Computers and Electronics in Agriculture 223 (2024) 109093

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Near real-time spatial interpolation of hourly air temperature and humidity for agricultural decision support systems

Damien Jean Rosillon [a,*], Alban Jago [a], Jean Pierre Huart [a], Patrick Bogaert [b], Michel Journée [c], Sébastien Dandrifosse [a], Viviane Planchon [a]

[a] Walloon Agricultural Research Centre, 9 rue de Liroux, 5030 Gembloux, Belgium
[b] UCLouvain – Earth and Life Institute, Croix du Sud 2, 1348 Louvain-la-Neuve, Belgium
[c] Royal Meteorological Institute of Belgium, 3 avenue Circulaire 1180 Uccle, Belgium

A R T I C L E I N F O

Keywords:
Agriculture
DSS
Kriging
Agrometeorology
Reanalyses
Precision farming

A B S T R A C T

Food production will have to increase in the future to face the growing world population. Agricultural decision support systems (DSSs) are part of the solution since they aim at protecting crops against fungal diseases, a significant contributor to yield losses, while minimising pesticide use. DSSs are mainly driven by weather data which, currently, are usually obtained from the nearest available weather station. Since the latter is sometimes located far away from a farmer's field, this can lead to inaccurate recommendations. In order to provide better local weather data, spatial interpolation is a solution. However, since it must be delivered in near real-time, integrating a spatial interpolation process into an operational application necessitates addressing four constraints: *Accuracy*, *Robustness*, *Reliability* and *Latency*.

This study aimed at developing an operational application for a near real-time spatial interpolation of air temperature and relative humidity at hourly and daily timescales. The first objective was to select the best spatial interpolation models among five algorithms: nearest neighbour, inverse distance weighting, multiple linear regression, ordinary kriging and kriging with external drift. The best models were based on kriging with elevation as external drift. They largely reduced the mean absolute error (MAE) compared to using the nearest station: for hourly air temperature MAE dropped from 0.93 °C to 0.59 °C. It performed also better than multiple linear regression (MAE = 0.68 °C).

The second objective was to evaluate the impact of increasing station density by adding stations from the Belgian synoptic network. Additional stations improved *Accuracy* (MAE = 0.57 °C) but to a lesser extent than expected and had no clear impact on *Robustness*. The third objective was to assess the interest of using reanalyses (i.e. climate model outputs) as dynamic predictor variables. Reanalyses did not improve *Accuracy* (MAE = 0.62 °C) because, compared to elevation, they did not provide useful additional information that can be leveraged by the interpolation models. Using such a dynamical input also impacted *Reliability* negatively due to potential availability issues. Kriging models presented the highest computing times. However, *Latency* caused by the interpolation process itself was very small compared to the entirety of data processing.

The selected models were implemented on an online application "Agromet.be". Near real-time dissemination of interpolated weather data enables to produce local warnings helping farmers to take better decisions about spraying schedules. As a future improvement of spatial interpolation, integrating numerous personal weather stations owned by farmers seems promising.

1. Introduction

By 2050, food production will have to increase by 50 % to feed the projected 10 billion of inhabitants on earth (FAO, 2019; Searchinger et al., 2018). Crop pathogens and pests reduce yields and quality of agricultural productions (Savary et al., 2019). In particular, fungal diseases such as potato late blight (*Phytophthora infestans*) (Haverkort et al., 2009; Yuen, 2021) or wheat septoria tritici blotch (*Septoria tritici*) are responsible for a yield loss of 20 % worldwide (Fisher et al., 2018). Integrated pest management is an approach to crop production and

* Corresponding author.
E-mail address: d.rosillon@cra.wallonie.be (D.J. Rosillon).

https://doi.org/10.1016/j.compag.2024.109093
Received 9 November 2023; Received in revised form 19 April 2024; Accepted 23 May 2024
Available online 6 June 2024

protection that combines different management strategies and agronomical practices in order to grow healthy crops and minimise the use of pesticides (Duvivier et al., 2015). One strategy is to use efficient pesticides with adapted doses and timing (Duvivier et al., 2015; Fry, 2008). Decision support systems in agriculture (DSSs) make it possible by anticipating the risk of pest development. DSSs are profitable (El Jarroudi et al., 2015) and improve the efficiency of fungicide use when compared to a weekly spray schedule (Small et al., 2015a).

DSSs are partly based on crop and management tactics information (Small et al., 2015a) but weather data are the main driver of disease forecasting systems. Weather-based criteria indicate the pest emergence or infection periods, which are crucial to identify the right times to spray fungicides. An example of weather-based criterion is the sum of air temperatures (degree-days) to compute the time between an inductive rain and the exact day of orange wheat blossom midge emergence (Jacquemin et al., 2014). Another example of criterion is reaching a threshold of 90 % of relative air humidity during several hours for Cercospora leaf spot in sugar beet (El Jarroudi et al., 2021). DSSs leverage observed or forecasted weather data, and often both. Forecasted data are useful for preventive treatments, when there is a need to anticipate the pest development, but they are associated with a higher uncertainty. Observed weather data remain the core of many DSSs, when update with ground truth values is needed or that preventive action is not necessary. This study focuses on weather observations. An issue is that DSSs are usually fed with observed weather data from a limited number of automatic weather stations. Farmers have thus to refer to the nearest available stations (Small et al., 2015b), which leads to recommendations that are not always adapted to their situation, e.g., if their field is located far away (Rosillon et al., 2019). Pest management must be fine-tuned at farm or field-scale to be relevant (Been et al., 2009) and spatial interpolation of weather data is a solution to improve it. Recent studies used global gridded temperature datasets for application in agricultural systems modelling (Araghi et al., 2022). However, their spatial resolution (typically 0.1°x 0.1° equivalent to 7 km x 11 km in Wallonia) is too coarse for a field-scale weather data assessment.

Pest management DSSs usually require hourly data whereas spatial interpolation of weather data is mainly conducted at a daily, monthly or yearly frequency (Appelhans et al., 2015; Fonseca and Santos, 2018; Nashwan et al., 2019; Razafimaharo et al., 2020). In a review of spatial interpolation methods (Li and Heap, 2008), only 2 out of 13 identified meteorological applications dealt with hourly data and both concerned rainfall. Spatial interpolation of hourly air temperature and relative humidity is a field deserving further study.

Besides the need for data with fine spatial and temporal resolutions, successful pest management requires decision-making throughout the crop growing season (Small et al., 2015b). Weather data must therefore be delivered to DSSs in near real-time, that is, data should be updated every hour. For this reason, it is mandatory to deal with operational constraints (Pleau et al., 2002). In this study, the following four operational constraints were considered:

- *Accuracy*: “predictive accuracy based on the differences between the predicted values for, and the observed values of, new samples” (Li, 2017). Spatial interpolation needs to produce accurate weather data to properly model the risk of pest development.
- *Robustness*: “the insensitivity of products or processes against different sources of variation” (Kemmler et al., 2015). Spatial interpolation needs to be resilient when facing unusual weather conditions and to be performant in locations with peculiar topoclimatic context.
- *Reliability*: “the ability to perform as required, without failure, for a given time interval, under given conditions” (IEC 2015 cited by Kemmler et al., 2015). Interpolation procedures must be reliable in real-time conditions and cannot fail because DSSs cannot run with missing data.
- *Latency*: “the time interval between user input and system response” (Attig et al., 2017). Under favourable weather conditions, some fungal diseases have a very high development rate and DSSs must be highly responsive in assessing the risk of infection. Spatial interpolation needs to run quickly, within a few minutes.

Most of the studies about spatial interpolation of weather data focus on *Accuracy* (Appelhans et al., 2015; Fonseca and Santos, 2018; Nashwan et al., 2019). The other three constraints are common in the engineering research fields where the above definitions come from (IT for *Latency*, electronics for *Robustness* and *Reliability*) but not in geostatistics. One originality of this study is to propose a methodology to consider these constraints in the development of an operational application for near real-time spatial interpolation. Some commercial products currently provide near real-time interpolated weather data for an operational agricultural monitoring (i.e., iMetos Virtual Weather Stations, https://metos.at/en/imetos-vws/?cn-reloaded=1, consulted on 19th October 2023 or Weather Measures, https://weather-measures.com/, consulted on 19th October 2023) but not enough documentation is provided about the interpolation algorithms or the validation procedures neither are there peer-reviewed papers to validate the quality of the study. Only a few studies focused on interpolating weather data for operational DSSs, such as in Germany at a national scale (Zeuner and Kleinhenz, 2007; Paolo et al., 2011) and in France at a vineyard-specific scale (Bois et al., 2014). However, all those studies were based on static predictor variables (position and elevation) while our study proposes to go further by assessing the interest of integrating meteorological reanalyses as a dynamic predictor variable. In meteorology, reanalyses are gridded weather data recalculated on an hourly basis required to define the initial state of a climate model.

This study aimed at developing an operational online application for a near real-time spatial interpolation of air temperature and relative humidity at hourly and daily time-scale for agricultural DSSs in Wallonia. Since the appropriate method for spatial interpolation depends on the interpolated variable, the quality and density of observational data (Daly, 2006), the spatial and temporal resolution (Razafimaharo et al., 2020) and the topoclimatic context, conducting an in-depth study was essential to choose the best model for our context. In this paper, a “model” was defined as the combination of an interpolation algorithm and a set of predictor variables. In practical terms, the objectives were (1) to select among five spatial interpolation algorithms models based on longitude, latitude and elevation and supplied with data from the 28 automatic weather stations of a regional agrometeorological network (the Pameseb network); this selection process was conducted separately for air temperature and relative humidity and for both hourly and daily timescales, (2) to evaluate the impact of increasing station density by incorporating additional stations from a secondary network (the national official synoptic network) into the training dataset and (3) to assess the interest of integrating meteorological reanalyses as a predictor variable. A methodology to integrate operational constraints into the selection of the spatial interpolation model was also proposed.

