: 2010/01/08.

A. Edgar. *Habermas: the key concepts*. Routledge key guides. Routledge, 2006. ISBN 9780415303781. URL http://books.google.com/books?id=BIVymZuLF_4C.

D. Edge. Quantitative measures of communication in science: A critical review. *History of Science*, 17:102--134, 1979.

EFSA. Opinion on a request from efsa related to the default q10 value used to describe the temperature effect on transformation rates of pesticides in soil. scientific opinion of the panel on plant protection products and their residues (ppr-panel) (question no efsa-q-2007-048). *The EFSA Journal*, 622:1--32, 2007.

J. Elliott, S. Heesterbeek, C.J. Lukensmeyer, and N. Slocum. Participatory methods toolkit. a praticioner's manual. Technical report, King Baudouin Fondation, Brussels, 2005. pp. 199.

EPA. *Guidance for Quality Assurance Project Plans for Modeling*. EPA (U.S. Environmental Protection Agency), 2002. QA/G-5M.

Epicure. Lettre à hérodote in "trois lettres", trad. o. hamelin. *Rev. de métaphysique et de Morale*, XVIII:397--440, 1910.

S. Epstein. The construction of lay expertise: Aids activism and the forging of credibility in the reform of clinical trials. *Science, Technology & Human Values*, 20:408--437, 1995.

Jan Faye. Copenhagen interpretation of quantum mechanics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition, 2008.

L. Ferry. Préface de Luc Ferry. In *E. Kant. Critique de la raison pure*. GF-Flammarion, 1987.

P. Feyerabend. *Against Method*. Verso, 3d edition, 1993. 279pp.

D.J. Fiorino. Citizen participation and environmental risk: A survey of institutional mechanisms. *Science, Technology & Human Values*, 15: 226--243, 1990.

F. Fischer. *Citizens, Experts and the Environment. The Politics of Local Knowledge.* Duke University press., Durham, 2000.

FOCUS. Leaching models and eu registration. the final report of the work of the regulatory modelling work group of focus. Technical Report DOC 4952/VI/95, European Commission, 1995.

FOCUS. Surface water models and eu registration of plant protection products. Technical Report EC 6476/VI/97, European Commission, 1997. 231 pp.

FOCUS. Focus groundwater scenarios in the eu review of active substances. report of the focus groundwater scenarios workgroup. Technical Report SANCO/321/2000 rev.2, European Commission, 2000. URL http://arno.ei.jrc.it/focus.htm. 202 p.

FOCUS. Generic guidance for focus groundwater scenarios, version: 1.1., by the groundwater scenarios workgroup of focus. Technical report, European Commission, April 2002.

FOCUS. Assessing potential for movement of active substances and their metabolites to ground water in the eu—report of the focus ground water work group. Technical Report Sanco/13144/2010 version 1, 604 pp., European Commission, 2009.

FOCUS Steering Committee. Communication from the focus steering committee.focus constitution, April 2001a. URL http://focus.jrc.ec.europa.eu/docs/FOCUS%20Constitution%202001.doc. Date of access: 2006/01/08.

FOCUS Steering Committee. Minutes of the steering committe meeting, December 2001b.

FOCUS Steering Committee. Minutes of the steering committe meeting, April 2002a.

FOCUS Steering Committee. Minutes of the steering committe meeting, December 2002b.

FOCUS Steering Committee. Minutes of the steering committe meeting, April 2003.

C. Folland and C. Anderson. Estimating changing extremes using empirical ranking methods. *JOURNAL OF CLIMATE*, 15(20): 2954--2960, OCT 2002. ISSN 0894-8755.

G. Fourez. *La construction des sciences*. Sciences, éthiques, sociétés. De Boeck Université, 4 edition, 2001.

R. Frigg and S. Hartmann. Models in science. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2009 edition, 2009. URL http://plato.stanford.edu/archives/sum2009/entries/models-science/.

S.O. Funtowics and J.R. Ravetz. *Robert Costanza (ed.), Ecological Economics*, chapter A New Scientific Methodology for Global Environnemental Issues, pages 137--152. Columbia University Press, New York, 1991.

