



# A coupled human–natural system to assess the operational value of weather and climate services for agriculture

Yu Li<sup>1</sup>, Matteo Giuliani<sup>2</sup>, and Andrea Castelletti<sup>2,1</sup>

<sup>1</sup>Institute of Environmental Engineering, ETH Zurich, Wolfgang-Pauli-Str. 15, 8093 Zurich, Switzerland

<sup>2</sup>Department of Electronics, Information and Bioengineering, Politecnico di Milano, Piazza L. da Vinci, 32, 20133 Milan, Italy

*Correspondence to:* Andrea Castelletti (andrea.castelletti@polimi.it)

Received: 22 May 2017 – Discussion started: 29 May 2017

Revised: 1 August 2017 – Accepted: 9 August 2017 – Published: 18 September 2017

**Abstract.** Recent advances in weather and climate (W&C) services are showing increasing forecast skills over seasonal and longer timescales, potentially providing valuable support in informing decisions in a variety of economic sectors. Quantifying this value, however, might not be straightforward as better forecast quality does not necessarily imply better decisions by the end users, especially when forecasts do not reach their final users, when providers are not trusted, or when forecasts are not appropriately understood. In this study, we contribute an assessment framework to evaluate the operational value of W&C services for informing agricultural practices by complementing traditional forecast quality assessments with a coupled human–natural system behavioural model which reproduces farmers' decisions. This allows a more critical assessment of the forecast value mediated by the end users' perspective, including farmers' risk attitudes and behavioural factors. The application to an agricultural area in northern Italy shows that the quality of state-of-the-art W&C services is still limited in predicting the weather and the crop yield of the incoming agricultural season, with ECMWF annual products simulated by the IFS/HOPE model resulting in the most skillful product in the study area. However, we also show that the accuracy of estimating crop yield and the probability of making optimal decisions are not necessarily linearly correlated, with the overall assessment procedure being strongly impacted by the behavioural attitudes of farmers, which can produce rank reversals in the quantification of the W&C services operational value depending on the different perceptions of risk and uncertainty.

## 1 Introduction

Weather and climate (W&C) services, defined as information on past, present, and future weather and climate useful to assist decision making (GFCS, 2014), can provide valuable aid to a variety of economic sectors, including hydropower production (e.g. Garcia-Morales and Dubus, 2007), drought management (e.g. Mwangi et al., 2014), flood protection (e.g. Cloke et al., 2017), and disease spread control (e.g. Thomson et al., 2006). These services are particularly important in agriculture (Hammer et al., 2001), where weather-sensitive decisions, such as crop choices or irrigation scheduling (e.g. Dutra et al., 2013; Winsemius et al., 2014; Wetterhall et al., 2015), are frequently to be taken. Here, W&C services are expected to be even more helpful over the next years, when extreme weather conditions will be more frequent and intense (Dai, 2011).

Over past decades, W&C services have undergone broad development in many parts of the world (Cloke and Pappenberger, 2009; Bauer et al., 2015; Brunet et al., 2015). The existence of slow, and hence predictable, variations in sea surface temperature, sea ice, soil moisture, and snow cover, which interact with the atmosphere and impact on the global climate, can be used to extend predictability at the seasonal timescale (Palmer and Hagedorn, 2006). Although some limitations still exist (e.g. Palmer et al., 2005; Lee et al., 2011), the recent increase in model resolutions (e.g. Prodhomme et al., 2016a), the improvement of initialization procedures (e.g. Prodhomme et al., 2016b), and the more accurate representation of some physical processes (e.g. Hourdin et al., 2013) has considerably advanced the ac-

curacy of W&C services, with current state-of-the-art products showing good forecast skills even over seasonal and longer timescales (Doblas-Reyes et al., 2013).

A causal link between better forecast quality and higher operational value is, however, not necessarily straightforward (e.g. Ritchie et al., 2004; Ramos et al., 2013), especially when forecasts do not reach their final users, when the provider is not trusted, or when forecasts are not appropriately understood (e.g. Ramos et al., 2010; Frick and Hegg, 2011). In other words, while quantifying forecast quality is a necessary step in the assessment of W&C services, other indicators should be considered for capturing the stakeholders' judgment of the value of the forecast products, i.e. their operational value, particularly when this evaluation differs from the opinion of forecasters (Hartmann et al., 2002). Yet, most assessments reported in the literature focus solely on forecast quality, defined as the similarity between the forecast estimates and the actual observations of weather or hydrological variables based on some statistically formulated performance metrics (Murphy, 1993).