2. Material & methods

2.1. The study area

The study area corresponds to Wallonia, i.e., the southern part of Belgium, that covers about 16,000 km^2. Belgium has a temperate oceanic climate highly influenced by its position at an average latitude and by the proximity of the North Sea. As air masses often come directly from the ocean, the Belgian climate is temperate, with relatively cool wet summers and relatively mild rainy winters. These conditions are favouring the development of fungal diseases.

In Wallonia, the distance from the sea varies between 43 and 263 km, while the elevation above sea level ranges from 10 to 694 m. The annual normal on the period 1991–2020 of the average air temperature exhibited spatial variations of 4 °C, while annual precipitations on the same period ranged from 700 mm (in Hesbaye, North of Wallonia) to 1400 mm (in Hautes Fagnes area, extreme East of Wallonia). In addition to these regional differences, meteorological variations due to local mesoclimates can be observed particularly in the rugged area of eastern and southern Wallonia. Such a spatial heterogeneity is in favour of spatial interpolation of weather data rather than relying on the nearest station for DSSs.

2.2. The datasets

2.2.1. Weather observations

The dataset included six years (01/01/2016–31/12/2021) of hourly and daily (minimum, mean and maximum) air temperature and relative humidity. Weather observations came from two distinct sources, namely the Pameseb network and the Belgian network of synoptic stations.

The Pameseb network consisted of 30 automatic weather stations across Wallonia (Fig. 1a) recording air temperature, relative humidity (both at 1,5 m), wind speed at 2 m, solar radiation, precipitation and leaf wetness. The network was dedicated to provide weather data to agricultural DSSs in near real-time. The stations were owned and maintained by the Walloon Agricultural Research Centre (CRA-W – https://www.cra.wallonie.be/en). The stations were visited every three months to carry out preventive maintenance (e.g., rain gauge and solar panel cleaning, lawn mowing, battery checks, etc.). In particular, temperature and relative humidity probes HMP 155 (Vaisala, Vantaa, Finland) were recalibrated on average every 15 months. The probes were placed in a naturally ventilated shelter meteoShield Professional (Barani, Bratislava, Slovakia) to protect them from direct solar radiations. Air temperature and humidity data were recorded every minute by a datalogger netDL 1000 (OTT HydroMet GmbH, Kempten, Germany). Every hour, minute data were transmitted by GPRS to a server where they were aggregated. Pameseb hourly data were thus the mean of 60-minute data points. A quality control based on automatic procedures and human daily supervision was done to ensure data quality. Each data was flagged in the database with its status ‘corrected’ or ‘not corrected’. ’Corrected’ flagged data were not selected as they were not fully representative of the weather conditions at the time and place. As far as possible, World Meteorological Organization good practices for station siting were respected (WMO, 2018). Out of the 30 available

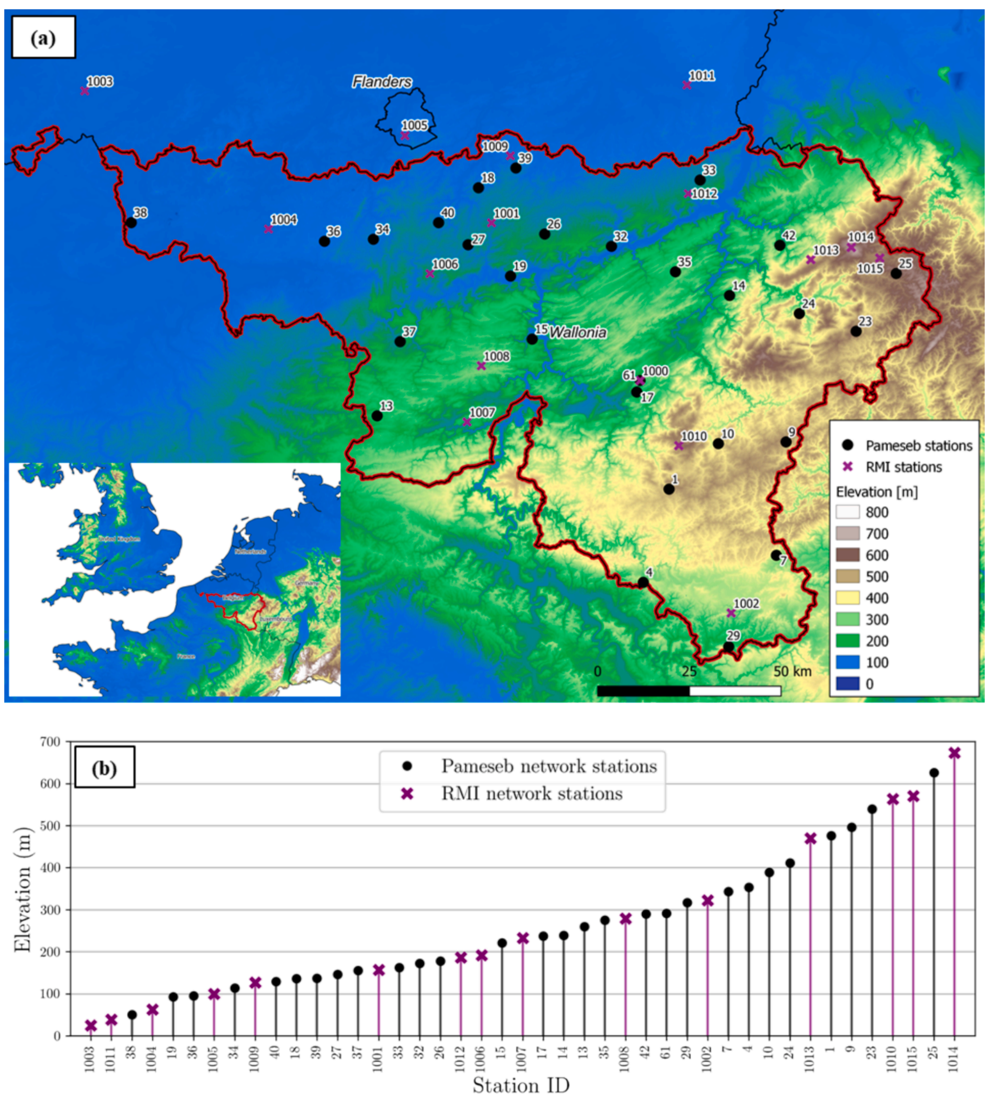

Fig. 1. Presentation of the study area. Fig. 1a displays a map of the stations of the Pameseb and RMI networks that were used in this study. The location of Wallonia in Europe is indicated by a red shape on the map in the lower left corner. Black dots correspond to Pameseb stations and purple crosses correspond to RMI stations. Elevation is used as background. All stations are labelled with their ID. The topographic map has been provided by EuroGeographics 2022 and is available for free at https://www.mapsforeurope.org/licence. Fig. 1b displays the elevation of the selected stations. They are ranged by increasing elevation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stations, 2 were not considered since they were not located in agricultural areas (in a forest area for ID 41 and industrial area for ID 30). The 28 selected stations were all located in fields or meadows. Their density was 1 station per 602 km^2. Data were available on Agromet.be (https://agromet.be/fr/pages/home/, consulted on 19th October 2023).

The Belgian network of synoptic stations (hereafter called 'RMI network') consisted of 28 automatic weather stations across Belgium recording ground weather parameters such as air temperature, relative humidity, precipitation, wind speed and pressure. These stations were operated by the Royal Meteorological Institute of Belgium (RMI, civil meteorology, www.meteo.be, consulted on 20th October 2023), Skeyes (civil aviation, www.skeyes.be, consulted on 20th October 2023) and Meteo Wing (military aviation, https://www.mil.be, consulted on 20th October 2023). They were placed on a ground covered with short grass according to the World Meteorological Organization recommendations (WMO, 2018). Data were acquired at high sampling frequency and then aggregated to 10-min values (Gonzàlez Sotelino et al., 2014). Concerning the Skeyes and Meteo Wing stations, only the data from the synoptic messages were available, i.e., 10-min averages between h40 and h50 every hour. RMI stations were subject to regular preventive maintenance. In particular, routine calibration was performed according to vendor requirements. The air temperature and relative humidity probes were placed in artificially ventilated RMI Stevenson shelters (Gonzàlez Sotelino et al., 2018). In order to identify and correct invalid or missing values, RMI conducted a quality control procedure based on automatic processes (Gonzàlez Sotelino, 2022) and human supervision (Bertrand et al., 2013). Out of the 28 stations, 15 of them were selected for their location inside Wallonia (12 stations) or close to it (three stations). In order to avoid information redundancy during the calibration and validation of the spatial interpolation, station ID 1000 (located in Humain, Wallonia) was discarded because it was too close (a few metres) from a Pameseb station located on the same site. The density of RMI selected stations was 1 station per 1.123 km^2. RMI network was considered as a secondary network for agricultural DSSs mainly due to the few numbers of stations located in Wallonia and to the fact that most of them were not specifically located in an agricultural context (e.g. airport).

2.3. Predictor variables

Two types of predictor variables were considered to feed the algorithms: static variables (latitude, longitude and elevation at stations location measured by GPS) and a dynamic variable (INCA reanalyses data, called hereafter 'INCA data'). INCA data are outputs from a climate model. They were produced by RMI with their nowcasting (i.e. weather forecasting on a very short term) system called 'INCA-BE'. They were based on a combination of Numerical Weather Prediction (NWP) forecasts and observation at surface stations (Reyniers et al., 2021). INCA data were considered as 'dynamic' predictor variables because values were updated every hour in contrast to 'static' predictor variables such as latitude, longitude and elevation which remained fixed over time. This dynamic aspect was expected to be relevant for interpolating changing variables such as weather data. Moreover, NWP models were based on mathematical equations representing the physical behaviour of the atmosphere. Their outputs were expected to add a physical information layer that was not captured by geostatistical models making use of topographic data. INCA data were produced on a basis on a 1 km^2 grid. Each Pameseb station was associated with the value of the INCA data at the corresponding pixel. To test the impact of INCA data for daily interpolations, hourly INCA data were aggregated to produce daily (minimum, mean and maximum) values.