H.-G. Gadamer. *Vérité et méthode, les grandes lignes d'une herméneutique philosophique*. Seuil, 1996-1960. Edition intégrale. 534 pp.

H. Garfinkel. *Studies in Ethnomethodology*. Prentice Hall., Englewood Cliffs, 1967.

M.J. Garland. Experts and the public: A needed partnership for genetic policy. *Public Understanding of Science*, (8):241--254, 1999.

N. Goodman. *Languages of art: an approach to a theory of symbols*. Hackett, 1976-1968. ISBN 9780915144341.

N. Goodman. *Langages de l'art. Une approche de la théorie des symboles*. Chambon, 1990-1968.

D. Gries and F.B. Schneider. *A Logical Approach to Discrete Math*. Springer-Verlag, 1993.

J.-P. Grémy. Simulation. In *Encyclopædia Universalis*. Encyclopædia Universalis s.a., 2009. URL http://www.universalis-edu.com.

J. Habermas. *La Technique et la science comme idéologie*. Gallimard., Paris, 1973.

J. Habermas. *Connaissance et intérêt*. Gallimard., Paris, 1976.

A.E. Hassan. Validation of numerical ground water models used to guide decision making. *GROUND WATER*, 42(2):277--290, MAR-APR 2004. ISSN 0017-467X.

G.W. Hazeu, M.J. Metzger, C.A. Mucher, M. Perez-Soba, C. Renetzeder, and E. Andersen. European environmental stratifications and typologies: An overview. *Agriculture, Ecosystems and Environment*, 2010. Article in press, doi:10.1016/j.agee.2010.01.009.

M. Heidegger, J. Stambaugh, and D.J. Schmidt. *Being and Time*. SUNY series in contemporary continental philosophy. State University of New York Press, 2010-1927. ISBN 9781438432762.

A. Heidemann. Still no decision on zonal authorization and mutual recognition. *SCC Newsletter*, 6(3):2, July 2006.

P. Hendley. Advances in validation of environmental exposure model predictions for regulatory purposes. In G. Ramo G. Voss, editor, *Chemistry of Crop Protection*. Wiley-VCH, 2003.

B. Hicks and F. Raveney. European pesticide directive conference. *Crop protection monthly*, July 2001. Issue 140.

R. Hiederer, R.J.A. Jones, and J. Daroussin. Soil profile analytical database for europe (spade): Reconstruction and validation of the measured data (spade/m). *Geografisk Tiddskrift - Danish Journal of Geography*, 106(1):71--85, 2006.

M. Hisschemöller and C.J.H. Midden. Improving the usability of research on the public perception of science and technology for policy-making. *Public Understanding of Science*, (8):17--33, 1999.

D.A. Howes. Decision support systems: Maximizing data value. *CSI Visions*, (Summer 2001), 2001a.

F. Howes. Eu plant protection products directive: Progress to date and projected economic impact. *CSI Visions*, (Summer 2001), 2001b.

E. Husserl. *La crise des sciences européennes et la phénoménologie transcendantale*. gallimard, 1976. Note: French Translation of Husserl's texts from 1926 to 1936.

IPCC Working Group III. *IPCC SPECIAL REPORT. EMISSIONS SCENARIOS. Summary for Policymakers. A Special Report of IPCC Working Group III*. WMO/UNEP, 2000. ISBN 92-9169-113-5.

IPCC Working Group III Technical Support Unit, editor. *Workshop on New Emission Scenarios, Laxenburg, Austria. Meeting report*, 29 June 1 July 2005. Intergovernmental Panel on Climate Change.

P. Jacob. Intentionality. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2010 edition, 2010.

D. Jacquette. Review of E.N. Zalta's "intensional logic and the metaphysics of intentionality". *Philosophy and Phenomenological Research*, 51(2):439--444, 1991.

M. Jamagne, J. Daroussin, M. Eimberck, D. King, J. J. Lambert, C. Le Bas, and L. Montanarella. Soil geographical database of eurasia and mediterranean countries at 1:1 m. wcss 2002 proceedings., 2002. URL http://search.ebscohost.com/login.aspx?direct=true&db=lah&AN=20033035790&lang=fr&site=ehost-live.

