

# A new framework for climate sensitivity and prediction: a modelling perspective

Francesco Ragone<sup>1,2</sup> · Valerio Lucarini<sup>2,3,4</sup> · Frank Lunkeit<sup>2</sup>

Received: 28 March 2014 / Accepted: 12 May 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** The sensitivity of climate models to increasing CO<sub>2</sub> concentration and the climate response at decadal time-scales are still major factors of uncertainty for the assessment of the long and short term effects of anthropogenic climate change. While the relative slow progress on these issues is partly due to the inherent inaccuracies of numerical climate models, this also hints at the need for stronger theoretical foundations to the problem of studying climate sensitivity and performing climate change predictions with numerical models. Here we demonstrate that it is possible to use Ruelle's response theory to predict the impact of an arbitrary CO<sub>2</sub> forcing scenario on the global surface temperature of a general circulation model. Response theory puts the concept of climate sensitivity on firm theoretical grounds, and addresses rigorously the problem of predictability at different time-scales. Conceptually, these results show that performing climate change experiments with general circulation models is a well defined problem from a physical and mathematical point of view. Practically, these results show that considering one single CO<sub>2</sub> forcing scenario is enough to construct operators able to predict the response of climatic observables to any other CO<sub>2</sub> forcing scenario, without the

need to perform additional numerical simulations. We also introduce a general relationship between climate sensitivity and climate response at different time scales, thus providing an explicit definition of the inertia of the system at different time scales. This technique allows also for studying systematically, for a large variety of forcing scenarios, the time horizon at which the climate change signal (in an ensemble sense) becomes statistically significant. While what we report here refers to the linear response, the general theory allows for treating nonlinear effects as well. These results pave the way for redesigning and interpreting climate change experiments from a radically new perspective.

**Keywords** Climate response · Climate sensitivity · Linear response theory · IPCC climate change scenarios · GCM ensemble simulations

## 1 Introduction

One of the main goals of climate science is to predict how modulations on different time scales of internal or external parameters impact the statistical properties of the climate system. Due to the observational evidence of the ongoing anthropogenic climate change, the problem of the response of the system to increasing greenhouse gases (GHGs) and in particular CO<sub>2</sub> concentration ([CO<sub>2</sub>]) is of particular social and environmental relevance (IPCC 2007a, 2013). The assessment of the future impacts of climate change under a variety of CO<sub>2</sub> forcing scenarios (IPCC 2007b, 2013) and the evaluation of the climate sensitivity to the increase of [CO<sub>2</sub>] (Knutti and Hegerl 2008; Galloway et al. 2013; Otto et al. 2013) mostly rely on the use of general circulation models (GCMs).

---

✉ Francesco Ragone  
francesco.ragone@uni-hamburg.de

<sup>1</sup> Klimacampus, Institut für Meereskunde, University of Hamburg, Bundesstrasse 53, 20146 Hamburg, Germany

<sup>2</sup> Klimacampus, Meteorologisches Institut, University of Hamburg, Hamburg, Germany

<sup>3</sup> Department of Mathematics and Statistics, University of Reading, Reading, UK

<sup>4</sup> Walker Institute for Climate System Research, University of Reading, Reading, UK

The intense efforts put on model development by the scientific community in the past decades have led to impressive improvements in the GCMs complexity and computational performances. However, even the evaluation of the most basic measure of climate sensitivity, the Equilibrium Climate Sensitivity (ECS, the change of the globally averaged surface temperature for doubling  $[\text{CO}_2]$ ), is still subject to large uncertainties (Otto et al. 2013; Sherwood et al. 2014). The relationship between the ECS and the Transient Climate Response (TCR), defined as the change in the globally averaged surface temperature at the end of 1 % per year increase of  $\text{CO}_2$  until its doubling (Otto et al. 2013), is also unclear. More in general, the response of the surface temperature (as well as of other relevant observables) at different time scales for  $\text{CO}_2$  forcing with non-trivial temporal evolution is a poorly understood problem, emphasized by the presence of substantial uncertainties in the predictive skills of the models at decadal time scales (IPCC 2013).

These issues are certainly partially due to technical difficulties related to the inaccuracies of the GCMs. However, the very efforts put in model development in the past decades and the results achieved show that these issues can not be approached only as a technological problem to be solved by building (arguably endlessly) bigger and more complex numerical models. The need for substantial advances in the scientific ideas at the basis of climate modelling strategies (and not only in the technological tools in use) is more and more clear in the climate community (Shukla et al. 2009). An illuminating perspective on mathematical frameworks suited for a theory of climate sensitivity has been recently presented (Chekroun et al. 2011a). Here we follow a complementary approach to define a robust theoretical framework for the use of GCMs in addressing the problem of climate response, sensitivity and prediction, based upon Lucarini and Sarno (2011).

In the past, several attempts have been aimed at constructing some sort of response operator for the CS able to relate a forcing (e.g. change in  $[\text{CO}_2]$ ) with a given spatial and temporal pattern to the time dependent climate response as measured by the change of chosen climatic observables (e.g. global surface temperature). In equilibrium statistical mechanics, the fluctuation–dissipation theorem (FDT) relates the response of a system to external perturbations to its internal fluctuations in the unperturbed state (Kubo 1966). The FDT has proved to be a very powerful tool for studying a variety of physical processes in condensed matter physics, acoustics, optics and many other fields of physics. FDT thus provides a tempting framework for deducing climate response from its internal variability. In fact, various authors have gone in this direction by adopting simplified versions of the FDT, with in general a satisfactory degree of success (Langen and Alexeev 2005; Gritsun and Branstator 2007; Abramov and Majda 2008).

However, it has been shown that the applicability and effectiveness of FDT-based methods to nonequilibrium systems crucially depend on the observable one wants to predict (Cooper and Haynes 2011; Cooper et al. 2013). See also the discussion in Wouters and Lucarini (2013). Here we approach the problem from a related but different point of view.

Ruelle introduced general methods for studying how nonequilibrium systems respond to external perturbations (Ruelle 1998a, b, 2009). Ruelle's response theory (RRT) clarifies the limits of FDT-based methods for nonequilibrium systems and shows that is indeed possible to compute deviations from a nonequilibrium steady state (NESS) due to weak forcing through explicit response formulas and using the statistical properties of the unperturbed state only. Recently RRT has been proposed as a rigorous framework for computing climate response and its applicability has been tested on the Lorenz 96 model (Lucarini and Sarno 2011). Going in the direction of bridging the gap between statistical mechanical theories and climate science, we show how, starting from an ensemble of runs forced by instantaneous  $\text{CO}_2$  doubling, the formalism of RRT can be used for predicting the response of the globally averaged surface temperature of a GCM to an arbitrary temporal evolution of the  $\text{CO}_2$  forcing.

## 1.1 Climate response

The standard approach to the problem of computing the response of a climate model to the forcing due to a increasing  $\text{CO}_2$  concentration is the following: take a model, run it to a stationary state, increase the  $\text{CO}_2$  concentration following one specific  $\text{CO}_2$  increase scenario, measure the increase of global surface temperature, define on it operational measures of the sensitivity of the system. Then, if you want to check what happens with another  $\text{CO}_2$  increase scenario, repeat the operation with the new scenario. What we suggest here is that repeating the simulations for every scenario could be a waste of computational resources. In this paper we show that is possible to extract from the results of one single scenario all the information needed in order to compute the effect of another scenario, without the need to perform simulations for the new scenario.

On the other hand, the standard approach to model inter-comparison of the climate response is to compare several operational definitions of climate sensitivity obtained from the simulations. The results are typically poor, in particular poorer than the agreement on the reproduction of the climatology, and the reasons of the disagreement are not well understood. The approach suggested here makes it possible to compare models in a new way, showing how the equilibrium climate sensitivity is just one point of a function that contain much more informations about the properties of the

response to components of the forcing at different time-scales. This could be useful to compare models and understand at which time-scales the response of the models differ the most, helping in the detection of differences among climate models. Also, we find a rigorous relation between transient climate sensitivity and equilibrium climate sensitivity could help in interpreting figures like Fig. 9.42 of Chapter 9 of the IPCC WG1 Fifth Assessment Report.

This new methodology has two possible benefits compared to the standard approach: making better use of the computational resources when performing climate change experiments for multiple scenarios, and allowing a new robust approach to model intercomparison. Moreover, all this is done setting the problem of studying climate response in a well defined theoretical framework, that is Ruelle response theory for nonequilibrium dynamical systems, a very successful and powerful branch of nonequilibrium statistical mechanics. While what we report here refers to linear response, the theory behind these findings allows for treating nonlinear effects as well (Ruelle 1998b; Lucarini 2009; Lucarini and Colangeli 2012).