2.4. Interpolation algorithms

Five spatial interpolation algorithms were investigated: nearest neighbour (NN), inverse distance weighting (IDW), multiple linear regression (mulLR), ordinary kriging (OK) and kriging with external drift (KED). NN estimates weather data at one location by attributing the weather observation at the nearest station. It is well-known that NN performs poorly but it remains an interesting reference algorithm because it is representative of the way most agricultural DSSs work: if no weather data are available for a farmer's field, weather data from the nearest station are used to run pest and disease models and to identify a potential risk for the farmer's crop (Small et al., 2015a). NN is also a comprehensible model that is convenient to refer to when interpreting the results: a good performance of NN is indicative of an easy-to-predict location, while a bad performance reveals a hard-to-predict one. IDW estimates weather data at one location as a weighted average of the neighbouring stations' observations. The weights are proportional to the inverse of the distance raised to some power. In this study, the power parameter was set to two, the most popular choice for spatial interpolation in environmental sciences (Li and Heap, 2008). MulLR estimates the weather data as a linear combination of the predictor variables. The parameters of the linear model are estimated from the observations at the stations. Kriging is a geostatistical method used to model a spatially correlated variable (which can be relevant for a weather data). As in the case of IDW, that variable is estimated as a weighted average of the observations at the neighbouring stations. However, the weights are more complex to account for the spatial structure of the observations. Kriging can be based directly on raw weather data (OK) or on detrended data (KED). Detrended data are computed by removing an "external" trend, such as elevation, which can partly explain the behaviour of the interpolated variable.

The ten studied models are presented in Table 1. For each one, an identifier is defined using a code that refers to the algorithm (i.e., mulLR,

Table 1
Interpolation algorithms and predictor variables used in each model. The columns "Data analysis" list the models tested for each analysis.

Model	Algorithm	Predictor variables	Data analysis		
			Pam.	Pam. + RMI	Pam. + INCA
Linear Regression					
mulLR-xy	Multiple linear regression	lon,lat	Yes	Yes	Yes
mulLR-xyz	Multiple linear regression	lon,lat,elev	Yes	Yes	Yes
mulLR-xyI	Multiple linear regression	lon,lat,INCA	No	No	Yes
mulLR-xyzI	Multiple linear regression	lon,lat,elev,INCA	No	No	Yes
Kriging					
OK-xy	Ordinary Kriging	lon,lat	Yes	Yes	Yes
KED-xyz	Kriging with external drift	lon,lat,elev	Yes	Yes	Yes
KED-xyI	Kriging with external drift	lon,lat,INCA	No	No	Yes
KED-xyzI	Kriging with external drift	lon,lat,elev,INCA	No	No	Yes
Others					
IDW	Inverse Distance	lon,lat	Yes	Yes	No
NN	Nearest Neighbour	lon,lat	Yes	Yes	No

OK, KED, IDW or NN) and to the predictor variables (i.e., x, y, z for coordinates and I for INCA data).

2.5. Models evaluation

Three analyses were conducted in accordance with the objectives of the paper. The first analysis ("Pameseb only") aimed at choosing a model trained with the Pameseb stations and static predictor variables. The second analysis ("Pameseb + RMI") aimed at assessing the contribution of stations from a secondary network (RMI network) to the training dataset. The models were trained with the Pameseb and RMI stations with static predictor variables. The third analysis ("Pameseb + INCA") aimed at assessing the interest of integrating meteorological reanalyses as dynamic predictor variables. The models were trained with the Pameseb stations, static predictor variables and the INCA data as dynamic predictor variables.

All models in Table 1 were evaluated in each analysis with the following exceptions: (1) models integrating INCA data were only evaluated for the analysis "Pameseb + INCA" and (2) NN and IDW were not evaluated for the analysis "Pameseb + INCA" because they could only integrate latitude and longitude as predictor variables. The models were trained for each hour independently on a training dataset containing the Pameseb observations, either temperature or relative humidity. RMI observations and INCA data were also used in the "Pameseb + RMI" and "Pameseb + INCA" analyses, respectively. The temperature and the relative humidity were predicted for each Pameseb station and each hour by a leave-one-out cross-validation (LOOCV) method. LOOCV is not a perfect method because errors can only be evaluated where a station exists (Daly, 2006). However, since the 28 Pameseb stations cover a broad diversity of topoclimates, LOOCV enables to make relevant assumptions about models' performances. The absolute error for each hour and each Pameseb station ($AE_{i,j}$) was computed as

$$AE_{i,j} = \left|\widehat{y}_{i,j} - y_{i,j}\right| \tag{1}$$

where $\widehat{y}_{i,j}$ is the predicted value and $y_{i,j}$ the actual observation for hour i and station j. General statistics of $AE_{i,j}$ (mean, median, maximum and standard deviation) over the whole six-years period including all Pameseb stations were computed. The mean absolute error per-station (MAE_j) was computed as

$$MAE_j = \frac{1}{n}\sum_{i=1}^{n} AE_{i,j} \tag{2}$$

Theoretically, the training dataset contained 52.608 h but in order to ensure a reasonable training quality, only hours with at least 20 available Pameseb measurements were accounted for. For the models using INCA data, only hours with at least 20 available Pameseb measurements and with available INCA data were accounted for. Thus, the size of the training dataset slightly changed with the analysis. The description above outlines the evaluation of models for hourly data. The same methodology was applied to daily data. Data analysis was done in Python using various libraries: pandas, numpy, sklearn, pykrige, pyarrow and matplotlib.

2.6. Semivariogram estimation

Kriging algorithms require the use of a semivariogram (VGM) that needs to be estimated from the data. In this paper, two approaches were tested to avoid the tedious task of manually fitting VGM. The first one was an automated estimation of a VGM for each hourly set of data. This approach is referred to as 'dynamic' hereafter, as the VGM may differ from one hour to another. The second approach was to estimate a single VGM that would be relevant for all hourly data, thus leading to a 'static' (i.e. time-invariant) VGM. Two possible ways were considered for modelling static VGMs: (1) to construct an experimental VGM based on all the observations over the whole six-years period and to fit a linear or spherical VGM (hereafter called 'static – average'); or (2) to choose an optimised VGM that minimises the MAE of hourly predictions on the whole six-years period (hereafter called 'static – grid search'). For this second approach, a grid search was conducted over all possible combinations of VGM parameters (type of model, range and nugget) for selecting the model that yielded the lowest MAE value when used in kriging. A part of this grid and its results is shown in Table 3 for the case of KED. To reduce the computational burden, analyses for VGM estimation were computed on 10 % of the dataset. This data subset was randomly selected and was identical for all the analyses. As the kriging prediction variance was not used in this study, it was only needed to model the spherical and exponential VGM models using the ratio between the nugget and the partial sill, which is just referred to as "nugget" hereafter.

3. Results and discussion

3.1. Prerequisites

The evaluation of the models delivers raw performance statistics such as AE or MAE but in order to integrate operational constraints, results are also discussed in terms of *Accuracy*, *Robustness*, *Reliability* and

Table 2
For each model, mean, median, maximum and standard deviation of the absolute errors of all LOOCV computed over the six-years period for hourly air temperature. Even if the training datasets for "Pameseb + INCA" and for "Pameseb only" were slightly different, this did not impact the results. Statistics for KED-xyz, mulLR-xyz and OK-xy for "Pameseb + INCA" were the same as for "Pameseb only" and were not reported in this table.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	0.59	0.37	13.78	0.69
mulLR-xyz	0.68	0.44	13.93	0.78
OK-xy	0.71	0.51	13.25	0.71
IDW	0.79	0.58	12.70	0.76
mulLR-xy	0.89	0.66	14.98	0.84
NN	0.93	0.70	13.40	0.96
Pameseb + RMI				
KED-xyz	0.57	0.35	17.76	0.69
mulLR-xyz	0.66	0.43	13.39	0.74
NN	0.94	0.60	27.30	1.01
Pameseb + INCA				
KED-xyzI	0.59	0.37	13.42	0.69
KED-xyI	0.62	0.39	12.77	0.71
mulLR-xyzI	0.65	0.40	13.25	0.77
mulLR-xyI	0.66	0.42	13.48	0.76

Table 3
MAE for KED-xyz and for different VGM models using all six-years data of Pameseb stations for the spatial interpolation of hourly air temperature. (a) Nugget (%) = 100* (Nugget / Partial sill) where the partial sill is arbitrarily set to 100.

Type of VGM	Linear	Spherical				Exponential			
Dynamic	0.64	0.62				0.62			
Static – Average	0.66	0.64				0.66			
Static – Grid search									
Range (km)		1	10	100	1000	1	10	100	1000
Nugget[a]									
0 %	**0.60**	0.77	0.76	0.62	0.60	0.77	0.76	0.62	0.60
5 %	0.59	0.77	0.76	0.61	0.60	0.77	0.76	0.62	0.59
90 %	0.67	0.77	0.76	0.73	0.70	0.77	0.76	0.65	0.68

Table 4
Average computing times (in millisecond) over 8,784 simulations for training and evaluating each model. The "Evaluation" column is the time to predict temperature across the 16,818 pixels of a 1x1 km grid covering Wallonia. The "Total" column is the sum of the "Training" and "Evaluation" columns. Models are sorted by increasing order of "Total" values.