R. Jin, W. Chen, and T.W. Simpson. Comparative studies of metamodelling techniques under multiple modelling criteria. *Struct. Multidiscip. Optim.*, 23:1--13, 2001.

R.L. Jones and G. Mangels. Review of the validation of models used in federal insecticide, fungicide, and rodenticide act environmental exposure assessments. *ENVIRONMENTAL TOXICOLOGY AND CHEMISTRY*, 21(8):1535--1544, AUG 2002. ISSN 0730-7268.

L. Joumier. *Lire Husserl*. Philo. ellipses, 2007.

JRC. European soil database v 1.0, metadata june 1999: Soil profile database of europe ver 2.1.0.0, 1999. URL http://eusoils.jrc.ec.europa.eu.

Soil Geographical Database of Eurasia at scale 1:1,000,000 version 4 beta, 25/09/2001. JRC, 2001. URL http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/esdb/sgdbe/SGDBE_metadata.doc.

JRC. Spade-2 objectives. JRC web site, 2009. URL http://eusoils.jrc.ec.europa.eu/projects/spade/Objectives.html. Date of access: 2009/01/09.

JRC. European soil database v.2.0. JRC web site, 2011. URL http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/fr_intro.htm. Date of access: 2011/07/29.

S. Juan. *Méthodes de recherche en sciences sociales*. PUF, Paris, 1999.

M. Kah and C.D. Brown. Describing the behaviour of ionic compounds for pesticide registration. research project final report to defra. Technical report, Central Science Laboratory, UK, 2007.

E. Kant. *Critique of Pure Reason*. Macmillan & Co, 1929-1781. originally published in 1781; transl.: Norman Kemp Smith.

E. Kant. *Critique de la raison pure*. GF-Flammarion, 1987-1781. (First published in 1781).

D.M. Kaplan. *Encyclopedia of Philosophy*, chapter Ricœur, Paul. Macmillan Reference USA. Thomson Gale, Detroit/New York, 2nd edition, 2006. 10 Volumes.

J.P.C. Kleijen. Validation of models: statistical techniques and data availability. In *Winter Simulation Conference (WSC'99)*, volume 1, pages 647--654, 1999.

J.P.C. Kleijnen. Verification and validation of simulation models. *European Journal of Operational Research*, 82(1):145--162, April 1995.

J.P.C. Kleijnen and R.G. Sargent. A methodology for fitting and validating metamodels in simulation. *Eur. J. Oper. Res.*, 120:14--29, 2000.

H. Knoche, M. Klein, P. Lepper, M. Herrchen, C. Köhler, and U. Storm. Entwicklung von kriterien und verfahren zum vergleich und zur Übertragbarkeit regionaler umweltbedingungen innerhalb der eu-mitgliedstaaten (development of criteria and methods for comparison and applicability of regional environmental conditions within the eu member countries). Technical Report Report No: 126 05 113, Berlin Umweltbundesamt, 1998.

J.J. Kockelmans. *Boston studies in the philosophy of science*, volume 192, chapter Hermeneutic vs. empiricist philosophy of science, pages 191--215. Springer, Dordrecht, PAYS-BAS, 1997.

T.S. Kuhn. *The structure of scientific revolutions.* University of Chicago Press, 1996-1962.

M. Lamotte, C.F. Sacchi, and P. Blandin. Écologie. In *Encyclopædia Universalis*. Encyclopædia Universalis s.a., 2009. URL http://www.universalis-edu.com.

B. Latour. *Politiques de la nature. Comment faire entrer les sciences en démocratie.* Coll. armillaire. Ed. La Découverte, Paris, 1999.

B. Latour. *L'espoir de Pandore: pour une version réaliste de l'activité scientifique.* Armillaire. La Découverte, 2001.

B. Latour. *La science en action. Introduction à la sociologie des sciences.* La découverte, 2005.

B. Latour, P. Mauguin, and G. Teil. *Gestion de la recherche, nouveaux problèmes, nouveaux outils.*, chapter Pour une cartographie des innovations. Le graphe socio-technique., pages p.419--480. De Boeck, Brussels, 1991.