Note that this approach targets the way we use numerical models in climate science. The evaluation of the sensitivity of the real climate depends of course on the quality and realism of the models we use. In this sense, RRT does not lead directly to an improvement of the skills of the models in representing the actual climate and climate change. What RRT inform us about is not the realism of the answer, but the consistency of the following scientific question: how does the state of a GCM change when changing one or more of its parameters? If this question proved to be ill posed (that in general with a nonequilibrium system like a complex GCM could be), the whole project of the numerical simulation of climate change would be essentially flawed, independently on the (apparent) level of realism of the results. The applicability itself of RRT demonstrates that the problem of studying climate change with GCMs is mathematically and physically well posed, and faceable with rigorous techniques widely used in other fields of physics. RRT provides a way to reformulate the question of climate change numerical simulation in a more theoretically grounded way, which hopefully would also help in understanding how to improve the quality of the answer and of the tools used in order to get to it. The issue of understanding and improving the skills of climate models per se is a challenging one, but of essentially different nature (IPCC 2007a, 2013).

The paper is organized as follows. In Sect. 2 we give a brief summary of the basic concepts of Ruelle's linear response theory, we describe the numerical experiments we have performed, and we describe how we have analyzed the data. In Sect. 3 we present the results of the numerical simulations and data analysis demonstrating the

predictive power of the theory, we show how the concept of climate sensitivity assumes a natural role in the context of linear response theory, and we present some guidelines for designing future experiments based on a scale analysis. Eventually, in Sect. 4 we discuss our results and we propose possible future lines of research.

## 2 Methods and materials

### 2.1 Linear response theory

RRT strictly applies to Axiom A dynamical systems (Ruelle 1998a, b, 2009). Roughly speaking, a dynamical system is Axiom A if (1) the dynamics is uniformly hyperbolic on the attractor and (2) the set of the periodic points is dense on the attractor. Axiom A systems possess a Sinai-Ruelle-Bowen (SRB) invariant measure, which guarantees (a) the asymptotic equivalence of time and ensemble averages of observables (that it is not, despite intuition, a general property of nonequilibrium systems) and (b) the stability of the statistical properties when a weak stochastic forcing is applied. In this sense Axiom A systems are good physical systems, for which the statistical properties of the observables are well defined. Note that property (b) requires to be far from tipping points, in the neighbourhood of which response formulas are expected to fail (indeed a tipping point could be rigorously defined by the breaking of the applicability of RRT).

Demonstrating that a dynamical system is Axiom A is a difficult task even for simple systems, and impossible for a numerical model of the complexity of a GCM. However, the use of response formulas in most cases of physical interest is justified thanks to the Chaotic Hypothesis (Gallavotti 1996), which states that chaotic systems with many degrees of freedom effectively behave as Axiom A systems in terms of properties (a) and (b) even if they do not satisfy rigorously requirements (1) and (2), at least when considering the statistical properties of coarse-grained observables (e.g. globally or regionally integrated quantities). See also Penland (2003) for a discussion in a geophysical context. In a sense the Chaotic Hypothesis is the extension to nonequilibrium systems of the ergodic hypothesis for equilibrium systems. As an example, when we compute the expectation value of an observable in a numerical model as the long-term average on a stationary state, we are in fact implicitly assuming that the system is Axiom A-like. Similarly, operations like the computation of Lyapunov exponents or the setup of data assimilation systems (Kalnay 2003) implicitly assume the same conditions. In this paper we assume that the Chaotic Hypothesis allows using RRT to study climate response with a GCM. Note that in this sense the high dimensionality and complexity of the system enforce

the applicability of theoretical instruments, rather than the opposite. If we consider low dimensional models, or dynamical systems living at the edge of chaos, the basic tenets of response theory may fail (Chekroun et al. 2011b).

Let us briefly recapitulate the basic elements of RRT. See Ruelle (2009) for a detailed presentation and Lucarini and Sarno (2011) or Lucarini et al. (2014) for a geophysically oriented discussion. Let us consider the dynamical system described by the set of ordinary differential equations corresponding to a numerical model of the climate system. We apply a weak forcing to the system so that the evolution equations can be written as  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \mathbf{X}(\mathbf{x})f(t)$ , where  $\mathbf{x}$  is the state vector of the system,  $\mathbf{F}(\mathbf{x})$  represents the unperturbed dynamics with set boundary conditions,  $\mathbf{X}(\mathbf{x})$  is a vector field defining the pattern of the forcing in the phase space, and  $f(t)$  is the time modulation of the forcing. We consider the case where the base dynamics is autonomous (no explicit time dependence in the unperturbed dynamics, e.g. absence of daily or seasonal cycle), but one could accommodate more general time-dependent flows by resorting to the concept of pullback attractor (Chekroun et al. 2011a). Ruelle showed that the expectation value of an observable  $\Phi$  in the forced system can be computed as a perturbative expansion

$$\Phi_f(t) = \Phi_0 + \sum_{n=1}^{+\infty} \Phi_f^{(n)}(t) \quad (1)$$

where  $\Phi_0$  is the expectation value in the unperturbed state and the perturbative terms  $\Phi_f^{(n)}(t)$  can be computed as convolution integrals of  $f(t)$  and suitable functions determined by the unperturbed dynamics only (Ruelle 1998a, b, 2009). The linear response is given by the first term in the series, that can be computed as

$$\Phi_f^{(1)}(t) = \int_{-\infty}^{+\infty} d\sigma_1 G_\Phi^{(1)}(\sigma_1) f(t - \sigma_1) \quad (2)$$

where  $G_\Phi^{(1)}(t)$  is the first order Green function of the observable  $\Phi$ , that can be computed with an explicit formula defined by the unperturbed flow only as described in Ruelle (1998a, b, 2009).

Response theory guarantees general properties for  $G_\Phi^{(1)}(t)$  and its Fourier transform  $\chi_\Phi^{(1)}(\omega)$ , the (linear) susceptibility of the observable  $\Phi$  (Lucarini et al. 2005; Ruelle 2009). The Green function is in general a causal function, that is  $G_\Phi^{(1)}(t) = 0$  if  $t < 0$ . By taking the Fourier transform of Eq. (1) we have

$$\tilde{\Phi}_f^{(1)}(\omega) = \chi_\Phi^{(1)}(\omega) \tilde{f}(\omega) \quad (3)$$

where  $\tilde{f}(\omega)$  is the Fourier transform of  $f(t)$ . The susceptibility gives therefore the structure of the response of the system to forcing at different frequencies (time-scales).

The real and imaginary parts of  $\chi_\Phi^{(1)}(\omega)$  represent the in- and out-of-phase response of the system respectively to a sinusoidal forcing at frequency  $\omega$ . Maxima in the absolute value of the susceptibility correspond to resonances of the system, where the response is enhanced due to positive feedbacks at the corresponding time-scales. Being  $G_\Phi^{(1)}(t)$  a causal function, its Fourier transform  $\chi_\Phi^{(1)}(\omega)$  obeys the following identity (Lucarini et al. 2005):

$$\chi_\Phi^{(1)}(\omega) = \frac{i}{\pi} P \int_{-\infty}^{+\infty} d\omega' \frac{\chi_\Phi^{(1)}(\omega')}{\omega' - \omega} \quad (4)$$

where  $P$  indicates integration in principal part and the susceptibility is related to its complex conjugate through  $\chi_\Phi^{(1)}(\omega) = [\chi_\Phi^{(1)}(-\omega)]^*$ . Equation (4) can be recast in terms of the so called Kramers–Kronig relations (KK), self-consistency relations linking the real and imaginary parts of  $\chi_\Phi^{(1)}(\omega)$  of wide use in several fields of physics, notably in optics (Lucarini et al. 2005; Lucarini 2008). RRT provides similar formulas to compute any order of nonlinearity of the response and the related nonlinear Green functions and susceptibilities (Lucarini 2009; Lucarini and Colangeli 2012). Note that limiting the attention to the linear response does not imply neglecting the nonlinear nature of the dynamics: the linear susceptibility accounts for the linear part of the response of the full nonlinear system.