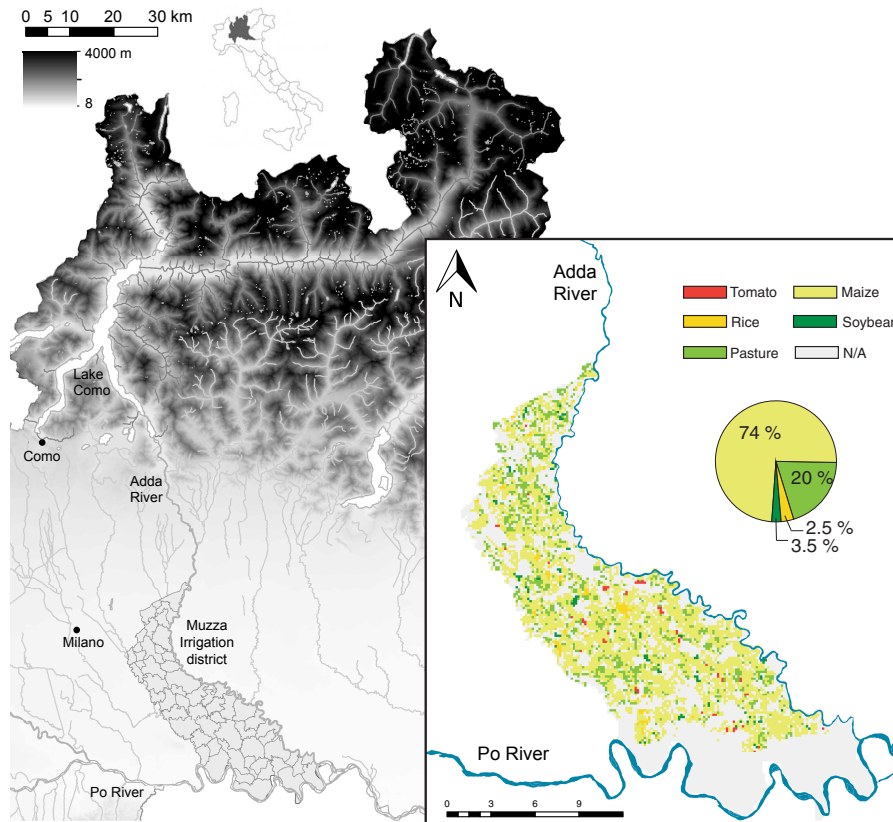
Recent attempts to assess the operational value of W&C services have tended to apply long-term forecasts for feeding simulation models in order to predict decision-relevant information, such as soil water availability for irrigation scheduling (e.g. Wang and Cai, 2009; Calanca et al., 2011) or crop production for cropping pattern decisions (e.g. Hansen, 2004; Baigorria et al., 2008). The use of process-based simulation models contributes to a better understanding of W&C services by stakeholders and users as it makes it possible to transform weather forecasts (e.g. precipitation and temperature) into decision-relevant information (e.g. crop yield) through a transparent, objective, and reproducible procedure. For example, although farmers can quantify the risks associated with predictions of a dry season, they would benefit much more from information on the anticipated crop yield and the associated risk of crop failure (e.g. Challinor et al., 2005). In addition, the relationship between weather and decision-relevant variables is often nonlinear, and an error in the weather forecast will not be linearly propagated into an error of the same magnitude in the crop yield prediction. The quality of forecast products evaluated on weather variables can differ from the quality of evaluations performed on crop yield: two forecast products characterized by different levels of accuracy in predicting weather variables can provide similar predictions of crop productivity; vice versa, two products having similar skills in predicting temperature and precipitation can show different performance in predicting crop yield. Quantifying the value in terms of forecast accuracy in predicting decision-relevant information is therefore crucial for improving stakeholders' trust in W&C services.

Although the model-based prediction of decision-relevant information is surely a step forward from the end users' perspective, high-quality forecasts may still be unused by stakeholders (e.g. Rayner et al., 2005; Coulibaly et al., 2015). For example, an attempt to increase forecast accuracy by pro-

viding more early warnings often implies a risk of increasing the number of false alarms, ultimately discouraging the use of W&C services in an operational context due to different perceptions of risk and uncertainty (Demeritt et al., 2007). In addition, many studies have shown how stakeholders' adoption of weather forecasting is influenced by their social context (e.g. Hansen, 2002; Suarez and Patt, 2004; Crane et al., 2010). Such evidence motivates exploring how users' behavioural factors influence the uptake and the use of W&C services, and suggests the need to quantify the operational value of W&C services and the improvement in the system performance obtained by informing stakeholders' decisions with W&C services (e.g. Zhu et al., 2002; Mylne, 2002; Giuliani et al., 2015; Denaro et al., 2017).