B. Latour, P. Mauguin, and G. Teil. Une nouvelle méthode de suivi socio-technique: le graphe socio-technique, 1992. URL http://www.ensmp.fr/~latour/articles/article/044.html. Date of access: 2003.

B. Latour, P. Mauguin, and G. Teil. A new method to trace the path of innovations. the "socio-technical graph"., 2007. URL http://www.bruno-latour.fr/articles/article/44bis-GRAPH-GB.pdf. Date of access: 2011.

A.M. Law and W.D. Kelton. *Simulation Modeling and Analysis.* McGraw-Hill, New York, 1991. 759pp.

J. Law and A. Mol, editors. *Complexities. Social Studies of Knowledge Practices*. Duke University Press, Durham, NC, 2002.

John Law. Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systemic Practice and Action Research*, 5:379--393, 1992. ISSN 1094-429X. URL http://dx.doi.org/10.1007/BF01059830.

D. Lecourt, editor. *Dictionnaire d'histoire et philosophie des sciences*. QUADRIGE/PUF, 2003.

Bertrand Leterme, Marnik Vanclooster, Mark D.A. Rounsevell, and Patrick Bogaert. Discriminating between point and non-point sources of atrazine contamination of a sandy aquifer. *Science of The Total Environment*, 362(1-3):124 -- 142, 2006. ISSN 0048-9697. doi: DOI:10.1016/j.scitotenv.2005.06.010. URL http://www.sciencedirect.com/science/article/pii/S0048969705004444.

R. Levins. The strategy of model building in population biology. *Am. Sci.*, (54):421--431, 1966.

L.Pintér, D. Swanson, and J. Chenje, editors. *IEA Training Manual: A training manual on integrated environmental assessment and reporting*. UNEP and IISD, 2008. ISBN 978-92-807-2983-2. URL http://www.unep.org/ieacp/iea/training/manual/.

M. Lynch. *Art and Artifact in Laboratory Science. A Study of Shop Work and Shop Talk in a Research Laboratory*. Routledge & Kegan Paul., London, 1985.

D. Mackenzie. *Statistics in Britain, 1895-1930. The Social Construction of Scientific Knowledge*. Edinburgh University Press., Edinburgh, 1981.

Mohammed Mahmoud, Yuqiong Liu, Holly Hartmann, Steven Stewart, Thorsten Wagener, Darius Semmens, Robert Stewart, Hoshin Gupta, Damian Dominguez, Francina Dominguez, David Hulse, Rebecca Letcher, Brenda Rashleigh, Court Smith, Roger Street, Jenifer Ticehurst, Mark Twery, Hedwig van Delden, Ruth Waldick, Denis White, and Larry Winter. A formal framework for scenario development in support of environmental decision-making. *Environmental Modelling & Software*, 24(7):798 -- 808, 2009. ISSN

1364-8152. doi: DOI:10.1016/j.envsoft.2008.11.010. URL http://www.sciencedirect.com/science/article/B6VHC-4V936S9-1/2/7c019171b37aaabd589d795686c3c460.

M. Makins, editor. *Collins English Dictionary, third edition.* Harper-Collins, 1991.

H. Maraite, W. Steurbaut, and P. Debongnie. Project cp/33 - development of awareness tools for a sustainable use of pesticides - final report. Technical report, the Belgian Science Policy (BELSPO), scientific support plan for a sustainable development policy (SPSD II), part 1: sustainable production and consumption patterns., 2004.

J.-M. Marin, K. Mengersen, and C. Robert. *Handbook of Statistics*, volume 25, chapter Bayesian modelling and inference on mixtures of distributions. Elsevier Sciences, 2005.

D. Marker. *Model Theory: An Introduction.* Graduate Texts in Mathematics. Springer, 2002. ISBN 9780387987606.

M. Markou and S. Singh. Novelty detection: a review--part 2:: neural network based approaches. *Signal Processing*, 83(12):2499 -- 2521, 2003. doi: DOI:10.1016/j.sigpro.2003.07.019. URL http://www.sciencedirect.com/science/article/pii/S0165168403002032.