Equations (2, 3) show that RRT boils down to a surprisingly simple formalism for computing the linear response of a system to an arbitrary forcing, once the Green function and the susceptibility are computed. When dealing with an high dimensional system like a GCM, the direct computation of  $G_\Phi^{(1)}(t)$  on the unperturbed dynamics following Ruelle (1998a, b, 2009) can be extremely difficult. However, the above formulas allow computing  $G_\Phi^{(1)}(t)$  and  $\chi_\Phi^{(1)}(\omega)$  from a set of suitable forcing experiments in a simple way. Consider performing an experiment with a prescribed forcing  $f(t)$ . Let us suppose that we are in the linear regime, so that the observed response of the observable  $\Phi$  is approximately  $\Phi_f^{(1)}(t)$ . Equations (2, 3) provide then the basis for deriving  $G_\Phi^{(1)}(t)$ , or equivalently  $\chi_\Phi^{(1)}(\omega)$ , from the output of the numerical experiments and the (known) functional form of the forcing (Lucarini and Sarno 2011; Lucarini et al. 2014). Note that Eq. (3) implies that for reconstructing  $\chi_\Phi^{(1)}(\omega)$  one must consider a broadband forcing, since the response  $\tilde{\Phi}_f^{(1)}(\omega)$  does not contain frequencies that are not present in the forcing  $\tilde{f}(\omega)$ .

Knowing  $G_\Phi^{(1)}(t)$ , we can then use Eq. (2) to perform projections at any lead-time  $t$  when a different temporal evolution  $g(t)$  of the forcing is used, while knowing  $\chi_\Phi^{(1)}(\omega)$  informs us about the properties of the response of the system to forcing on a wide horizon of time-scales. Demonstrations of the validity of RRT for simple systems of

geophysical interest have been proposed in the past (Abramov and Majda 2008; Lucarini and Sarno 2011); here we assess the predictive power of the theory with a real GCM, following the approach described above.

## 2.2 Experimental settings

### 2.2.1 Description of the model

The numerical model used in this study is PLASIM (Fraedrich et al. 2005), a simplified GCM developed at the University of Hamburg. Even if indeed not competitive with state-of-the-art GCMs, PLASIM produces a fairly realistic present climate and is representative of the class of complex numerical models used for operational climate prediction. The dynamical core is based on the Portable University Model of the Atmosphere PUMA (Fraedrich et al. 1998). The primitive equations are solved by a spectral transform method. The model has a full set of physical parameterizations for unresolved processes. Parameterizations include long and short-wave radiation with interactive clouds, horizontal and vertical diffusion, boundary layer fluxes of latent and sensible heat. Stratiform precipitation is generated in supersaturated states, and the Kuo scheme is used for deep moist convection, while shallow cumulus convection is parameterized by means of vertical diffusion. See the Reference Manual freely available together with the code at <http://www.mi.uni-hamburg.de/plasim> for a detailed and referenced description of the parameterizations included in the model.

The atmospheric model is coupled to a 1-layer slab model of the oceanic mixed layer with a depth of 50 m, which includes a thermodynamic sea-ice module. The climate response and sensitivities evaluated with PLASIM are therefore due only to the fast feedbacks (Lunt et al. 2010; Previdi et al. 2013), missing contributions from ocean, continental ice-sheets, vegetation and interactive carbon cycle. For the present study we set the model at T21 horizontal resolution and 10 levels vertical resolution, for a total of  $O(10^5)$  degrees of freedom, with a time-step of 45 min. In order to simplify the analysis, we have removed daily and seasonal cycles, so that the evolution equations in the reference state do not explicitly depend on time.

### 2.2.2 Description of the experiments

We consider as our observable the globally averaged surface temperature  $T_s$  and as forcing the convergence of radiative fluxes due to the increase in the logarithm of  $[\text{CO}_2]$ , since the radiative forcing scales approximately logarithmically with  $[\text{CO}_2]$  within a reasonable range of concentrations (IPCC 2007b, 2013). Therefore from now on  $\Phi_f^{(1)} = T_{sf}^{(1)}$  represents the expectation value of the increase of the

globally averaged surface temperature  $T_s$  in the linear regime, and  $f(t)$  represents the temporal evolution of the radiative forcing corresponding to a chosen  $\text{CO}_2$  forcing scenario. Linear RRT in principle applies only in the limit of infinitesimal forcing. Nonetheless, we show that its range of validity extends to rather intense finite forcing.

We consider a control run of 2400 years. The  $[\text{CO}_2]$  is set to 360 ppm, representative of the present-day value. We then perform two sets of forcing experiments, prototypical of the scenarios proposed by the IPCC protocol for estimating the climate sensitivity and the impact of  $\text{CO}_2$  forcing on the CS. In the first set of experiments we double instantaneously the  $\text{CO}_2$  concentration and we keep it fixed at 720 ppm afterwards. This corresponds to a constant radiative forcing after the  $[\text{CO}_2]$  doubling. This is the scenario used for computing the ECS. In the second set of experiments  $[\text{CO}_2]$  is increased by 1 % per year until reaching 720 ppm, and it is kept fixed afterwards. This corresponds to a radiative forcing linearly increasing for about the first 70 years until  $[\text{CO}_2]$  has doubled, and constant afterwards. This is the scenario in which one computes the TCR. In both experiments the forcing is applied homogeneously at each grid-point.

In both cases we have performed an ensemble of 200 experiments starting from different initial conditions (Lucarini 2009; Lucarini and Sarno 2011). The 200 initial conditions are taken from the control run at intervals of 10 years starting from year 200, in order to guarantee their statistical independence and performing in this way a reasonable sampling of the attractor of the unperturbed system. Each run is 200 years long. For each set of forcing experiments we consider yearly averaged values of  $T_s$  as output. We then compute the expectation values of the temperature difference  $T_{sf}^{(1)}$  and  $T_{sg\tau}^{(1)}$  by averaging over all the ensemble members and subtracting the average value computed from the control run.

The effectiveness of the sampling of the attractor, hence the number of realizations needed in order to perform our analysis, depends on the choice of the observable and of the time-scales at which we limit the estimate. In our case, the “climate” of the annual global surface temperature in PLASIM can be defined over averages of about 10 years length (we have taken as initial conditions of the realizations of the forced experiments states of the control run distant 10 years from each other, noting that they were totally uncorrelated). Therefore, having 200 realizations of 200 years each, we have an equivalent of 4000 independent “climates” to build the statistics. In order to perform this analysis on observables representative of the fine details of the dynamics, the requirements on the statistics and on the sampling of the attractor would be much more stringent. We have chosen a coarse-grained observable (both in space and time) to avoid these problems in this first attempt

of application of the technique to the output of a complex model like PLASIM.

### 2.2.3 Data analysis

In general, the susceptibility  $\chi_{T_s}^{(1)}(\omega)$  can be computed from the response of the surface temperature  $T_{sf}^{(1)}(t)$  by taking its Fourier transform  $\tilde{T}_{sf}^{(1)}(\omega)$  and using Eq. (2), being the Fourier transform of the forcing  $\tilde{f}(\omega)$  known. The Green function  $G_{T_s}^{(1)}(t)$  can then be derived by taking the inverse Fourier transform of  $\chi_{T_s}^{(1)}(\omega)$ . For special choices of the forcing,  $G_{T_s}^{(1)}(t)$  can be computed directly from  $T_{sf}^{(1)}(t)$ . If we consider the case of instantaneous  $[\text{CO}_2]$  doubling, the time evolution of the forcing is given by  $f(t) = f_{\text{CO}_2}^{2x} H(t)$ , where  $f_{\text{CO}_2}^{2x}$  is a constant and  $H(t)$  is the Heaviside function

$$\begin{aligned} \Re\{\chi_{T_s}^{(1)}(\omega)\} - \Re\{\chi_{T_s}^{(1)}(\omega_1)\} &= \frac{2(\omega^2 - \omega_1^2)}{\pi} P \int_0^{+\infty} d\omega' \frac{\omega' \Im\{\chi_{T_s}^{(1)}(\omega)\}}{(\omega'^2 - \omega^2)(\omega'^2 - \omega_1^2)} \\ \omega^{-1} \Im\{\chi_{T_s}^{(1)}(\omega)\} - \omega_1^{-1} \Im\{\chi_{T_s}^{(1)}(\omega_1)\} &= -\frac{2(\omega^2 - \omega_1^2)}{\pi} P \int_0^{+\infty} d\omega' \frac{\Re\{\chi_{T_s}^{(1)}(\omega)\}}{(\omega'^2 - \omega^2)(\omega'^2 - \omega_1^2)} \end{aligned} \quad (7)$$

( $H(t) = 0$  for  $t \leq 0$  and  $H(t) = 1$  for  $t > 0$ ), setting  $t = 0$  at the instant at which the  $[\text{CO}_2]$  doubles. Inserting this expression of  $f(t)$  in Eq. (2) and differentiating both sides of the equation gives

$$f_{\text{CO}_2}^{2x} G_{T_s}^{(1)}(t) = \frac{d}{dt} T_{sf}^{(1)}(t) \quad (5)$$

The Green function can therefore also be computed differentiating the response signal, and the susceptibility can then be computed by taking its Fourier transform.