Model	Training	Evaluation	Total
mulLR-xy	0.3	0.2	0.5
mulLR-xyz	0.3	0.2	0.5
NN	0.2	3.8	4.0
IDW	0.01	5.9	5.9
OK-xy	0.4	10.4	10.8
KED-xyz	3.1	10.9	14.0

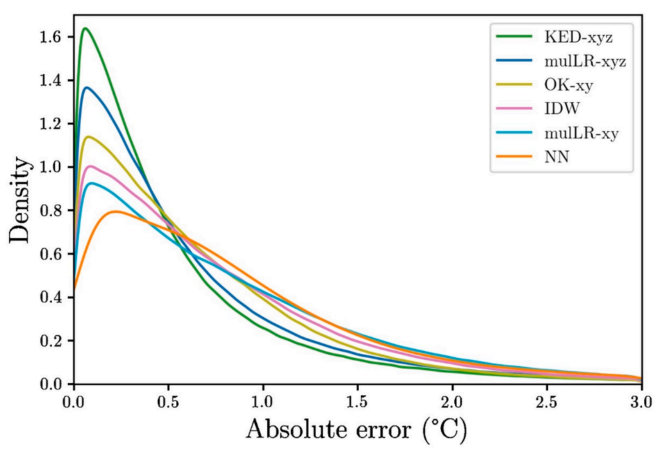

Fig. 2. Estimated probability density functions of absolute error (AE) for the models "Pameseb only". For the sake of clarity, plots are limited to a maximum AE of 3 °C.

Latency. *Accuracy* of the models was assessed by comparing MAE. *Accuracy* reveals the ability of a model to produce accurate predictions on average over the whole period and study area. As spatial interpolation performances may vary in space and time, a complementary analysis was required. *Robustness* of the models was assessed by an in-depth analysis of MAE by station and a focus on peculiar cases. It allowed us to investigate and understand the behaviour of the models for hard-to-predict locations and events. *Reliability* was assessed by analysing the risk of failure in the interpolation process and the possibility to access data in real-time. *Latency* was assessed by comparing computing times to train the models and to perform spatial interpolation on a 1 km^2 grid covering the whole Wallonia.

Within the six years investigated, a high diversity of weather phenomena was encountered. For example, summer 2021 was the wettest summer ever recorded in Belgium, while 2018 was on the contrary an exceptionally dry year. Years 2020 and 2022 were also the two warmest years ever recorded in Belgium. The models discussed hereafter were thus tested and compared on a large number of situations that covered a wide range of weather conditions including rare and hard-to-predict ones. This large panel permits strong statements about *Accuracy* and *Robustness* of the models. Moreover, the models present the advantage of being highly adaptive. For KED and mulLR models, the relationship between temperature (or humidity) and elevation could change but, as the models were retrained for each hour (or day), the external trend (for KED) or the linear relationship (for mulLR) was estimated on the current weather conditions that allowed it to better fit to the observations.

Focusing on the spatial interpolation of hourly temperature, for each model, Table 2 gives the statistics over the six-years period of all $AE_{i,j}$. Models are ranked from best to worst with respect to MAE. Fig. 2 displays the estimated probability density functions of AE for the models "Pameseb only". The same analysis was performed for daily temperature and hourly and daily relative humidity. Results and conclusions were similar to those for the hourly temperature and are therefore not detailed hereafter. Tables with detailed statistics are included in the appendix. Tables A.1–A.3 refer to daily temperature, Table A4 refers to hourly relative humidity and Tables A.5–A.7 refer to daily relative humidity.

3.2. Model based on Pameseb stations

This section addresses the question of building a model based on the Pameseb stations only and on static predictor variables. The selected model is used as a reference in Sections 3.3 and 3.4 to assess the interest of complexifying the model by integrating either RMI stations or INCA data.

3.2.1. Accuracy analysis

Table 2 shows that the best model for predicting hourly air temperature was kriging with elevation as an external drift (KED-xyz). From Fig. 2, it can be seen that KED-xyz showed the narrowest AE distribution with AE values higher than 2.5 °C that are exceptional. It is worth comparing KED-xyz performance with those for NN and mulLR-xyz. On the one hand, NN is currently used and advocated by most of the agricultural DSSs, as it is argued that hourly temperatures are highly variable in time and may exhibit complex behaviours in space, so that it would be safer to refer to the nearest weather station instead of interpolating. On the other hand, mulLR-xyz is the current model used to spatialise hourly air temperature for agricultural DSSs in Germany (Paolo et al., 2011). In our case however, the results show that KED-xyz outperformed both NN and mulLR-xyz, with reduced MAE values by 58 % and 15 %, respectively. KED-xyz showed the best MAE and therefore had the best *Accuracy*.

3.2.2. Robustness analysis

The global performances discussed above are valuable for determining the average performance of each model over the study area and over the six-years period. However, model performances are prone to variation depending on the station and the meteorological conditions. In this section, the errors of the models are analysed at station level with a focus on hard-to-predict locations and at meteorological event level that

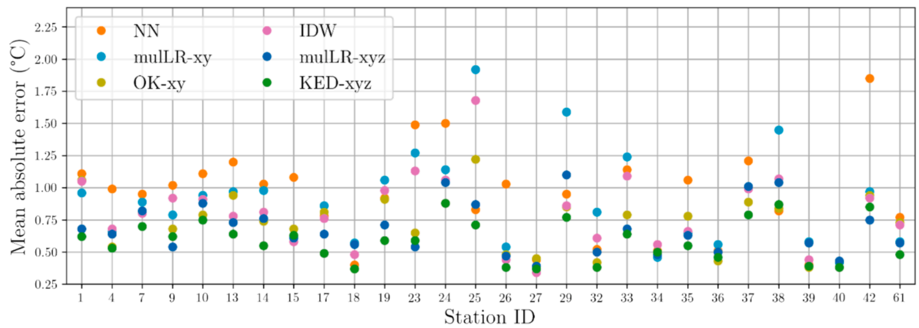

Fig. 3. Comparison of MAE for each Pameseb station and for all models.

are not revealed by the *Accuracy* analysis. This approach is crucial because the interpolated data are used to feed agricultural DSSs and one cannot be satisfied with a model that just performs well "on average". It is important to identify locations and particular meteorological events for which predictions are less reliable. In the following analysis, it is worth noting that the hard-to-predict meteorological events were not always related to hard-to-predict locations, illustrating the complementarity of both perspectives.

The station level analysis was based on Fig. 3. It can be seen that KED-xyz outperformed other models. It was the best-performing model for most of the stations (19 stations out of 28) while the second-best performing model was mulLR-xyz for three stations only. IDW and OK-xyz performed the best for two stations. The high variability observed in performance among the stations and the models deserves to be discussed. For a few stations, all models performed equally well. In these cases, the nice performances of NN suggest that the pairs of nearest stations (e.g.: IDs 27 and 40; 18 and 39; 34 and 36) were highly correlated. Those locations were not challenging: whatever the model, spatial interpolation performed well. On the opposite, all models performed poorly for a handful of stations (IDs 25, 29 and 38) that correspond to edge geographical conditions within the study area (in terms of latitude and longitude but also in terms of elevation for IDs 25 and 38 – see Fig. 1.). For such locations, as LOO prediction was used, data were not interpolated but rather extrapolated beyond the geographical conditions defined by the other stations to predict temperature at those points. For those hard-to-predict locations, KED-xyz was the best or among the best models. Fig. 3 also shows an unexpected hard-to-predict location (ID 42) because it was neither at the edge of the study area nor isolated. The NN score suggests that the station differs greatly from its neighbours. Indeed, it was located in the 'Région herbagère', an agricultural region composed of open landscapes, hilly meadows and hedgerows while most of its nearest stations were located in 'Haute Ardenne', a region of high elevation with more forests, a harsh climate with wide temperature variations and high rainfall. Based on this station level analysis, KED-xyz showed the best ability to produce local prediction for a majority of stations and therefore had the best *Robustness*.

The event level analysis was based on Fig. 4. The circled dot and the top-left part of the star's tail represent observations that were badly predicted by all the models, KED-xyz included. These two case studies are investigated hereafter. They illustrate that, on some rare occasions, spatial interpolation performed very poorly for a particular meteorological event.

The first event is picked up from the singular taillike area observed on the upper-left part for all models (Fig. 4). This area shows that, for very cold temperatures, prediction errors could sometimes be very high. Among others, black stars correspond to the data of station 10 on the 20th of January 2017. When compared with the observations of the two nearest stations (Fig. 5a), it appears that station 10 behaved differently. Recorded temperatures were much colder but plausible. This is a case of thermal inversion, a well-known weather phenomenon that often occurs during winter under anticyclonic conditions. The consequent clear and windless nights cause cold air to accumulate at the valley bottoms. It results locally in much lower temperatures than at higher elevations. This explanation is corroborated by the positioning of station 10 within a valley. Since spatial interpolation models generally consider a decrease in temperature with elevation, this leads to high prediction

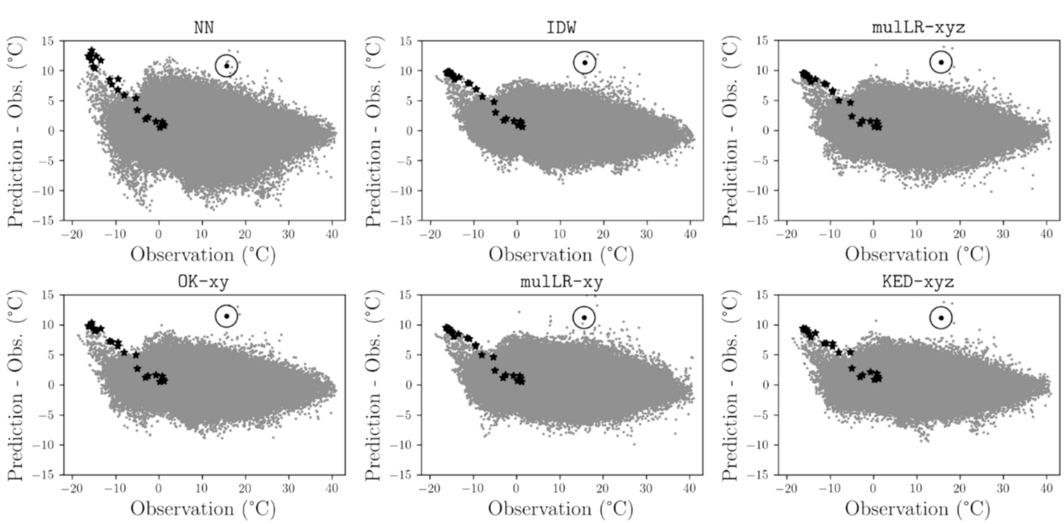

Fig. 4. For each model, scatter plots of the difference between prediction and observation as a function of observation for each LOOCV simulation over years 2016–2021. Each point refers to a given station and a given hour. The circled dot and the stars refer to hard-to-predict events for the station 27 and station 10 respectively.