In this work, we propose a new framework for assessing the operational value of W&C services, which puts human in the loop by integrating traditional forecast quality assessments with a behavioural model reproducing farmers' decisions. The proposed framework relies on a three-stage procedure, which starts by investigating the quality of post-processed forecast products. These forecasts are then used as input to an integrated model representing a coupled human–natural system (CHNS; see Liu et al., 2007). This includes process-based models of the physical environment to predict decision-relevant information, coupled with decision models which describe the farmers' decision-making process. Given the predicted climate forcing as inputs, the integrated CHNS model simulates production of different crops, from which each farmer selects the crop to cultivate by maximizing the expected net profit at the end of the agricultural season (Giuliani et al., 2016). This combination of process-based and decision models contributes a comprehensive and complete framework for assessing W&C services and allows the evaluation of both the forecast quality and operational value. In addition, the decision model includes heterogeneous behavioural factors, specifically diverse levels of farmers' risk aversion (or degree of trust) with respect to forecast uncertainty, which enable the exploration of the sensitivity of the overall assessment of W&C services with respect to variability of stakeholders' behaviours.

We demonstrate the potential of our approach by developing an application in the Muzza agricultural district, in northern Italy. The district is organized as 39 irrigation units, each including a number of farms receiving a continuous water supply through an extensive irrigation network. A set of state-of-the-art long-range climate forecast products are collected from the European Centre for Medium-Range Weather Forecasts (ECMWF), National Centres for Environmental Prediction (NCEP), and the Canadian Seasonal to Inter-annual Prediction System (CanSIPS). The forecast horizon ranges from 7 months to 10 years. Post-processing (i.e. downscaling and bias correction) is then used to address the mismatch of temporal and spatial resolution between the simulation models and the raw forecast products, as well as to resolve the systematic biases and uncertainty in the ensem-



**Figure 1.** The Muzza district with the cropping pattern observed in 2004.

ble forecasts. Finally, by simulating the combined process-based and decision models over the period 2001–2005, with 2003 and 2005 being extreme drought years, we perform the proposed three-stage assessment of forecast quality and operational value of W&C services. First, we assess the traditional forecast quality by comparing forecast meteorological variables against observed data. Then we measure, via model simulations, the prediction accuracy of crop yield as an intermediate assessment of decision-relevant information for supporting farmers in improving their practices. Finally, we quantify the operational value in terms of the payoff (or opportunity cost) of using W&C services for informing the selection of the cropping pattern. This value is contrasted with the upper bound of the system performance obtained using “perfect forecasts” as well as a baseline situation where farmers use few simple empirical forecast models, including climatology or past observations. In addition, our decision models allow us to explore alternative uses of W&C services, which depend on the personal behavioural attitude of farmers and on their level of trust in the forecast products. In particular, we explore three different levels of farmers’ risk aversion, namely risk averse, risk neutral, or risk prone, which create a spectrum of possible behavioural attitudes (e.g. Mosley and Verschoor, 2005)

The paper is organized as follows: in the next section we describe the study area, while Sect. 3 provides details about the methodology, including the data preparation and the modelling framework. Results and discussion are then reported in Sects. 5–6. Finally, conclusions and directions for future research are presented in Sect. 7.

## 2 Study site

In this work, the assessment of W&C services is conducted on the CHNS of the Muzza irrigation district, located south-east of the city of Milan (see Fig. 1). The selected district is one of the largest agricultural areas in the region with arable land of about 700 km<sup>2</sup>. Maize (ca. 74 % of the surface) and temporary grasslands (ca. 20 % of the surface) are currently the major cultivated crops, with minor crops including rice, soybean, wheat, tomato, and barley. Irrigation is provided through an extensive irrigation network (more than 4000 km in total length) served by the Adda River and feeding 39 irrigation units, which are organized in 1722 *comizi* and include around 12 000 farms.

Historically, water availability has not been a major limiting factor to the economic development of this area. Rather, farmers have operated the canal network to contribute to

flood protection. Yet, in recent decades, climate change has exacerbated the severity of extreme events (García-Herrera et al., 2010). For example, two severe droughts in 2003 and in 2005 generated acute crop failures and exacerbated the conflicts between agriculture and other sectors (Anghileri et al., 2013). These critical events are predicted to become more frequent over the next years (Lehner et al., 2006), representing a major challenge for the sustainability of the agricultural practices in this region.