C. Marris, B. Wynne, P. Simmons, and S. Weldon. Public perceptions of agricultural biotechnologies in europe - final report of the pabe research project. Technical report, European Commission, Fourth Framework Programme (Environment and Climate), Shared Cost Action, 2001.

D.G. Mayer and D.G. Butler. Statistical validation. *Ecol. Model.*, (68): 21--32, 1993.

M. Merleau-Ponty. *Oeuvres.* Quarto. Gallimard, 2010.

M. Merleau-Ponty. *Oeuvres*, chapter Phénoménologie de la perception. Quarto. Gallimard, 2010-1945.

J.S. Milton and J.C. Arnold. *Introduction to probability and statistics. Principles and applications for engineering and the computing sciences.* McGraw-Hill, 1990.

B.P. Minogue. Van fraassen's semanticism. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1:115--121, 1984. Contributed papers.

J.-M. Monnoyer. Validation. In *Dictionnaire d'histoire et philosophie des sciences*. QUADRIGE/PUF, 2003.

J. Montenot, editor. *Encyclopédie de la philosophie*. La pochothèque. Le livre de poche, 2002. 1800 pp.

M. Mormont and F. Mélard. Conférence citoyenne sur les tests génétiques "lire dans mes gènes" - rapport d'évaluation. Technical report, Fondation Roi Baudouin, Bruxelles, 2003.

N. Mouloud, B. Jaulin, M.-A. Tonnelat, J. Goguel, S. Guinand, R. Boudon, J.-F. Richard, B. Victorri, and H. Damisch. Modèle. In *Encyclopædia Universalis*. Encyclopædia Universalis s.a., 2009. URL http://www.universalis-edu.com.

C. Mucher, K. Steinnocher, F. Kressler., and C. Heunks. Land cover characterization and change detection for environmental monitoring of pan-europe. *International Journal of Remote Sensing*, 21(6): 11591181, 2000.

M.J. Mulkay. *Sociology of science. A sociological pilgrimage*. Open University Press., Milton Keynes, 1991.

M.J. Mulkay, G.N. Gilbert, and S. Woolgar. Problem areas and research networks in science. *Sociology*, 9:187--203, 1975.

B.S. Murdock, C. Wiessner, and K. Sexton. Stakeholder participation in voluntary environmental agreements: Analysis of 10 project xl case studies. *Science Technology Human Values*, (30):223--250, 2005.

E.M. Mutelesi. *Subjectivité comme auto-organisation. Une étude du constructivisme radical au départ de Husserl*. PhD thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 1998.

P. Nouvel, editor. *Enquête sur le concept de modèle*. Science, histoire et société. Presses Universitaires de France, Paris, 2002.

OECD. Savings to governments and industry resulting from the oecd environmental health and safety programme. Technical Report

ENV/EPOC/MIN(98)5, Environmental Health and Safety Division, OECD, February 1998.

N. Oreskes, K. Shrader-Frechette, and K. Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(4):641--46, 1994.

G. Pearson. The participation of scientists in public understanding of science activities: The policy and practice of the u.k. research councils. *Public Understanding of Science*, (10):121--137, 2001.

PELCOM. Pelcom project, development of a consistent methodology to derive land cover information on a european scale from remote sensing for environmental modelling, final report. Technical report, European Commission Directorate General XII-D Science, Research and Development, 2000.

A. Pickering, editor. *Science as practice and culture*, chapter Constructing quaternations: on the analysis of conceptual practice, pages 139--167. Chicago Univesity Press, 1992.

L.L. Pipino, Y.W. Lee, and R.Y. Wang. Data quality assessment. *Communications of the ACM*, 45(4ve):211--218, April 2002.

A. Pissard, V. Van Bol, J.D. Piñeros Garcet, P. Harcz, and L. Pussemier. Calcul d'indicateurs de risques liés à l'utilisation de produits phytosanitaires. etude préliminaire : détermination du niveau d'utilisation de pesticides en région wallonne. rapport final. Technical report, CERVA-CODA-VAR (Centre d'Etude et de Recherches Vétérinaires et Agrochimiques), Tervuren, Belgium, 2005. URL http://environnement.wallonie.be/eew/documents.aspx?doc=etude.