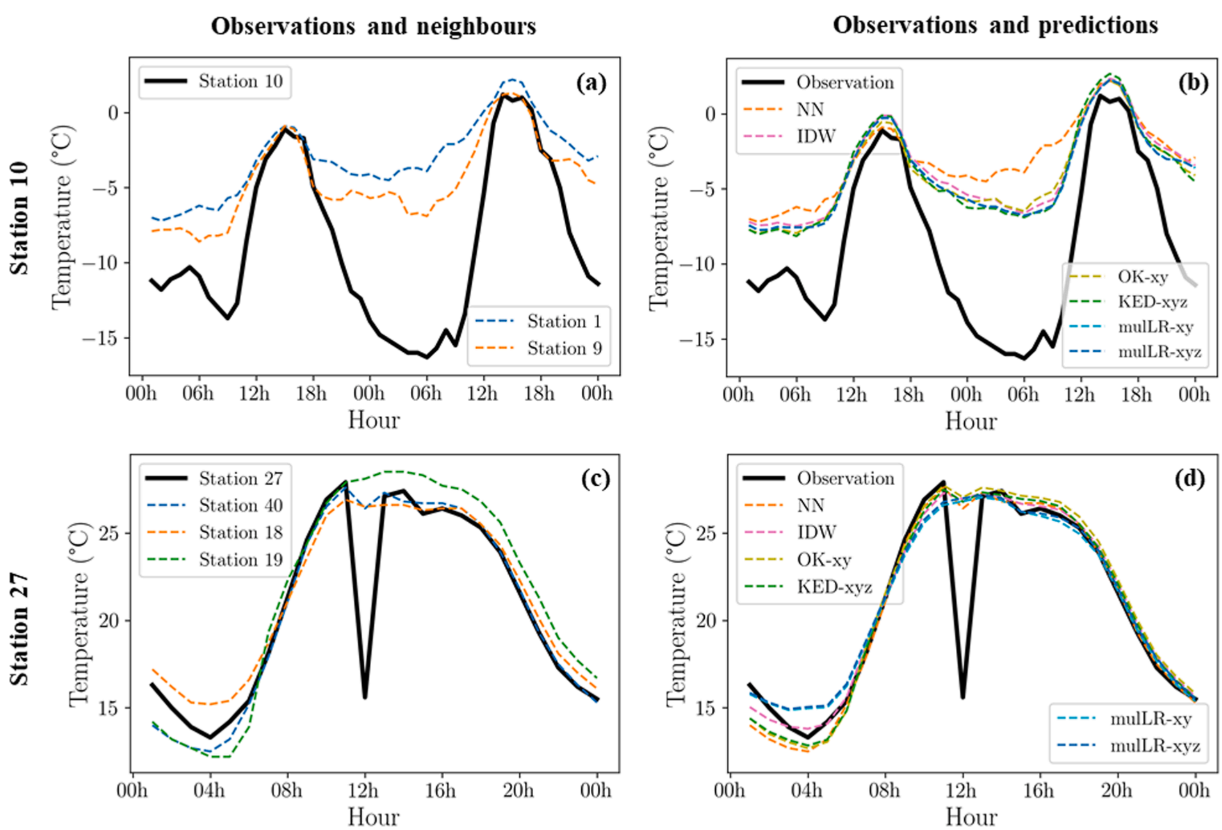

Fig. 5. Illustration of hard-to-predict events for station 10 and station 27. Fig. 5a and 5b compare hourly temperature observations for station 10 to neighbouring stations and predicted temperatures, respectively. Fig. 5c and 5d compare hourly temperature observations for station 27 to neighbouring stations and predicted temperatures, respectively.

errors. NN was unsurprisingly the worst model (Fig. 5b) but even the other models, KED-xyz included, were not able to properly account for thermal inversion. From this viewpoint, *Robustness* of all models was poor even for KED-xyz. This analysis is therefore not a way to differentiate models but is interesting for identifying the limitations of the approach: under some specific meteorological conditions, spatial interpolation performed poorly. This point is discussed in Section 3.8.

The second event explores one of the highest prediction errors. It corresponds to station 27 on the 15th of June 2017. By comparing with neighbouring stations (Fig. 5c) and analysing the dynamic of the temperature, it appears that this event was due to wrong data that were not filtered out during the quality control process. This event stresses the importance of the quality of the weather data used to train the models. For a near real-time interpolation, a strong automatic quality control is necessary to remove implausible observations from the training dataset. Fig. 5d, shows that all models were able to predict a sound value. It supports that LOO prediction could be an interesting way to detect implausible data into an automatic quality control process (Eischeid et al., 1995; Andresen et al., 2002) or for gap-filling. However, this question needs further study because, as demonstrated for temperature inversion, LOO predictions are not always representative of reality.

3.3. *Integration of RMI stations*

The spatial interpolation studied in the previous section was based on the 28 Pameseb stations only. It is expected that increasing the station density should improve the performances of the models (Berndt and Haberlandt, 2018). This section addresses the interest of integrating additional stations from the RMI network to the training dataset. The analysis focuses on KED-xyz and mulLR-xyz, (the two best performing models) and on NN (the model used by most of the agricultural DSSs).

3.3.1. *Accuracy analysis*

Integrating RMI stations for predicting hourly air temperature

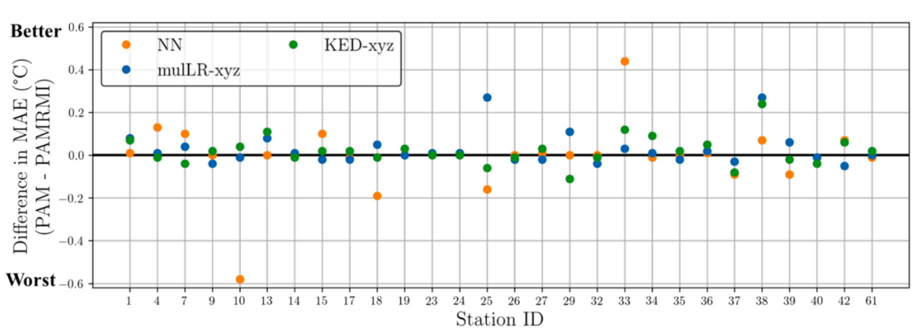

Fig. 6. Difference in MAE between benchmark "Pameseb only" and benchmark "Pameseb + RMI" for three models and for each Pameseb station.

improved the performances for KED-xyz and mulLR-xyz, but decreased the performances for NN (Table 2). RMI stations therefore improved *Accuracy* with MAE values reduced by 3 % for both KED-xyz and mulLR-xyz but the benefit was lower than expected. There are two possible explanations for this. The first one is that the number of RMI stations (15 stations of which 12 in Wallonia) was too small compared to the 28 Pameseb stations. The supplementary stations should probably be much more numerous for bringing a significant improvement. Moreover, the RMI stations were not particularly well positioned to fill the gaps between the Pameseb stations (Fig. 1a). Several RMI stations were close to Pameseb stations, while large areas were not covered by any of the networks. The second reason is that even if data quality for Pameseb and RMI datasets was considered equally good thanks to the quality of the sensors, the preventive maintenance and the data quality control, Pameseb and RMI networks were not really comparable. Indeed, unlike Pameseb stations, RMI stations were not specifically located in agricultural areas. Moreover, sensors (in particular radiation shelters) were of different types, thus potentially leading to differences for air temperature and relative humidity measurements. Finally, hourly data aggregation algorithms based on minutely data were different. It is difficult to know whether this impacted the results negatively but both datasets were most likely not fully equivalent.

3.3.2. Robustness analysis

Complementary to the analysis of *Accuracy*, it is also interesting to see if RMI stations improved *Robustness* of the models based on Pameseb stations, by analysing the changes in models performance at station level. Fig. 6 shows that the impact depends on the model and the station. For both KED-xyz and mulLR-xyz, the integration of RMI stations reduced MAE for 15 stations out of 28, it had no impact on two stations and degraded the MAE for 11 stations. The integration of RMI stations apparently had no clear impact on performances. It is interesting to focus on three peculiar cases to understand why, in such cases, the effect was notably negative, positive or unexpected.

(i) The highest improvement, NN excepted, was observed for station 38, one of the four hard-to-predict locations detailed in Section 3.2.2. It appears that this station was isolated from other Pameseb stations (Fig. 1a) and was the Pameseb station with the lowest elevation (Fig. 1b). Prediction at this location benefited from the addition of RMI stations, as two of them were in its close vicinity (IDs 1003 and 1004) and three of them were at similar elevations (IDs 1003, 1004 and 1011). (ii) Station 10 is also interesting because NN exhibited the highest score degradation. By adding RMI stations, RMI station 1010 became the nearest station but, with twice the elevation difference to station 1, its climate was less representative of station 10. Such occasional but important degradations explain the slight degradation of the global score for NN in Table 2. It is interesting to note that for this station, KED-xyz benefited from the addition of RMI stations even though the nearest station was not well correlated. (iii) For station 25, MulLR-xyz performed much better but surprisingly, the performance was degraded for KED-xyz. Station 25, the highest Pameseb station, was a hard-to-predict location because data were extrapolated for this location, as explained in Section 3.2.2. Thanks to the addition of a higher neighbouring station (ID 1014), temperature is now interpolated in the dimension of elevation. This additional station was expected to yield improved results for these models because they both integrate elevation.

Out of the four hard-to-predict locations investigated in Section 3.2.2, RMI stations improved the prediction of KED-xyz for two of them (IDs 38 and 42) but decreased the prediction for the two others (IDs 29 and 25). Integrating RMI stations was thus not systematically beneficial. We cannot therefore conclude to a *Robustness* improvement.