In this context, the use of W&C services offers a promising option for supporting agricultural activities as the improved forecast skill over medium to long lead times provides valuable information about the future agricultural season prior to the sowing date. Such information is key for better informing cropping pattern decisions to select the best crops with respect to farmers' objectives (e.g. the one characterized by the highest expected profit). Moreover, most W&C services are freely available online and thus represent a cost-effective solution to improve the resilience of agricultural systems without introducing infrastructural changes, such as modifying or expanding the irrigation canal network.

### 3 Methodology

The overall workflow of our assessment framework is composed of three main steps, as detailed in Fig. 2: (i) forecast quality assessment of post-processed W&C services using retrospective forecast (i.e. hindcast) products; (ii) extension of the forecast quality analysis via model-based prediction of decision-relevant variables, namely crop production at the end of the agricultural season; and (iii) evaluation of cropping pattern decisions in terms of farmers' payoff at the end of the agricultural season as simulated by the integrated CHNS model. In this step, different levels of risk aversion can be simulated to explore the sensitivity of the overall assessment with respect to farmers' behavioural attributes.

The first step of the framework (upper block in Fig. 2) starts with the post-processing of the hindcast data (see Sect. 3.1) in order to remove the modelling biases generally affecting the simulation of coupled ocean–atmospheric models, such as high precipitation frequency and low precipitation intensity (Ines and Hansen, 2006). The bias-corrected dataset is further downscaled using a stochastic weather generator in order to resolve the spatial and temporal scale mismatch between hindcast data and model inputs. In particular, the weather generator allows us to perform not only the spatial downscaling but also the temporal disaggregation to obtain forecasts of daily precipitation and temperature from the forecast products, in the case that they have a monthly time resolution (see Table 1). The comparison of the post-processed precipitation and temperature forecast products with the on-site historical observations provide a first estimate of the forecast quality.

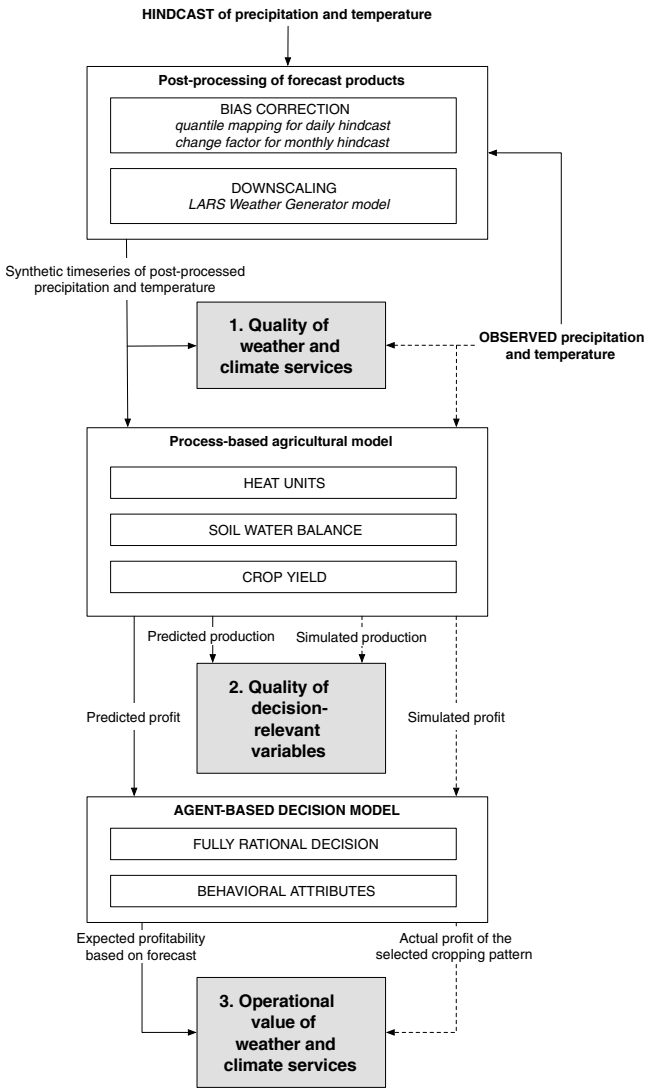
The post-processed hindcast dataset is then fed into the process-based component of our integrated CHNS model (middle block in Fig. 2). This includes a spatially distributed process-based representation of the Muzza irrigation district (see Sect. 3.2), which extends the assessment of the forecast quality by looking at the difference between forecast crop yields and that simulated using observed time series of precipitation and temperature, assuming that the expected crop yield represents the main determinant of farmers' cropping pattern decisions as in other similar applications (Hansen, 2004; Baigorria et al., 2008).