All the computations have been performed in MATLAB® V. 7.9 environment using the standard Fast Fourier Transform (fft) algorithm. When applying fft, we implicitly force the input function to be periodic outside the time domain where it is defined. Applying this method to the time series of  $T_{sf}^{(1)}$  introduces a frequency-dependent bias in the estimate of the spectrum due to the long-term behavior of  $T_{sf}^{(1)}$ . We need to keep the information on such long-term behavior, in order to estimate properly the low-frequency variability. It is possible to cure this issue by subtracting a suitably defined function before applying fft (Nicolson 1973).

The analysis of the spectral properties of the response has been carried out starting from Eq. (4). In this case, we have used a slightly modified version (Lucarini et al. 2005) of the KK relations usually referred to as singly subtractive Kramers–Kronig relations (SSKK). Conventional KK relations stems from considering the real and imaginary parts of Eq. (4) and exploiting the symmetries of the susceptibility, obtaining

$$\begin{aligned} \Re\{\chi_{T_s}^{(1)}(\omega)\} &= \frac{2}{\pi} P \int_0^{+\infty} d\omega' \frac{\omega' \Im\{\chi_{T_s}^{(1)}(\omega)\}}{\omega'^2 - \omega^2} \\ \Im\{\chi_{T_s}^{(1)}(\omega)\} &= -\frac{2\omega}{\pi} P \int_0^{+\infty} d\omega' \frac{\Re\{\chi_{T_s}^{(1)}(\omega)\}}{\omega'^2 - \omega^2} \end{aligned} \quad (6)$$

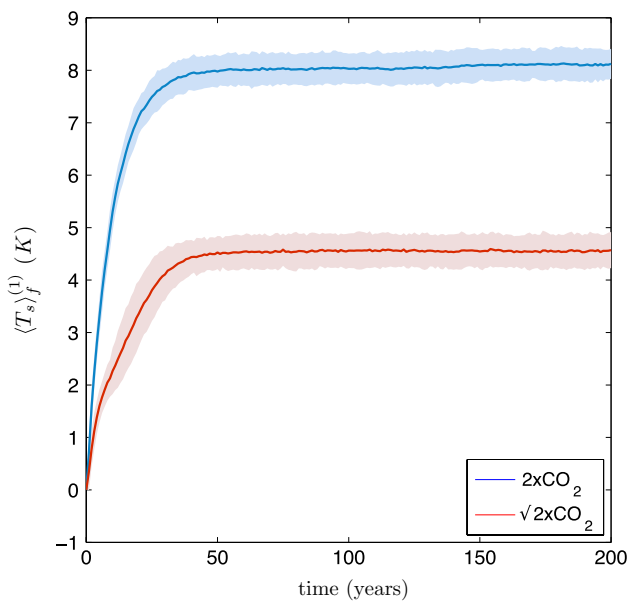
KK relations provide a self-consistency test for measured or observed data in many fields of physics, notably in optics (Lucarini et al. 2005). The accuracy of the KK inversion depends on the quality of the observed data. The SSKK method consists in referring to an anchor point  $\omega_1$  where it is supposed that the direct estimate of  $\chi_{T_s}^{(1)}(\omega)$  from the data is reliable (Lucarini et al. 2005). The SSKK relations then read

With a careful choice of the anchor point the SSKK relations improve the accuracy of the KK analysis (Lucarini et al. 2005). In our case we have taken  $\omega_1 = 2\pi\xi_1$  with  $\xi_1 = 0.1 \text{ years}^{-1}$ , obtaining a rather good quality of the inversion. We remark anyway that the results are rather robust with respect to different choices of the anchor point  $\omega_1$ .

## 3 Results

### 3.1 Climate response and prediction

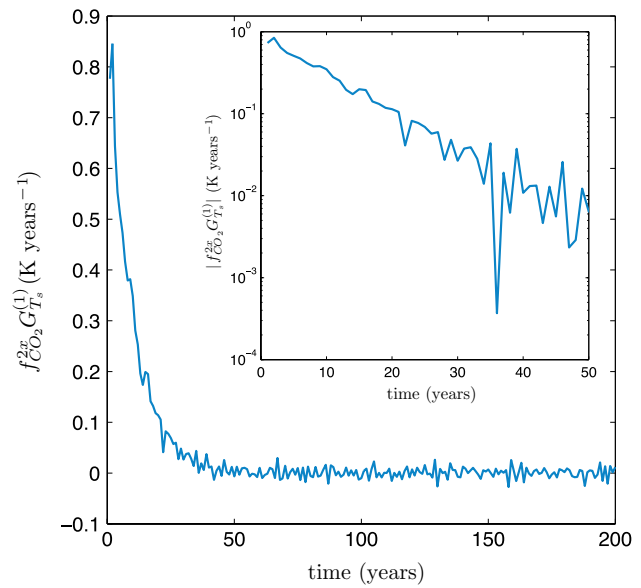
Figure 1 shows the time series of the ensemble average of the increase of the globally averaged surface temperature for the instantaneous doubling scenario (blue). The shaded area represents the 95 % (two standard deviations) of the ensemble variability. The long-term increase of the surface temperature for the doubling scenario (the equilibrium climate sensitivity) is rather high if compared with what is typically obtained with standard IPCC models, being 8.1 K against typical estimates between 1.5 and 4.5 K (IPCC 2007b, 2013). This is due to the fact we have chosen a simplified setup without the daily and in particular the seasonal cycle, which greatly enhances the system response to the radiative forcing. While this is mathematically more convenient at the price of decreasing the realism of our experiments, it also pushes the RRT more to its limits, thus providing a more stringent test for our approach. In order to test the limits of validity



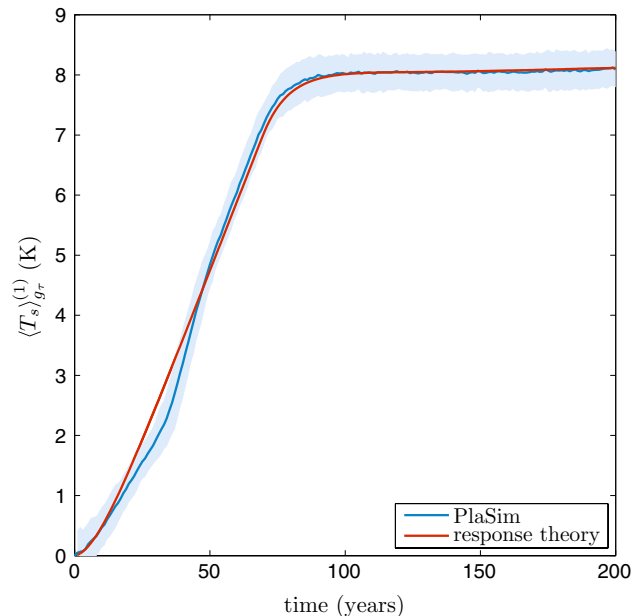
**Fig. 1** Ensemble average of increase of global surface temperature after instantaneous increase of the  $[CO_2]$  by a factor 2 (blue) and square root of 2 (red). The upper and lower limit of the bands are computed as two standard deviations of the ensemble distribution

of the linearity of the response, we have performed additional experiments (red) increasing instantaneously the  $[CO_2]$  by a factor  $\sqrt{2}$  (about 510 ppm), checking that the observable indeed scales reasonably linearly with the logarithm of the  $[CO_2]$ , the actual behavior being only slightly sublinear, with discrepancies of the order of 10 %.