3.3.3. Reliability analysis

To complete the analysis, the impact of using RMI data on the *Reliability* of the system should also be discussed. As an external data source, RMI data should be automatically retrieved every hour from an external server. A failure of this retrieval leads to missing data, but the interpolation process itself does not fail. In such cases, the model can still be trained on the Pameseb stations. From this viewpoint, integrating RMI data does not affect *Reliability*. However, at the time of writing, RMI data were not available in real-time open access. This significantly affects the *Reliability* and using RMI data is therefore not an option for an operational system. If RMI ever opened its data, their integration in the operational spatial interpolation should be considered.

3.4. Integration of INCA data

3.4.1. Accuracy analysis

This section addresses the benefit of using INCA data as a dynamic predictor variable. From Table 2, it can be seen that the improvement of MAE when using INCA was almost equivalent to the use of elevation for both kriging and multilinear models. Indeed, for kriging models, MAE decreased from 0.71 °C for OK-xy to 0.59 °C for KED-xyz and 0.62 °C for KED-xyI. For multilinear models, MAE decreased from 0.89 °C for mulLR-xy to 0.68 °C for mulLR-xyz and 0.66 °C for mulLR-xyI, respectively. Moreover, MAE was nearly identical for mulLR-xyz and mulLR-xyzI on one side and for KED-xyz KED-xyzI on the other side. INCA data did not significantly improve the performance of the models that were already integrating elevation. One explanation is that the INCA system already incorporated elevation in its process (Haiden et al., 2011) but it did not provide useful additional information that can be leveraged by the tested interpolation algorithms. As a result, integrating INCA data did not improve *Accuracy* of the models.

3.4.2. Reliability analysis

Beyond the question of *Accuracy*, the use of INCA data raises some operational considerations and impacts the *Reliability* of the system. As an external source, INCA data must be automatically retrieved every hour from an external RMI server. A failure in this retrieval leads to missing INCA data and, unlike RMI stations, it does negatively impact *Reliability*. Indeed, as INCA is used as a predictor variable, missing data prevent prediction. Furthermore, as a dynamical model, INCA data are prone to errors. Although very rare, some largely erroneous values were identified in the INCA dataset considered in this study, leading to invalid predictions. As using INCA data introduces a potential source of error without offering any clear benefit for our operational application, such models were discarded.

3.5. Semivariogram selection

For the sake of clarity, the selection of the VGM model used for the kriging algorithms was not previously discussed. It will be addressed here, with a focus on the analysis of KED-xyz for hourly air temperature spatial interpolation based on Pameseb stations only.

A dynamic VGM was expected to have a better *Accuracy* because it can adapt to peculiar or extreme weather conditions that could occasionally occur. However, Table 3 shows that the dynamic VGM did not have the lowest MAE and therefore offers no benefit from an *Accuracy* viewpoint compared to some static VGM. Regarding *Latency*, a static VGM is computationally more cost-effective. Finally, a dynamic VGM has a lower *Robustness* as it is more impacted by occasional erroneous data.

Among the static VGM, the best performances were presented for the grid search VGM. Among all grid search VGM, the linear model with a 5 % nugget had the best *Accuracy*. However, the model without nugget is simpler and has the advantage of producing a map without discontinuities at the stations. For this reason and given the very small score difference, it was preferred for our application and was used in all kriging algorithms previously discussed.

When using kriging, a lot of attention is usually paid to the VGM modelling. However, within the context of this study, it appeared that fine-tuning the VGM (by precisely adapting the parameters of the common exponential, spherical, and linear models) did not have a substantial impact on the performance. Indeed, MAE varied little except when using VGM with ranges as low as 1 and 10 km or with high nuggets. For the sake of brevity, details are only given here for hourly air temperature, but similar analyses for relative humidity and daily statistics lead to the same conclusions. As a result, the static linear VGM with no nugget was selected due to its operational convenience and its good performance.

3.6. Latency analysis

To ensure near real-time data dissemination on an operational application, it is crucial to use fast-running models. Since the models need to be trained every hour and for each weather variable, minimising *Latency* caused by interpolation tasks is essential. *Latency* was estimated as the average computing time for training and running the models. For every hour of year 2020 (a total of 8,784 h), models were trained on the Pameseb stations and then used to predict temperature on the 16,818 pixels of a 1x1 km grid covering Wallonia. Table 4 shows the average computing times of each model.

mulLR models were significantly faster compared to kriging. However, when considering the complete data acquisition process, an additional computing time of a few milliseconds becomes negligible. The entirety of data processing which involved tasks like transmitting data from stations to server, formatting and storing data, automatic quality checks and performing spatial interpolation, required approximately 20 min. This duration was primarily attributed to the inclusion of security delay to ensure the completeness of data transmission. Consequently, the actual *Latency* caused by the interpolation process itself was not a determining factor in the model selection. Despite a higher *Latency*, KED-xyz remained the preferred model due to a superior *Accuracy* and *Robustness*.

3.7. The online Agromet.be application

Qualitative and local weather data are fundamental for successful fungicides application. Spatial interpolation is a solution to provide field-scale weather data but they must be easily available in near real-time to users (Been et al., 2009). An online application called "Agromet.be" (https://www.agromet.be) was developed for data dissemination. It provides interpolated (hourly and daily) temperature and relative humidity in near real-time for any location inside Wallonia. Efforts have been made to maximise data accessibility. Data can be accessed freely by a large majority of users (farmers, research, agricultural advisors and government) while access must be negotiated and is potentially free for private companies. Data can be manually exported through a graphical user interface or automatically exported thanks to an Application Programming Interface (API). Interpolated data can be extracted at coordinates encoded by the user or as a map of 1 km^2 pixels. This map can also be easily displayed on the application. Clicking on a pixel displays the time series for that specific point (Fig. 7). Historical data are available for download for the past two years. Hourly interpolated data are available about 20 min after the end of the corresponding hour.

3.8. Limitations of the study

This study aimed at identifying the best models for spatial interpolation of hourly and daily temperature and relative humidity. KED-xyz gave the best performances with low prediction errors, but this model is not perfect. Errors can be very high in some rare instances, e.g. when thermal inversions occur. As the final goal is to use interpolated data with agricultural DSSs to deliver advice to farmers, it is important to analyse the sensitivity of the underlying models to weather data. Agricultural models frequently work with threshold values on weather data (e.g., in the modelling of potato late blight, a threshold of 90 % relative humidity must be exceeded to initiate a risk of infection). A small error around this value can totally change the simulated risk and deliver

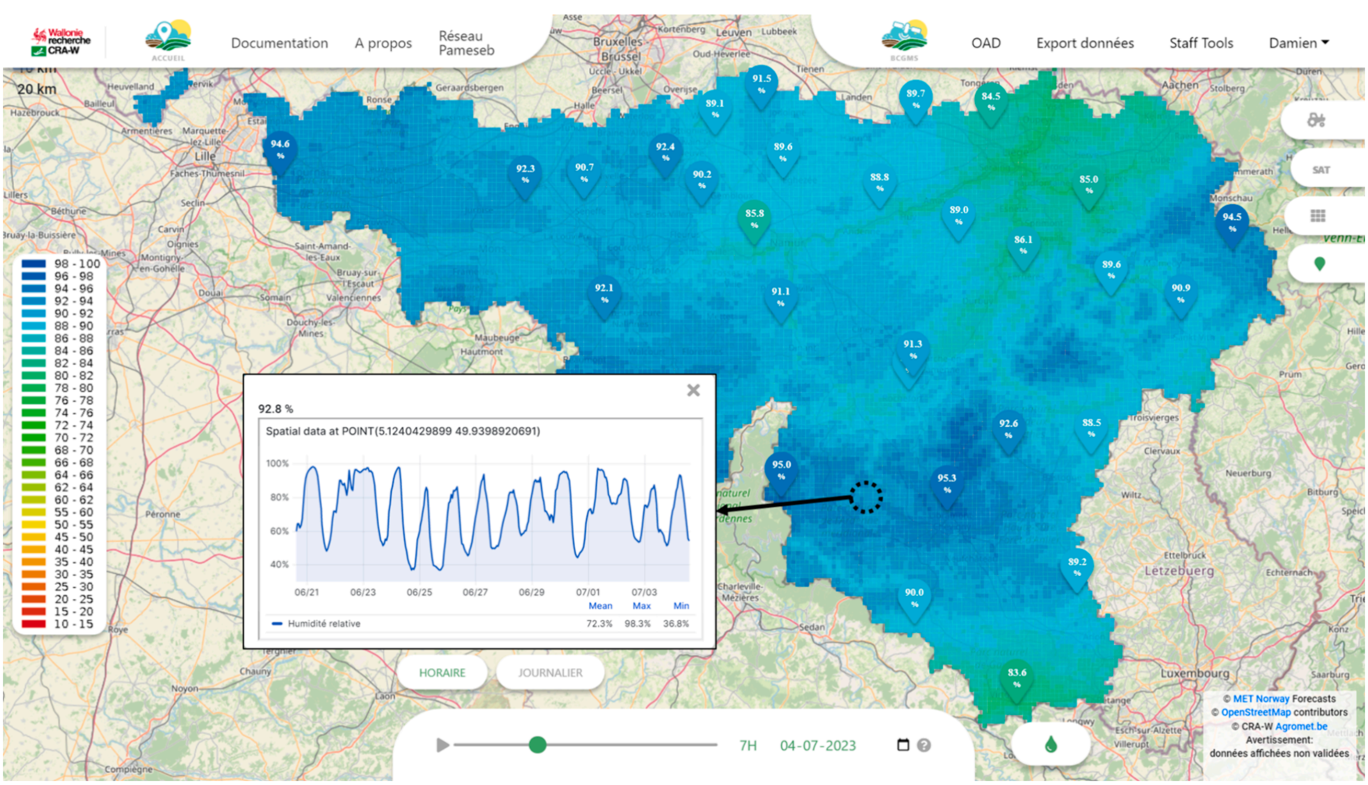

Fig. 7. Screenshot of the application www.agromet.be, showing the spatial interpolation of hourly relative humidity on the 4th of July 2023, 5 h UTC.

wrong advice. Other DSSs work with cumulative temperatures (e.g. wheat phenology modelling) and a small but systematic error in the predicted temperature can largely influence the output.