J.D. Piñeros Garcet. Scénarios, modèles mathématiques et expertise environnementale. entre construction du réel et concertation politique. Master's thesis, Université catholique de Louvain, département des sciences politiques et sociales, Louvain-la-Neuve, Belgium, 2003.

J.D. Piñeros Garcet, M. Deblonde, F. Mélard, M. Louviaux, L. Goorden, and M. Mormont. Pepam. feasibility of a participatory modelling process for pesticides risk assessment. Technical Report Research contract 0A/00/027, BELSPO, 2006a.

J.D. Piñeros Garcet, B.K. Nguyen, L. Pussemier, and E. Haubruge. Calcul d'indicateurs de risque liés à l'utilisation de produits phytosanitaires : Calcul du risque (type pribel) pour les abeilles, agrégation et cartographie pour la région wallonne. Technical report, VAR and FUSAGX, 2006b. Financed by DGRNE.

J.D. Piñeros Garcet, A. Ordoñez, J. Roosen, and M. Vanclooster. Metamodelling: Theory, concepts and application to nitrate leaching modelling. *Ecological Modelling*, 193, Issues 3-4:629--644, March 2006c.

J.D. Piñeros Garcet, V. Van Bol, and L. Pussemier. Aggregation of indicators, in: Hair, harmonised environmental indicators for pesticide risk. Technical Report Contract n°SSPE-CT-2003-501997, EU-FP6 project, 2007.

M. Power. The predictive validation of ecological and environmental models. *Ecol. Model.*, (68):33--50, 1993.

PPR Unit. Outcome of the public consultation on the project plan for the development of a new guidance document on emissions from protected crops systems. EFSA scientific Report 284,1-11, EFSA, 2009.

K. Racke. Development and registration of pesticides with reduced characteristics. In *10th IUPAC International Congress on the Chemistry of Crop Protection Products*, Basel, August 2002. Syngenta.

S. Rana. Heat mounts over eu regulatory reforms. *AGROW*, (569): 14--17, June 12th 2009. URL http://www.agrow.com.

C.A. Reeves and D.A. Bednar. Defining quality: Alternatives and implications. *The Academy of Management Review*, 19(3):419--445, 1994. Special Issue: "Total Quality".

P. Ricœur. *Du texte à l'action. Essais d'herméneutique II.* coll. Esprit. Editions du Seuil, Paris, 1986.

P. Ricœur. *From text to action. Essays in Hermeneutics, II.* Northwestern University Press, 1991-1986. ISBN 9780810109926.

A. Roger. Quelle implication des destinataires de la norme ? la voie de la corégulation. *VertigO - la revue électronique en sciences de l'environnement*, Hors série 6, 2009. URL http://vertigo.revues.org/8956.

C. Romano. *Au coeur de la raison, la phénoménologie*. folio essais. Gallimard, 2010. 1137 pp.

G. Rowe and L.J. Frewer. Public participation methods: A framework for evaluation. *Science Technology Human Values*, (25):3--29, 2000.

G. Rowe and L.J. Frewer. A typology of public engagement mechanisms. *Science Technology Human Values*, pages 251--290, 2005.

G. Rowe, R. Marsh, and L.J. Frewer. Evaluation of a deliberative conference. *Science Technology Human Values*, (29):88--121, 2004.

G. Rowe, T. Horlick-Jones, J. Walls, and N. Pidgeon. Difficulties in evaluating public engagement initiatives: reflections on an evaluation of the uk gm nation? public debate about transgenic crops. *Public Understanding of Science*, (14):331--352, 2005.

M.H. Russell, R.J. Layton, and P.M. Tillotson. The use of pesticide leaching models in a regulatory setting: An industrial perspective. *Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering*, 29(6):1105--1116, 1994. URL http://www.informaworld.com/10.1080/10934529409376097.

E.J. Rykiel. Testing ecological models: the meaning of validation. *Ecological Modelling*, 90(3):229 -- 244, 1996. ISSN 0304-3800. doi: DOI:10.1016/0304-3800(95)00152-2. URL http://www.sciencedirect.com/science/article/B6VBS-3VWK7V1-7/2/6a5c8fc2b57707e1f028e00432baa7b2.