In Fig. 2 we show  $G_{T_s}^{(1)}(t)$ , derived from the doubling experiment output. In the inset, the Green function is shown in logarithmic scale for the first 50 years. The Green function  $G_{T_s}^{(1)}(t)$  is consistent with a relaxation process with a characteristic time scale of about 10 years (roughly correspondent to the relaxation time-scale of the mixed layer ocean), even if relevant differences emerge. Hasselmann et al. (1993) introduced the concept of Green function of the global surface temperature of a GCM when addressing empirically the problem of the so called cold start of climate simulations. This is the warming negative bias obtained when, due to limitations in computational resources, a climate change experiment was not started from the preindustrial stationary state but from a stationary state corresponding to an higher GHGs concentration. In that case, the response of  $T_s$  was fitted with a prescribed functional form and then  $G_{T_s}^{(1)}(t)$  computed analytically with Eq. (5). Here, instead, we are computing  $G_{T_s}^{(1)}(t)$  numerically directly from the output of the GCM without loss of information.



**Fig. 2** Green function of global surface temperature, computed from the  $[CO_2]$  doubling experiment. In the inset, the absolute value is displayed in logarithmic scale for the first 50 years, showing good agreement with an exponential. The time scale of the exponential decay is about 25 years and agrees with the relaxation time scale of the mixed layer ocean

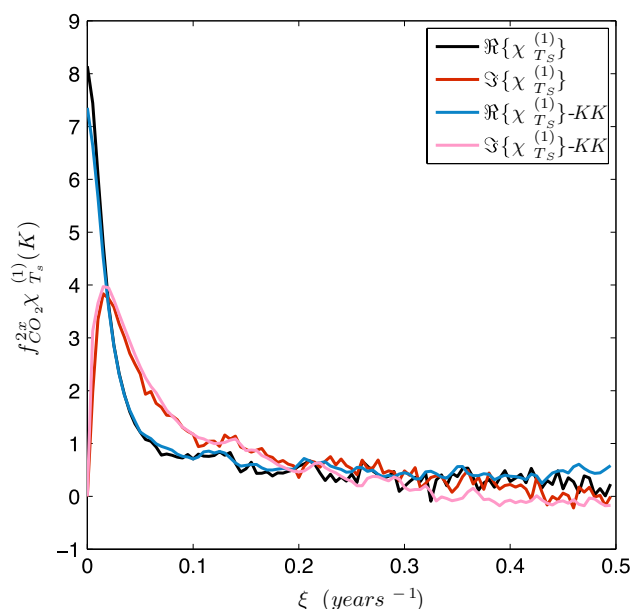


**Fig. 3** Comparison between GCM simulation (blue) and response theory prediction (red) for 1 % per year increase of the  $CO_2$  concentration. The upper and lower limit of the bands are computed as two standard deviations of the ensemble distribution. The agreement between the GCM run and the prediction via linear response theory is remarkable, with only a slight discrepancy during the transient, most probably connected to the activation of the ice-albedo feedback

In order to test the predictive power of the theory, we consider another scenario where  $[\text{CO}_2]$  is increased by 1 % per year starting from the present-day value of 360 ppm until the value of 720 ppm is reached, and kept fixed afterwards. Since to a very good approximation the radiative absorption is proportional to the logarithm of  $[\text{CO}_2]$ , the corresponding radiative forcing is, to a good approximation, a ramp function  $g_\tau(t) = f_{\text{CO}_2}^{2x} t/\tau$  for  $0 \leq t \leq \tau$  and  $g_\tau(t) = f_{\text{CO}_2}^{2x}$  for  $t > \tau$  with  $\tau \approx 70$  years. In Fig. 3 we compare the ensemble average of the simulations and the prediction for  $T_{sg\tau}^{(1)}$  obtained using the estimate of the Green function shown in Fig. 2. The agreement is excellent both on the short and long term. Discrepancies of less than 10 % are present during the transient in the window between 25 and 100 years, because of the strong nonlinearities due to the activation of the ice-albedo feedback. The degree of precision of the prediction obtained with the linear RRT is remarkable, considering the complexity of the climate model in use and the presence of strong nonlinearities in the underlying equations.

### 3.2 Climate sensitivity in response theory framework

Figure 4 shows the real  $\Re\{\chi_{T_s}^{(1)}(\omega)\}$  and imaginary  $\Im\{\chi_{T_s}^{(1)}(\omega)\}$  part of the susceptibility, computed as the Fourier transform of  $G_{T_s}^{(1)}(t)$ . The function is plotted against the frequency  $\xi = \omega/2\pi$  measured in  $\text{years}^{-1}$ . The susceptibility to  $\text{CO}_2$  forcing at a frequency  $\xi'$  gives a precise quantification of the climate response at the corresponding time



**Fig. 4** Real (black) and imaginary (red) parts of the susceptibility computed as the Fourier transform of the Green function of Fig. 1. Real (blue) and imaginary (pink) parts of the susceptibility obtained via singly subtractive KK relations

scale  $t' = 1/\xi'$ . The estimate of the susceptibility becomes rather noisy for frequencies higher than  $0.25 \text{ years}^{-1}$  (time-scales  $\leq 4$  years), where we are not able to compute reliable climate projections. The overall quality of the experimentally obtained  $\chi_{T_s}^{(1)}(\omega)$  is confirmed by the good agreement between the measured and KK-reconstructed real and imaginary parts. We have truncated the integral in Eq. (4) at the highest measured frequency  $\xi_h = 0.5/\Delta t$ , where  $\Delta t = 1$  year, and used the SSKK technique for the reconstruction (Lucarini et al. 2005). One must emphasize that all the bumps around  $\xi \approx 0.15 \text{ years}^{-1}$  and  $\xi \approx 0.20 \text{ years}^{-1}$  found in the  $\chi_{T_s}^{(1)}(\omega)$  are not due (only) to the presence of the noise, but correspond indeed to physical processes of the system (multi-annual variability), as they are consistently captured by the KK relations. These spectral features are not consistent with a simple relaxation model of the climate response.

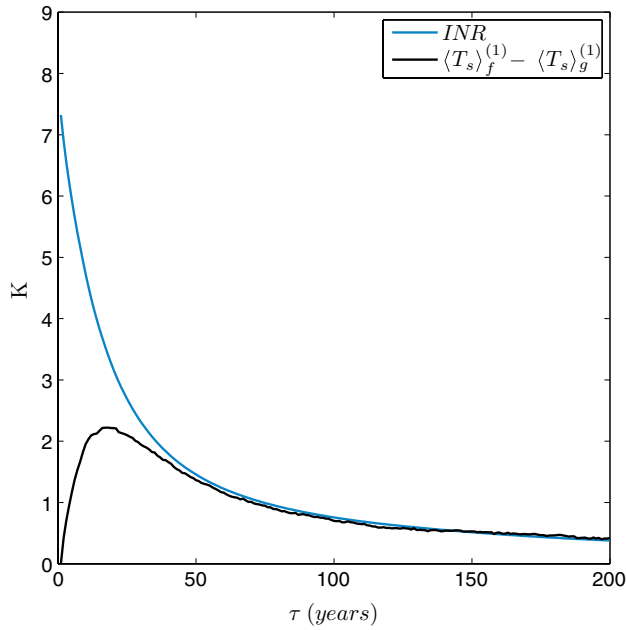
While the importance of  $G_{T_s}^{(1)}(t)$  is due to its predictive power,  $\chi_{T_s}^{(1)}(\omega)$  is a powerful tool for defining accurately climate sensitivity at different time scales. The equilibrium climate sensitivity is defined as  $ECS = \lim_{t \rightarrow +\infty} T_s^{(1)}(t)$  after an instantaneous doubling of  $[\text{CO}_2]$ . Using Eq. (3) and the definition of  $\chi_{T_s}^{(1)}(\omega)$  one can show that, in linear approximation,

$$ECS = f_{\text{CO}_2}^{2x} \chi_{T_s}^{(1)}(0) \quad (8)$$

Therefore  $ECS$  is proportional to the zero frequency value of the susceptibility. From Fig. 4 (and Fig. 1) we see that  $f_{\text{CO}_2}^{2x} \chi_{T_s}^{(1)}(0) \approx 8.1 \text{ K}$ . Therefore, when computing the  $ECS$  as long-term average of the  $T_s$  increase from a GCM run, we actually compute  $\chi_{T_s}^{(1)}(\omega)$  for a specific frequency  $\omega = 0$ . By using Eqs. (3, 4) and the expression of  $\tilde{f}(\omega)$  we obtain:

$$ECS = \frac{2}{\pi} P \int_0^{+\infty} d\omega \frac{\text{Im}[f_{\text{CO}_2}^{2x} \chi_{T_s}^{(1)}(\omega)]}{\omega} = \frac{2}{\pi} P \int_0^{+\infty} d\omega \text{Re}[\tilde{T}_{sf}^{(1)}(\omega)] \quad (9)$$

The RRT formalism allows linking the  $ECS$  to the imaginary part of the susceptibility at all frequencies or, equivalently, to the real part of the response at all frequencies for the imposed step function forcing  $f(t)$ . Equation (9) clarifies that the values of the response at all frequencies are relevant for determining the long-term response. One may compare the integrand of Eq. (9) obtained for two different models in order to test their agreement. In so doing, one would find out which time-scales (and therefore which physical processes) are mostly responsible for possible discrepancies in their  $ECS$ s. Alternatively, one may find that two models with similar  $ECS$ s differ substantially regarding their response at different time scales, detecting possible misleading compensating effects.