A second limitation concerns the dataset used in this study. It was the result of a quality control based on automatic procedures but also on human supervision. The errors of spatial interpolation models were estimated on a clean dataset. To ensure that this study is representative of reality, the real-time operational system should also rely on a clean dataset and therefore requires a strong automatic quality control.

4. Conclusions

Near real-time spatial interpolation of weather data is a promising option to improve current agricultural DSSs which are still largely based on the nearest weather station, sometimes located far away from the place of interest. Local weather data, and therefore local warnings, enable farmers to take better decisions, to save time and money and to reduce environmental impacts due to pesticides. This study showed that the best spatial interpolation models based on kriging (KED-xyz) largely reduced the MAE compared to using the nearest station: for hourly air temperature MAE dropped from 0.93 °C to 0.59 °C. It performed also better than multiple linear regression (MAE = 0.68 °C).

Due to the need to ensure near real-time data dissemination, four major operational constraints were identified to evaluate the different models, namely *Accuracy*, *Robustness*, *Reliability* and *Latency*. With respect to these constraints, the KED-xyz, algorithm based on kriging with elevation as external drift, Pameseb stations only and with a static semivariogram, was considered as the optimal choice, both for temperature and humidity and for hourly and daily frequencies. Using RMI stations improved *Accuracy* (MAE = 0.57 °C) but to a lesser extent than expected. *Robustness* was not clearly improved by RMI stations but they significantly impacted *Reliability* because they were not available in real-time open access. Using INCA data as dynamic predictor variables did not improve *Accuracy* (MAE = 0.62 °C), most likely because, compared to elevation, INCA data did not provide useful additional information that can be leveraged by the tested interpolation algorithms. Also, using such a dynamical input negatively impacted *Reliability* due to potential availability issues. Integrating dynamic components (predictor variables or semivariogram estimation) to the models was expected to improve their performances. Nevertheless, it did not make models more accurate, while it increased their complexity. Models with static components were therefore favoured.

It is worth noting that the methodology used in this paper can be replicated for other weather variables or for other countries. However, it is important to note too that the optimal models could differ from those selected in Wallonia. Taking the time to look beyond the overall statistics through an in-depth investigation of the simulations is crucial. Such an analysis helps to highlight singular local behaviours and identify the limitations of the models. This process is not fully automated and necessitates expert knowledge to interpret the results.

Several avenues could be explored in the future. First, the use of outputs of climate models is worth investigating. Numerical Weather Prediction models are continuously improved (Alley et al., 2019) and they are expected to become more and more relevant as dynamic predictor variables. Secondly, it would be interesting to investigate the optimal density of weather stations to improve *Accuracy*. This issue is not trivial and deserves a dedicated study because it likely depends on various parameters including the topographic variability of the study area. Finally, integrating farmers' stations seems a promising option. In Wallonia, the number of these stations is estimated to be around 850 (WalDigiFarm asbl personal communication) and this number increases at a fast pace. However, several issues need to be addressed to leverage such data in near real-time, this including the quality of and the access to the data.

5 Funding

This project was funded by the Walloon Agricultural Research Centre (CRA-W) by a tax exemption operation of research in Belgium under the Moerman law.

6. Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Chat GPT 3.5 (OpenAI, USA) in order to check grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Damien Jean Rosillon: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Alban Jago:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – review & editing. **Jean Pierre Huart:** Data curation, Software, Writing – review & editing. **Patrick Bogaert:** Methodology, Writing – review & editing. **Michel Journée:** Methodology, Writing – review & editing. **Sébastien Dandrifosse:** Writing – review & editing. **Viviane Planchon:** Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Pameseb weather observation data used in this paper will be made available on request. The authors have no permission to share RMI data.

Acknowledgment

The authors are grateful to Rodolphe Geradin for his work on the PAMESEB weather stations, to Cyril Vansteenberge and Valéry Michaud for their involvement in human data quality control and to Manfred Röhrig (ISIP) and Olivier Deudon (Arvalis) for the discussions about operational spatial interpolation challenges. The authors also thank the Royal Meteorological Institute of Belgium for sharing weather data and INCA data.

Appendix

Table A1
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **daily mean air temperature**.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	0.38	0.26	5.85	0.41
mulLR-xyz	0.43	0.29	6.21	0.46
OK-xy	0.49	0.37	6.04	0.44
IDW	0.58	0.44	6.11	0.52
mulLR-xy	0.68	0.52	6.09	0.6
NN	0.63	0.49	8.6	0.6
Pameseb + RMI				
KED-xyz	0.37	0.24	6.56	0.41
mulLR-xyz	0.41	0.27	5.83	0.44
NN	0.61	0.45	8.90	0.60
Pameseb + INCA				
KED-xyzI	0.37	0.25	5.73	0.40
KED-xyI	0.38	0.26	6.11	0.40
mulLR-xyzI	0.41	0.27	7.43	0.46
mulLR-xyI	0.42	0.28	6.13	0.46

Table A2
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **daily minimum air temperature**.

	Mean	Median	Max	Std
Pameseb only				
KED-xyz	0.77	0.49	8.69	0.83
OK-xy	0.84	0.58	9.32	0.85
mulLR-xyz	0.88	0.56	9.14	0.96
IDW	0.91	0.66	9.09	0.88
mulLR-xy	1.00	0.74	8.78	0.95
NN	1.14	0.80	12.20	1.20
Pameseb + RMI				
KED-xyz	0.76	0.45	9.57	0.87
mulLR-xyz	0.84	0.54	8.98	0.90
NN	1.16	0.70	12.70	1.27
Pameseb + INCA				
KED-xyzI	0.78	0.49	10.52	0.85
KED-xyI	0.81	0.51	11.33	0.89
mulLR-xyI	0.88	0.55	9.33	0.97
mulLR-xyzI	0.89	0.54	10.03	1.00

Table A3
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **daily maximum air temperature**.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	0.43	0.32	8.49	0.41
mulLR-xyz	0.50	0.36	9.73	0.49
OK-xy	0.62	0.50	5.98	0.51
IDW	0.70	0.54	6.94	0.60
mulLR-xy	0.77	0.59	11.11	0.69
NN	0.78	0.70	8.30	0.62
Pameseb + RMI				
KED-xyz	0.41	0.30	5.70	0.39
mulLR-xyz	0.48	0.35	7.06	0.46
NN	0.80	0.60	7.90	0.67
Pameseb + INCA				
KED-xyzI	0.43	0.32	7.10	0.40
mulLR-xyzI	0.45	0.34	6.58	0.42
KED-xyI	0.46	0.35	6.11	0.42
mulLR-xyI	0.47	0.36	6.23	0.43

Table A4
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **hourly relative humidity**.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	3.13	2.15	55.57	3.32
OK-xy	3.49	2.54	53.64	3.46
mulLR-xyz	3.59	2.49	57.90	3.71
IDW	3.70	2.71	56.30	3.60
mulLR-xy	4.02	2.92	57.01	3.87
NN	4.33	3.10	75.20	4.56
Pameseb + RMI				
KED-xyz	3.19	2.23	56.34	3.37
mulLR-xyz	3.55	2.53	53.18	3.54
NN	4.51	3.10	79.80	4.87
Pameseb + INCA				
KED-xyzI	3.16	2.17	55.73	3.34
KED-xyI	3.37	2.40	55.50	3.46
mulLR-xyzI	3.42	2.36	55.36	3.60
mulLR-xyI	3.58	2.56	58.98	3.59

Table A5
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **daily mean relative humidity**.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	2.03	1.55	37.29	1.88
mulLR-xyz	2.29	1.76	35.48	2.10
OK-xy	2.40	1.91	33.00	2.08
IDW	2.59	2.03	36.84	2.26
mulLR-xy	2.82	2.19	34.31	2.44
NN	2.99	2.36	56.48	2.70
Pameseb + RMI				
KED-xyz	2.13	1.63	42.89	1.95
mulLR-xyz	2.29	1.77	36.02	2.05
NN	3.10	2.40	56.48	2.89
Pameseb + INCA				
KED-xyzI	2.04	1.56	38.51	1.88
KED-xyI	2.21	1.73	37.91	1.98
mulLR-xyzI	2.22	1.68	35.59	2.06
mulLR-xyI	2.39	1.87	36.91	2.11

Table A6
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **daily minimum relative humidity**.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	2.59	1.94	51.10	2.47
mulLR-xyz	3.02	2.25	54.14	2.88
OK-xy	3.07	2.44	46.87	2.65
IDW	3.32	2.59	49.61	2.96
mulLR-xy	3.62	2.78	58.37	3.27
NN	3.77	3.10	59.50	3.21
Pameseb + RMI				
KED-xyz	2.59	1.98	47.87	2.43
mulLR-xyz	3.00	2.27	53.02	2.78
NN	3.65	2.90	61.90	3.21
Pameseb + INCA				
KED-xyzI	2.61	1.97	49.68	2.46
mulLR-xyzI	2.79	2.12	53.84	2.61
KED-xyI	2.91	2.30	47.28	2.54
mulLR-xyI	3.05	2.38	53.59	2.70

Table A7
For each model, mean, median, maximum and standard deviation of the absolute errors of all the leave-one-out simulations computed over the six-years period for **daily maximum relative humidity**.