G.B. Sahoo, C. Ray, and H.F. Wade. Pesticide prediction in ground water in North Carolina domestic wells using artificial neural networks. *Ecol. Model.*, 183:29--46, 2005.

B. Sapio. Search (scenario evaluation and analysis through repeated cross impact handling): a new method for scenario analysis with an application to the videotel service in italy. *International Journal of Forecasting*, 11(1):113 -- 131, 1995. ISSN 0169-2070. doi: DOI:10.1016/0169-2070(94)02002-7. URL http://www.sciencedirect.com/science/article/B6V92-3XWRP0G-10/2/4186173c37a184bc38a115ebe90d45de. Probability Forecasting.

J.P. Sartre. *L'Être et le néant: essai d'ontologie phénoménologique*. Gallimard, 2001-1943. ISBN 9782070293889.

Scientific Committee on Plants. Opinion of the scientific committee on plants regarding the evaluation of a document concerning the focus groundwater scenarios in the eu registration process. Technical Report SCP/GUIDE-FOC-GW/002-Final, European Commission, September 2000.

M. Singleton. Éthique des enjeux. Sillabus du cours SPED3200, UCL, 2000.

D.W. Smith. Phenomenology. In E.N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, Summer 2009.

J. Staum. Better simulation metamodeling: the why, what, and how of stochastic kriging. In M.D. Rossetti, R.R. Hill, B. Johansson, A. Dunkin, and R.G. Ingalls, editors, *Proceedings of the 2009 Winter Simulation Conference*, 2009.

I. Stengers. Le developpement durable, une nouvelle approche? *Alliage*, (40), 1999.

I. Stengers. Et si l'opinion avait parfois raison? In E. Zaccai and J.N. Missa, editors, *Le principe de précaution*. Éditions de l'Université de Bruxelles, Bruxelles, 2000.

B. Strengers, R. Leemans, B. Eickhout, B. de Vries, and L. Bouwman. The land-use projections and resulting emissions in the ipcc sres scenarios as simulated by the image 2.2 model. *GeoJournal*, 61 (4):381--393, December 2004. URL http://dx.doi.org/10.1007/s10708-004-5054-8.

H. Tijms. *Understanding probability: Chance Rules in Everyday Life*. Cambridge University Press, 2007.

A. Tiktak, D.S. de Nie, J.D. Piñeros Garcet, A. Jones, and M. Vanclooster. Assessment of the pesticide leaching risk at the pan-European level. The EuroPEARL approach. *Journal of Hydrology*, 289:222--238, 2004.

A. Tiktak, J.J.T.I. Boesten, A.M.A. van der Linden, and M. Vanclooster. Mapping ground water vulnerability to pesticide leaching with a process-based metamodel of europearl. *J. Environ. Qual.*, 35: 1213--1226, 2006.

A. Touraine. *La voix et le regard*. coll. « Sociologie permanente ». Seuil, 1978. 309 p.

D. Touret. Hans Kelsen, 2001. URL http://www.multimania.com/ideologues/Kelsen.html. Date of access: 2001/01/08.

A. Utaker. Analogies, métaphores et concepts. In P. Nouvel, editor, *Enquête sur le concept de modèle*, Science, histoire et société. Presses Universitaires de France, Paris, 2002.

N. Van Cauwenbergh. *Expert and local knowledge in decision support for natural resource management : analysis of capture and use*. PhD thesis, Université catholique de Louvain, 2008.

B. van Fraassen. *The question of hermeneutics: essays in honor of Joseph J. Kockelmans*, chapter Against transcendental empiricism, pages 309--336. Contributions to phenomenology. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.

B.C. van Fraassen. Representation: the problem for structuralism, 2004. URL http://philsci-archive.pitt.edu/2498/. Date of access: 2006/01/10.

B.C. van Fraassen. *Scientific representation: paradoxes of perspective*. Clarendon Press, Oxford, 2008.

M. Vanclooster. Apecop. Project presentation leaflet, 2002.

M. Vanclooster and J.J.T.I. Boesten. Application of nine pesticide leaching models to the vredepeel dataset . i. water, solute and heat transport. *Agricultural Water Management*, 1998.