**Fig. 5** Inertia of the system at time scale  $\tau$  (blue) and difference at lead time  $\tau$  between the response at the system to a step-like and ramp (of time scale  $\tau$ ) forcings (black)

As mentioned above, the Transient Climate Response (TCR) is defined as the  $T_s$  increase at the moment (after about 70 years)  $[\text{CO}_2]$  has doubled following a 1 % per year increase (Otto et al. 2013). In our case  $TCR(\tau) = T_{sg\tau}^{(1)}(\tau)$  when the forcing is given by the ramp function  $g_\tau(t)$ . The forcing is the same used for the scenario of Fig. 3, and we obtain  $TCR(\tau) = 7.2$  K. Considering the Fourier representation of  $T_{sg\tau}^{(1)}(t)$  evaluated for  $t = \tau$  and using Eq. (3)

$$\begin{aligned} T_{sg\tau}^{(1)}(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \tilde{T}_{sg\tau}^{(1)}(\omega) e^{-i\omega\tau} \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \chi_{T_s}^{(1)}(\omega) \tilde{g}_\tau(\omega) e^{-i\omega\tau} \end{aligned} \quad (10)$$

The forcing is the sum of the ramp function up to time  $\tau$  and the Heaviside function translated by  $\tau$ , therefore its Fourier transform is given by  $\tilde{g}_\tau(\omega) = f_{CO_2}^{2x} P (\pi \delta(\omega) e^{i\omega\tau} + i \text{sinc}(\omega\tau/2) e^{i\omega\tau/2} / \omega)$ , where  $\text{sinc}(x) = \sin(x)/x$ . Therefore

$$\begin{aligned} T_{sg\tau}^{(1)}(\tau) &= \frac{1}{2} f_{CO_2}^{2x} \chi_{\Phi}^{(1)}(0) \\ &\quad - P \int_{-\infty}^{+\infty} d\omega f_{CO_2}^{2x} \chi_{\Phi}^{(1)}(\omega) \frac{\text{sinc}(\omega\tau/2) e^{-i\omega\tau/2}}{2\pi i\omega} \end{aligned} \quad (11)$$

Using the Cauchy formula

$$\frac{1}{2} f_{CO_2}^{2x} \chi_{T_s}^{(1)}(0) = P \int_{-\infty}^{+\infty} d\omega \frac{f_{CO_2}^{2x} \chi_{T_s}^{(1)}(\omega)}{2\pi i\omega} \quad (12)$$

we derive

$$\begin{aligned} T_{sg\tau}^{(1)}(\tau) &= f_{CO_2}^{2x} \chi_{T_s}^{(1)}(0) \\ &\quad - P \int_{-\infty}^{+\infty} d\omega f_{CO_2}^{2x} \chi_{T_s}^{(1)}(\omega) \frac{1 + \text{sinc}(\omega\tau/2) e^{-i\omega\tau/2}}{2\pi i\omega} \end{aligned} \quad (13)$$

Therefore we can write

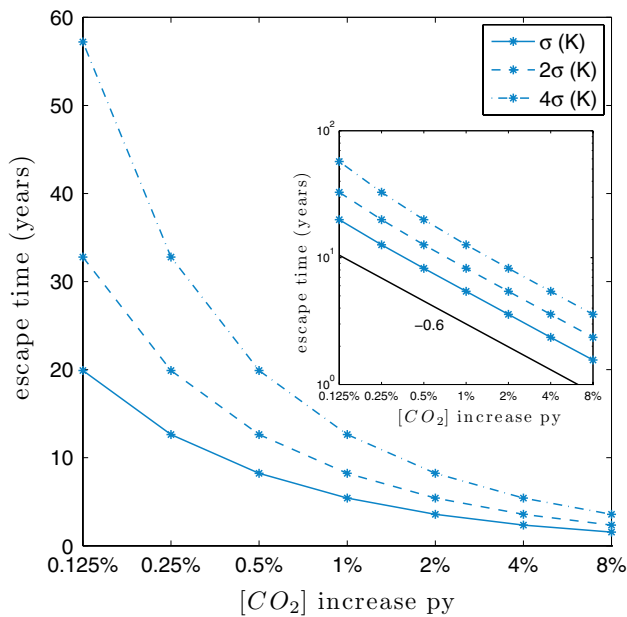
$$\begin{aligned} ECS - TCR(\tau) &= INR(\tau) \\ &= f_{CO_2}^{2x} P \int_{-\infty}^{+\infty} d\omega \chi_{T_s}^{(1)}(\omega) \frac{1 + \text{sinc}(\omega\tau/2) e^{-i\omega\tau/2}}{2\pi i\omega} \end{aligned} \quad (14)$$

The difference between  $ECS$  and  $TCR$  is given by a weighted integral of the susceptibility, accounting for the contribution of processes and feedbacks occurring at different time scales. The integral in Eq. (14) by KK relations gives  $\chi_{T_s}^{(1)}(0)$  in the limit  $\tau \rightarrow 0$ , decreases monotonically with  $\tau$ , and vanishes in the limit  $\tau \rightarrow \infty$ .  $INR(\tau)$  provides a measure of the inertia of the system at the timescale  $\tau$ , due to the overall contribution of the internal physical processes and characteristic time-scales of the relevant climatic subsystems (Saltzman 2001; Winton et al. 2010). By changing  $\tau$ ,  $INR(\tau)$  allows one to deal with different rates of increase of  $[\text{CO}_2]$ , as shown in Fig. 5. In Fig. 5 we also plot  $T_{sf}^{(1)}(\tau) - T_{sg\tau}^{(1)}(\tau)$ , which instead measures the difference in the transient response at time  $\tau$  between the case where the forcing is modulated by  $f(t)$  and by  $g_\tau(t)$ , respectively. This quantity approximately coincides with  $INR(\tau)$  for  $\tau > 50$  years but has a completely different behavior for small values of  $\tau$ .

### 3.3 Horizons of predictability

We can quantify the limits to the predictive skills of the theory with a scale analysis. We focus on the high-frequency range. Given an ensemble of  $N$  realizations, the error in the estimate of  $\tilde{\Phi}_f^{(1)}(\omega)$  with respect to the true expectation value can be represented as a random signal  $\sigma(\omega)$  such that  $|\sigma(\omega)| \approx \alpha N^{-1/2}$ , where  $\alpha$  is a suitable constant. If the forcing spectrum decays for high frequencies as  $|f(\omega)| \sim \beta \omega^{-\nu}$ , from Eq. (3) we derive that the error on the estimate of the susceptibility is  $|\delta[\chi_{\Phi}^{(1)}(\omega)]| \approx \alpha / \beta N^{-1/2} \omega^\nu$ . Assuming that the true susceptibility asymptotically decays as  $|\chi_{\Phi}^{(1)}(\omega)| \sim \gamma \omega^{-\kappa}$ , the signal to noise ratio approaches unity at

$$\omega_c = \left( \frac{\beta \gamma N^{1/2}}{\alpha} \right)^{1/(\nu+\kappa)} \quad (15)$$



**Fig. 6** Escape time of mean climate change signal from natural variability as function of the rate of increase of  $[CO_2]$ . The natural variability is computed as one, two and four times the standard deviation ( $\sigma = 0.17$  K) of the global surface temperature in the control run (blue solid, dashed and dot-dashed lines respectively). In the inset, the escape time is displayed in logarithmic scale. The linear relation indicates a power law relation between escape time and rate of increase of  $[CO_2]$ , with exponent  $n \approx -0.6$  for all values of  $\sigma$  (black)

For  $\omega \gtrsim \omega_c$  the estimate of the susceptibility will be seriously deteriorated by the presence of noise so that we will have no skills in prediction at time scales smaller than  $\tau_c = 2\pi/\omega_c$ . From our data  $N = 200$ , we estimate  $\alpha \approx 0.2$ ,  $\gamma \approx 0.63$ ,  $\kappa = 1$ , while we have by construction  $\beta = 1$  and  $\nu = 1$ . Using Eq. (15), we obtain an estimate of  $\tau_c \approx 1$  year, which fits with the qualitative information provided by Fig. 3. The design of the experiment and the processing of the output involve 3 parameters: the temporal resolution of the response signal  $\Delta t$  (typically a coarse graining of the raw output of the numerical model), the length of the simulations  $T$ , and the size of the ensemble  $N$ . They define respectively the high frequency cut-off  $\omega_h = \pi/\Delta t$ , the low frequency cutoff  $\omega_l = 2\pi/T$  (that is also the spectral resolution  $\Delta\omega = \omega_l$ ) and the critical value for predictability  $\omega_c$ . Given the time scales of interest, the appropriate temporal resolution, integration length and ensemble size to have good predictive skills can thus be determined. The quality of prediction differs for different observables and forcing, with *red* observables featuring a slowly decaying susceptibility and forced by *red* modulations being better candidates for a successful prediction.