Model	Mean	Median	Max	Std
Pameseb only				
KED-xyz	2.36	1.58	40.33	2.69
mulLR-xyz	2.53	1.64	40.61	2.96
OK-xy	2.57	1.76	42.87	2.91
IDW	2.63	1.81	40.85	2.90
mulLR-xy	2.76	1.88	40.24	2.97
NN	3.16	2.00	60.20	4.00
Pameseb + RMI				
KED-xyz	2.49	1.68	49.54	2.83
mulLR-xyz	2.51	1.68	38.39	2.83
NN	3.46	2.20	60.20	4.37
Pameseb + INCA				
KED-xyzI	2.40	1.59	100.00	2.81
mulLR-xyzI	2.53	1.64	39.96	2.99
KED-xyI	2.55	1.71	100.00	3.00
mulLR-xyI	2.63	1.77	39.77	2.94

References

Alley, R.B., Emanuel, K.A., Zhang, F., 2019. Advances in weather prediction. Science 363, 342–344. https://doi.org/10.1126/SCIENCE.AAV7274.

Andresen, L., Hellsten, E., Rissanen, P., Palsdóttir, T., Arason, T., 2002. Quality control of meteorological observations. Nor. Meteorol Inst.

Appelhans, T., Mwangomo, E., Hardy, D.R., Hemp, A., Nauss, T., 2015. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. Spat. Stat. Spatial and Spatio-Temporal Models for Interpolating Climatic and Meteorological Data 14, 91–113. https://doi.org/10.1016/j.spasta.2015.05.008.

Araghi, A., Martinez, C.J., Olesen, J.E., Hoogenboom, G., 2022. Assessment of nine gridded temperature data for modeling of wheat production systems. Comput. Electron. Agric. 199, 107189 https://doi.org/10.1016/j.compag.2022.107189.

Attig, C., Rauh, N., Franke, T., Krems, J.F., 2017. System Latency Guidelines Then and Now – Is Zero Latency Really Considered Necessary?, in: Harris, D. (Ed.), Engineering Psychology and Cognitive Ergonomics: Cognition and Design, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 3–14. doi: 10.1007/978-3-319-58475-1_1.

Been, T., Berti, A., Evans, N., 2009. Review of new technologies critical to effective implementation of Decision Support Systems (DSS's) and Farm Management Systems (FMS's).

Berndt, C., Haberlandt, U., 2018. Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. J. Hydrol. Reg. Stud. 15, 184–202. https://doi.org/10.1016/j.ejrh.2018.02.002.

Bertrand, C., Gonzàlez Sotelino, L., Journée, M., 2013. Quality control of 10-min air temperature data at RMI. Adv. Sci. Res. 10, 1–5. https://doi.org/10.5194/asr-10-1-2013.

Bois, B., Volta, A., Rega, M., Caffarra, A., Costa, F., Antolini, G., Tomei, F., Galizia, S., Nascimben, J., Crestini, D., others, 2014. GIS, mechanistic modelling and ontology: a performing mix for precision and sustainable viticulture. In: 37th World Congress of Vine and Wine. International Organisation of Vine and Wine, pp. 1–10.

Daly, C., 2006. Guidelines for assessing the suitability of spatial climate data sets. Int. J. Climatol. 26, 707–721. https://doi.org/10.1002/joc.1322.

Duvivier, M., De Proft, M., Bodson, B., Suffert, F., Planchon, V., 2015. Distribution of the airborne inoculum of wheat leaf rust and septoria tritici blotch: impact on epidemics in wheat fields and implications for integrated pest management. Prom Proft Michel.

Eischeid, J.K., Baker, C.B., Karl, T.R., Diaz, H.F., 1995. The Quality Control of Long-Term Climatological Data Using Objective Data Analysis. J. Appl. Meteorol. 1988–2005 (34), 2787–2795.

El Jarroudi, M., Kouadio, L., Beyer, M., Junk, J., Hoffmann, L., Tychon, B., Maraite, H., Bock, C.H., Delfosse, P., 2015. Economics of a decision–support system for managing the main fungal diseases of winter wheat in the Grand-Duchy of Luxembourg. Field Crops Res. 172, 32–41. https://doi.org/10.1016/j.fcr.2014.11.012.

El Jarroudi, M., Chairi, F., Kouadio, L., Antoons, K., Sallah, A.-H.-M., Fettweis, X., 2021. Weather-based predictive modeling of Cercospora beticola infection events in sugar beet in Belgium. J. Fungi 7, 777. https://doi.org/10.3390/jof7090777.

FAO, 2019. The State of Food and Agriculture 2019. Moving forward on food loss and waste reduction. Rome. Licence: CC BY-NC-SA 3.0 IGO.

Fisher, M.C., Hawkins, N.J., Sanglard, D., Gurr, S.J., 2018. Worldwide emergence of resistance to antifungal drugs challenges human health and food security. Science 360, 739–742. https://doi.org/10.1126/science.aap7999.

Fonseca, A.R., Santos, J.A., 2018. High-Resolution Temperature Datasets in Portugal from a Geostatistical Approach: Variability and Extremes. J. Appl. Meteorol. Climatol. 57, 627–644. https://doi.org/10.1175/JAMC-D-17-0215.1.

Fry, W., 2008. Phytophthora infestans: the plant (and R gene) destroyer. Mol. Plant Pathol. 9, 385–402. https://doi.org/10.1111/j.1364-3703.2007.00465.x.

Gonzàlez Sotelino, L., De Coster, N., Beirinckx, P., Peeters, P., 2018. Intercomparison of Shelters in the RMI AWS Network. World Meteorological Organisation.

Gonzàlez Sotelino, L., De Coster, N., Beirinckx, P., Peeters, P., 2014. New concept for Belgian AWS network, in: WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation, (WMO).

Gonzàlez Sotelino, L., 2022. Automatic Quality Control in the RMI AWS Network. Presented at the 2022 WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-2022), Paris (France).

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., Gruber, C., 2011. The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. Weather Forecast. 26, 166–183. https://doi.org/10.1175/2010WAF2222451.1.

Haverkort, A.J., Struik, P.C., Visser, R.G.F., Jacobsen, E., 2009. Applied Biotechnology to Combat Late Blight in Potato Caused by Phytophthora Infestans. Potato Res. 52, 249–264. https://doi.org/10.1007/s11540-009-9136-3.

Jacquemin, G., Chavalle, S., De Proft, M., 2014. Forecasting the emergence of the adult orange wheat blossom midge, Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae) in Belgium. Crop Protection 58, 6–13. https://doi.org/10.1016/j.cropro.2013.12.021.

Kemmler, S., Eifler, T., Bertsche, B., Howard, T., 2015. Robust reliability or reliable robustness? Integrated consideration of robustness- and reliability-aspects. OPUS - Online Publikationen der Universität Stuttgart. https://doi.org/10.18419/opus-4604.

Li, J., 2017. Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? PLoS ONE 12, e0183250. https://doi.org/10.1371/journal.pone.0183250.

Li, J., Heap, A.D., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. Geosci. Aust. Record 2008 (23), 1–137.

Nashwan, M.S., Shahid, S., Chung, E.-S., 2019. Development of high-resolution daily gridded temperature datasets for the central north region of Egypt. Sci. Data 6, 138. https://doi.org/10.1038/s41597-019-0144-0.

Paolo, R., Benno, K., Thorsten, Z., Barbara, K., Beate, T., Jeanette, J., 2011. Decision support systems in agriculture: administration of meteorological data, use of geographic information systems (GIS) and validation methods in crop protection warning service. Effic. Decis. Support Syst. Chall. Curr. Future. https://doi.org/10.5772/20809.

Pleau, M., Pelletier, G., Colas, H., Lavallée, P., Bonin, R., 2002. Reliability and robustness in real-time control applications. Global Solutions for Urban Drainage 1–12. https://doi.org/10.1061/40644(2002)322.

Razafimaharo, C., Krähenmann, S., Höpp, S., Rauthe, M., Deutschländer, T., 2020. New high-resolution gridded dataset of daily mean, minimum, and maximum

temperature and relative humidity for Central Europe (HYRAS). Theor. Appl. Climatol. 142, 1531–1553. https://doi.org/10.1007/s00704-020-03388-w.
Reyniers, M., Simon, A., Delobbe, L., Deckmyn, A., Goudenhoofdt, E., 2021. The INCA-BE system: ten years of operational nowcasting and its applications at the national meteorological service of Belgium. R. Meteorol. Inst. Belgium, Brussels, Belgium. https://doi.org/10.5281/zenodo.5798952.
Rosillon, D., Huart, J.P., Goossens, T., Journée, M., Planchon, V., 2019. The Agromet Project: A Virtual Weather Station Network for Agricultural Decision Support Systems in Wallonia, South of Belgium. In: Palattella, M.R., Scanzio, S., Coleri Ergen, S. (Eds.), Ad-Hoc, Mobile, and Wireless Networks, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 560–565. https://doi.org/10.1007/978-3-030-31831-4_39.
Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019. The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 3, 430–439. https://doi.org/10.1038/s41559-018-0793-y.
Searchinger, T., Waite, R., Hanson, C., Ranganathan, J., Dumas, P., Matthews, E., 2018. Creating a Sustainable Food Future. A Menu of Solutions to Feed Nearly 10 Billion People by 2050. World Ressources Institute, Washington DC.
Small, I.M., Joseph, L., Fry, W.E., 2015a. Development and implementation of the BlightPro decision support system for potato and tomato late blight management. Comput. Electron. Agric. 115, 57–65. https://doi.org/10.1016/j.compag.2015.05.010.
Small, I.M., Joseph, L., Fry, W.E., 2015b. Evaluation of the BlightPro decision support system for management of potato late blight using computer simulation and field validation. Phytopathology 105, 1545–1554. https://doi.org/10.1094/PHYTO-05-15-0117-R.
WMO, W., 2018. Guide to instruments and methods of observation, 2018th and 2021 editions ed. World Meteorological Organization WMO, Geneva.
Yuen, J., 2021. Pathogens which threaten food security: Phytophthora infestans, the potato late blight pathogen. Food Secur. 13, 247–253. https://doi.org/10.1007/s12571-021-01141-3.
Zeuner, T., Kleinhenz, B., 2007. Use of geographic information systems in warning services for late blight. EPPO Bull. 37, 327–334. https://doi.org/10.1111/j.1365-2338.2007.01134.x.