M. Vanclooster, J. Boesten, M. Trevisan, C. Brown, E. Capri, O.M. Eklo, B. Gottesbüren, V. Gouy, and A.M.A. van der Linden. A european test of pesticide-leaching models: methodology and major recommendations. *Agricultural Water Management*, (44):1--21, 2000.

M. Vanclooster, J.D. Piñeros Garcet, J. Boesten, F. van den Berg, M. Leistra, J. Smelt, N. Jarvis, S. Roulier, P. Burauel, H. Vereecken, A. Wolters, V. Linneman, E. Fernandez, M. Trevisan, E. Capri, L. Padovani, M. Klein, A. Tiktak, A. van der Linden, D. de Nie, G. Bidoglio, F. Baouroui, A. Jones, and A. Armstrong. Effective approaches for assessing the predicted environmental concentrations of pesticides (apecop).final report. Technical Report QLK4-1999-01238, Financed by European Commission, under Key Action 4 – Environment and Health, 2003.

F. Vercruysse and W. Steurbaut. Pocer, the pesticide occupational and environmental risk indicator. *Crop Protection*, 21(issue 4):307--315, May 2002.

S. Vergucht, J.D. Piñeros-Garcet, L. Pussemier, and W. Steurbaut. Development of a pesticide risk indicator for the evaluation of the belgian reduction plan. In C. Van Peteghem, S.De Saeger, and E. Daeseleire, editors, *Towards a safer food supply in Europe*. Belgian Science Policy (BELSPO), 2007. ISBN 978-90-8756-032-4. URL http://www.belspo.be/belspo/home/publ/pub_ostc/OA/rOA22_en.pdf.

D. Vinck. *Sociologie des sciences*. Collection "U", série Sociologie. Armand Colin, 1995.

E.P. Weber. The question of accountability in historical perspective: From jackson to contemporary grassroots ecosystem management. *Administration Society*, (31):451--494, 1999.

M. Weisberg. Models for modelling. draft published in http://philpapers.org, 2008.

E.W. Weisstein. Set extension. MathWorld--A Wolfram Web Resource., 2008a. URL http://mathworld.wolfram.com/SetExtension.html. Date of access: 2008/01/20.

E.W. Weisstein. Intension. MathWorld--A Wolfram Web Resource., 2008b. URL http://mathworld.wolfram.com/Intension.html. Date of access: 2008/01/20.

C.A. Williams and J.D. Albertson. Contrasting short- and long-timescale effects of vegetation dynamics on water and carbon fluxes in water-limited ecosystems. *Water Resources Research*, 41(6), June 2005.

S. Woolgar. Interests and explanation in the social study of science. *Social Studies of Sciences*, 11:365--394, 1981.

F. Worrall. *Sustainable Ground-water Development*, chapter Direct assessment of ground-water vulnerability from borehole observations, pages 245--254. Number 193 in Special Publications. Geological Society, London, 2002.

J.H.M Wosten, Y.A. Pachepsky, and W.J. Rawls. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251(3-4):123--150, OCT 1 2001. ISSN 0022-1694.

B. Wynne. Une approche réflexive du partage entre savoir expert et savoir profane. *Les Cahiers de la sécurité intérieure*, pages 219--236, 1999.

S. Yearley. Computer models and the public's understanding of science: A case-study analysis. *Social Studies of Science*, (29):845--866, 1999.

E. Zaccai and J.N. Missa, editors. *Le principe de précaution.* Éditions de l'Université de Bruxelles, Bruxelles, 2000.

E.N. Zalta. *Abstract objects: an introduction to axiomatic metaphysics.* Springer, 1983.

E.N. Zalta. *Intensional Logic and the Metaphysics of Intentionality.* MIT Press, Cambridge,Mass, 1988.

E.N. Zalta. Mally's determinates and husserl's noemata. In A. Hieke, editor, *Ernst Mally - Versuch einer Neubewertung*, pages 9--28. Academia-Verlag, 1998.

D. Zwillinger and S. Kokoska. *CRC Standard Probability and Statistics Tables and Formulae: Tables and Formulae.* CRC Press, 2000.