### 3.4 Escape time of climate change signal from natural variability range

Another interesting application the possibility to estimate the time horizon on which the mean climate change signal is distinguishable from the natural variability of the climate system for different rates of change of the forcing. Although there is nowadays a general consensus on the fact that the observed warming can not be explained by the natural fluctuations of the system (IPCC 2013), a precise statistical assessment of the anthropogenic nature of the global warming signal requires rejecting the hypothesis that the observed signal can be due to the internal variability of the system (e.g. Lovejoy 2014). Given the Green function of the system, we can compute the expected mean climate change signal for forcing corresponding to different rates of change of the  $CO_2$  concentration, and check after how many years the mean signal is larger than a chosen number of standard deviations of the observable in the unperturbed system. In Fig. 6 we show for our data this “escape time”  $\tau_e$  as a function of the rate of change of the  $CO_2$  concentration  $r$ . The x-axis is in logarithm scale. The different lines correspond to the escape time from one, two and four standard deviations of the observable in the control run ( $\sigma = 0.17$  K). In the inset we show the escape time in logarithm scale. The behavior is clearly linear, so that we can infer the relation  $\tau_e \propto r^{-n}$ , where  $n \approx 0.6$  for all the values of  $\sigma$  (the linear fit is shown in black in the inset). We can see that the time required in order to be able to reject with high confidence the hypothesis that the observed signal is due to the internal variability ranges between 10 up to more than 30 years, even with a model like PLASIM with a relatively small internal variability and a relatively large response with respect to state of the art GCMs.

## 4 Summary and discussion

The applicability of RRT has several important implications regarding climate change science. At a fundamental level, it demonstrates that the problem of climate change is indeed well-defined from a mathematical and a physical point of view, as we can study it in a statistical mechanical framework. Applying the FDT to a nonequilibrium system like the climate can introduce mathematically uncontrollable errors in the estimates of the response or, anyway, requires a very accurate representation of the attractor of the system. While there might be ways to circumvent this problem (Colangeli and Lucarini 2014), we expect to be able to reconstruct the climate response from its natural variability only in special cases. The approach proposed here bypasses some of these mathematical issues by exploiting formal properties of the

response and allows for constructing rigorous definitions of climate sensitivity at different time scales through the susceptibility function. We have provided a framework for relating the difference between transient and equilibrium climate sensitivity to the inertia of the CS, and have shown how these properties depend of the response of the system on all time scales. This partially addresses some issues debated in the literature regarding the specific relevance of the *ECS* (Allen and Frame 2007) and provide viable ways to compare GCMs. Deviations from a simple relaxation behaviour have been detected, and point to the complexity of the climate response at multi-annual time-scales. Considering the corresponding spectral features is necessary for having a consistent and integrated picture of the overall climate response at all frequencies.

Our approach helps clarifying the limits of simple linear feedbacks approximations. As an example, the linear inverse modelling technique (LIM, Sardeshmukh et al. 2000; Penland 1996) assumes the dynamics to be described by an Ornstein–Uhlenbeck (OU) process, that is a stochastic differential equation with a linear deterministic term and a state-independent Gaussian noise term. For this class of systems the fluctuation–dissipation theorem holds, and because of the specific form of the dynamics it boils down to the fact that the Green function of the system can be computed from its auto-correlation function in the unperturbed state. Our approach, based on Ruelle response theory, can be applied to a more general class of systems for which the underlying dynamics can not be reduced to a simple OU form, or for which the validity of the fluctuation–dissipation theorem itself is unclear. If one applies our analysis to a OU system, one would recover the same Green function as the one obtained with a LIM analysis. However, if the underlying dynamics is very different from a OU process, obtaining the Green function from the statistics of the unperturbed system can be extremely cumbersome (even when the form of the dynamics is known), or just impossible in case of lack of validity of the fluctuation–dissipation theorem. The approach proposed in this paper bypasses these problems, and recovers the results of a LIM analysis if applied to a system compatible with a OU dynamics. Note that with our analysis, however, we can recover the Green function and the deterministic part of the dynamics only. The covariance structure of the noise (or, more in general, the natural variability of the dynamics) can not be computed in general with the approach proposed in this paper, since Eq. (2) refers only to the expectation value of the response.

On the practical side, our results provide a way to perform climate prediction (in an ensemble sense) at all lead-times. We have also shown how to estimate the predictability horizon and assess how it scales with different sizes of the ensemble of simulations. Also, we have shown how

this method allows estimating the escape time of the mean climate change signal from the natural variability of the model. Inaccuracies in representing specific spectral features have serious impacts on our ability to predict climate response on the corresponding time scales, and our findings could help understanding why, e.g., climate response at decadal time scales may be hard to capture. We have highlighted that our ability to predict the response may vary substantially when different observables are considered. It appears that considering only one single CO<sub>2</sub> forcing scenario allows for reconstructing accurately the response of the system to other temporal patterns of changes in [CO<sub>2</sub>]. Exploiting the linearity of the response it is possible to compose the effect of multiple forcings (like changes in the solar constant or in other GHGs) by simply adding the linear corresponding Green functions and susceptibilities; the extension to nonlinear cases is more cumbersome but theoretically doable (Lucarini 2009; Lucarini and Colangeli 2012), and basically points to a generalisation of the factor separation technique (Stein and Alpert 1993).

We have limited our analysis to one single observable of primary climatic interest. The analysis of other observables could shed light on the mechanisms determining the response of the CS to CO<sub>2</sub> forcing. As an example, the analysis of the response of large-scale meridional gradients of temperature at surface and in the middle troposphere could provide information on changes in the mid-latitude circulation. We expect the analysis to work better for globally or regionally integrated, rather than for local, quantities, in terms of signal to noise ratio in the estimate of the susceptibility at high frequencies. In general different observables will show different ranges of linearity. The response of water vapor related observables, in particular, is expected not to scale linearly in the same range as the surface temperature, because of the Clausius–Clapeyron equation (Held and Soden 2006). In this case, if one wants to study large climate shifts the role of the nonlinear terms will likely be relevant. The existence of approximate functional relationship between the susceptibilities of different observables (Lucarini 2009) could set a way for rigorously defining the so-called emergent constraints (Bracegirdle and Stephenson 2013; Cox et al. 2013).

The analysis of fully coupled atmosphere–ocean GCMs is crucial in order to understand the important role of ocean and interactive sea-ice on climate response. Obviously, considering models with interactive ocean introduces longer time scales and makes the analysis technically more difficult. For a prediction in an ensemble sense, we need to be able to account for all time scales. This implies the use of very long integrations and a large number of ensemble members. A tentative intercomparison analysis on several GCMs could be performed using CMIP5 data, even if this goal seems to be out of reach given the limited length of the

hosted runs. This hints at the need of including in model data repositories longer runs to properly assess the effect of the climate change scenarios on the state of the system (Lucarini and Ragone 2011). As a first (still challenging but currently more doable) step in the direction of testing RRT in state of the art climate models, we are currently working on the large ensemble of simulations made available in the climateprediction.net project.

Regarding the important point of the comparison with observations, the operation is at the moment beyond the scope of the present paper and our technological capabilities. However, in general, considering reconstructions of historical climate and forcing (as available in IPCC AR5) and a realistic climate model, with this technique it is possible to derive pieces of a “realistic” comprehensive Green function of the climate system, and test it against historical climate records. As an example, we could use this technique to study the impact of volcanic eruptions (close to impulse forcing, and therefore of easy analysis) using a specific event and testing it for predicting the impact of further eruptions. As said, this is still beyond our technological capabilities and the scope of this exploratory, demonstrative paper, even if we are currently working in this direction.

RRT provides a well defined theoretical framework and tools that allows to diagnose rigorously discrepancies in the properties of the frequency dependent response of different models and to guide the design of the climate change experiments. On one side, this could support strategies for future GCM development aimed at improving the persisting deficiencies in the model performances highlighted in the last IPCC report (IPCC 2013) by exploiting rigorous analysis and intercomparison procedures. On the other side, the vast number of diverse scenarios considered nowadays in the IPCC protocol could be substituted by a more focused effort on few selected forcing experiments making use of ensemble methods. Or, conversely, given the present palette of forcing scenarios, one could fill in the gaps and create projections for combinations of scenarios without resorting to additional simulations. This is promising for applications, considering that simulations for climate change assessment are extremely expensive and, at the same time, it is necessary to inform policymakers on a wide range of climate change scenarios.

**Acknowledgments** The authors wish to thank C. Franzke and G. Gallavotti for commenting on an earlier version of the manuscript. F.R. wish to thank T. Bodai and S. Schubert for useful discussions. F.R. and V.L. acknowledge funding from the Cluster of Excellence for Integrated Climate Science (CLISAP) and from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement No. 257106. The authors acknowledge the Newton Institute for Mathematical Sciences (Cambridge, UK), hosting the 2013 programme “Mathematics for the Fluid Earth” during which part of this work was discussed.

## References

- Abramov RV, Majda A (2008) New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems. *J Nonlinear Sci* 18:303–341
- Allen MR, Frame DJ (2007) Call off the quest. *Science* 318:582–583
- Bracegirdle TJ, Stephenson BD (2013) On the robustness of emergent constraints used in multimodel climate change projections of arctic warming. *J Clim* 26:669–678
- Chekroun MD, Simonnet E, Ghil M (2011a) Stochastic climate dynamics: random attractors and time-dependent invariant measures. *Phys D* 240:1685–1700
- Chekroun MD, Kondrashov D, Ghil M (2011b) Predicting stochastic systems by noise sampling, and application to the El Niño–Southern Oscillation. *Proc Natl Acad Sci USA* 108:11766–11771
- Colangeli M, Lucarini V (2014) Elements of a unified framework for response formulae. *J Stat Mech* P01002. doi:10.1088/1742-5468/2014/01/P01002
- Cooper FC, Haynes PH (2011) Climate sensitivity via a nonparametric fluctuation-dissipation theorem. *J Atmos Sci* 68:937–953
- Cooper FC, Esler JG, Haynes PH (2013) Estimation of the local response to a forcing in a high dimensional system using the fluctuation-dissipation theorem. *Nonlinear Process Geophys* 20:239–248
- Cox PM, Pearson D, Booth BB, Friedlingstein P, Huntingford C, Jones CD, Luke CM (2013) Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature* 494:341–344
- Fraedrich K, Kirk E, Lunkeit FP (1998) Portable university model of the atmosphere. Technical Report 16, Deutsches Klimarechenzentrum
- Fraedrich K, Jansen H, Luksch U, Lunkeit F (2005) The planet simulator: towards a user friendly model. *Meteorol Z* 14:299–304
- Gallavotti G (1996) Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem. *J Stat Phys* 84:899–926
- Galloway J, Heimann J, Le Quéré C, Levitus S, Ramaswamy V (2013) Climate sensitivity in the Anthropocene. *Q J R Meteorol Soc* 139:1121–1131
- Gritsun A, Branstator G (2007) Climate response using a three-dimensional operator based on the Fluctuation Dissipation theorem. *J Atmos Sci* 64:2558–2575
- Hasselmann K, Sausen R, Maier-Reimer E, Reinhard V (1993) On the cold start problem in transient simulations with coupled atmosphere ocean models. *Clim Dyn* 9:53–61
- Held IM, Soden BJ (2006) Robust responses of the hydrological cycle to global warming. *J Clim* 19:5686–5699
- IPCC (2007a) In: Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (eds) *Climate Change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK
- IPCC (2007b) In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- IPCC (2013) In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate Change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change stocker*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Kalnay E (2003) *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, Cambridge

- Knutti R, Hegerl GC (2008) The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nat Geosci* 1:735–743
- Kubo R (1966) The fluctuation-dissipation theorem. *Rep Prog Phys* 29:255–284
- Langen PL, Alexeev VA (2005) Estimating  $2 \times \text{CO}_2$  warming in an aquaplanet GCM using the fluctuation-dissipation theorem. *Geophys Res Lett* 32(23):L23708
- Lorenz E (1979) Forced and free variations of weather and climate. *J Atmos Sci* 36:1367–1376
- Lovejoy S (2014) Scaling fluctuation analysis and statistical hypothesis of anthropogenic warming. *Clim Dyn* 42:2339–2351
- Lucarini V, Saarinen JJ, Peiponen K-E, Vartiainen EM (2005) Kramers-kronig relations in optical materials research. Springer, Heidelberg
- Lucarini V (2008) Response theory for equilibrium and non-equilibrium statistical mechanics: causality and generalized Kramers–Kronig relations. *J Stat Phys* 131:543–558
- Lucarini V (2009) Evidence of dispersion relations for the nonlinear response of Lorenz 63 System. *J Stat Phys* 134:381–400
- Lucarini V, Sarno S (2011) A statistical mechanical approach for the computation of the climatic response to general forcings. *Nonlin Processes Geophys* 18:7–28
- Lucarini V, Colangeli M (2012) Beyond the linear fluctuation-dissipation theorem: the role of causality. *J Stat Mech* P05013. doi:[10.1088/1742-5468/2012/05/P05013](https://doi.org/10.1088/1742-5468/2012/05/P05013)
- Lucarini V, Blender R, Herbert C, Pascale S, Ragone F, Wouters J (2014) Mathematical and physical ideas for climate science. <http://arxiv.org/abs/1311.1190>
- Lunt DJ, Haywood AM, Schmidt GA, Salzmann U, Valdes PJ, Dowset HJ (2010) Earth system sensitivity inferred from Pliocene modelling and data. *Nat Geosci* 3:60–64
- Nicolson AM (1973) Forming the fast Fourier transform of a step response in time-domain metrology. *Electron Lett* 9:317–318
- Otto A et al (2013) Energy budget constraints on climate response. *Nat Geosci* 6:415–416
- Penland C (1996) A stochastic model of Indo-Pacific sea surface temperature anomalies. *Phys D* 98:534–558
- Penland C (2003) Noise out of chaos and why it won't go away. *B Am Meteor Soc* 84:921–925
- Previdi M, Liepert BG, Peteet D, Hansen J, Beerling DJ, Broccoli AJ, Frolking S, Galloway JN, Heimann M, Le Quéré C, Levitus S, Ramaswamy V (2013) Climate sensitivity in the Anthropocene. *Q J R Meteorol Soc* 139:1121–1131
- Ruelle D (1998a) General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium. *Phys Lett A* 245:220–224
- Ruelle D (1998b) Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms. *Nonlinearity* 11:5–18
- Ruelle D (2009) A review of linear response theory for general differentiable dynamical systems. *Nonlinearity* 22:855–870
- Saltzman B (2001) Dynamical paleoclimatology. Academic Press, New York
- Sardeshmukh P, Compo GP, Penland C (2000) Changes in probability associated with El Niño. *J Clim* 13:4268–4286
- Sherwood SC, Bony S, Dufresne J-L (2014) Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature* 505:37–42
- Shukla J, Hagedorn R, Hoskins B, Kinter J, Marotzke J, Miller M, Palmer T, Slingo J (2009) Revolution in climate prediction is both necessary and possible: a declaration at the world modelling summit for climate prediction. *B Am Meteor Soc* 90:175–178
- Stein U, Alpert P (1993) Factor separation in numerical simulations. *J Atmos Sci* 50:2107–2115
- Winton M, Takahashi K, Held IM (2010) Importance of ocean heat uptake efficacy to transient climate change. *J Clim* 23:2333–2344
- Wouters J, Lucarini V (2013) Multi-level dynamics I systems: connecting the Ruelle response theory and the Mori-Zwanzig approach. *J Stat Phys* 151:850–860