The human component of the CHNS (bottom block in Fig. 2) is finally introduced in the form of an agent-based decision model (see Sect. 3.3), which allows us to simulate farmers' cropping pattern decisions driven by different forecast information. This decision model makes it possible to couple the simulation of the process-based model and the prediction of crops' profitability with the selection by each farmer-agent of the best cropping pattern as the one characterized by the highest profitability. The agent-based model allows us to test different behavioural criteria, capturing alternative levels of farmers' risk aversion (or degree of trust) with respect to the forecast uncertainty. In particular, we consider a spectrum of behaviours ranging from a fully optimistic farmer, who makes decision on the basis of the best possible situation, to an extremely pessimistic farmer, who, instead, looks at the worst-case performance. Then, given the selected cropping pattern of each farmer-agent, the model is simulated using the observed values of precipitation and temperature to obtain the production and the associated profit at the end of the agricultural season. The estimated agent's profit is compared with that obtained under the hypothesis of perfect foresight, which represents the ideal upper-bound of the system performance. The operational value of W&C services is finally estimated as the percentage of agents making optimal decisions using the forecast products, which represents the opportunity cost of using W&C services with respect to having perfect foresight. The results are then validated against the profit obtained by the agents when informed with simple empirical forecasts.

Details about each step of the proposed framework, corresponding to a different block in Fig. 2, are reported in the next sections.

#### 3.1 Post-processing of forecast products

The first step of the proposed procedure (upper box in Fig. 2) aims at post-processing the forecast products. Depending on the characteristic of the forecast, we perform bias correction by means of the change factor approach (Crochemore et al., 2016) or the quantile mapping technique (Déqué, 2007). Given the strong intra-annual seasonal cycle of our study site, the bias correction was applied on a monthly basis. This also means that the corrections are differentiated according to the lead time (Doblas-Reyes et al., 2013), as the forecasts are



**Figure 2.** Detailed workflow of the proposed three-stage assessment framework of W&C services (solid line) against observed conditions (dashed line).

considered for the same starting months (i.e. April, when the agricultural season starts).

The quantile-based mapping technique is a bias correction method which builds the transfer function by mapping the cumulated density function (CDF) of climate model outputs onto the site-based observations. The calibrated transfer function is used later on to derive corrected estimates from new incoming outputs by resolving the mismatch between the observed site measurements and the simulated climate outputs. Quantile-based mapping is applied to forecast products providing daily trajectories of precipitation and temperature, which enable a proper estimation of the corresponding CDFs. This step becomes questionable in the case of monthly hindcast due to the limited dimension of the dataset. In this case, we apply the change factor approach, in which a mul-

tiplicative factor is used to scale the value of precipitation, while an additive factor is used to adjust the temperature for each month.

Although the systematic bias in the hindcast dataset can be partially solved by using bias correction, the difficulty in dealing with the uncertainty of ensemble forecasts remains a challenge. Previous studies (e.g. Tippett et al., 2007) have suggested the probabilistic use of long-range weather forecasts by deriving the statistical signatures from ensemble forecasts, such as the mean or the anomaly values. This statistic is then compared with the climatology to indicate whether the incoming year is expected to be normal, wet, or dry. As a consequence, the information on the intra-annual variability of the climate, which is critical for crop growth and agricultural management, is not preserved. Rather, in this work the multi-ensemble data are assimilated into a stochastic weather generator, whose parameters are calibrated from observations and then perturbed based on the forecast conditions. This allows us to generate synthetic time series of precipitation and temperature maintaining the information estimated by the ensemble forecast. In addition, the stochastic weather generator can also disaggregate a monthly forecast into daily values, which are needed to run the process-based model in the next step, without losing the generality of the statistical behaviour of the variables. The LARS-WG model (Semenov and Barrow, 1997) is selected for this task as it has been reported to outperform many other weather generators (Hashmi et al., 2011).

The perturbation factors of the mean daily precipitation intensity ( $F_i^{P,\text{pert}}$ ) and of the mean temperature ( $F_i^{T,\text{pert}}$ ), along with the monthly average number of wet (dry) days ( $F_i^{\text{wet,pert}}, F_i^{\text{dry,pert}}$ ), are the key parameters of the weather generator. These are determined according to Eq. (1), where  $m$  is the number of days in the  $i$ th month and  $\mathbf{1}(\bullet)$  is the binary operator that returns 1 if the daily precipitation intensity  $P_{i,j}$  is larger than 1 mm (wet), and 0 otherwise (dry) (Ceballos et al., 2004). In particular, Eqs. (1a)–(1b) represent the expected change of precipitation frequency with respect to the average historical observations ( $P^h$ ) measured by the local stations located in the considered study area, while Eq. (1c) specifies the change of precipitation intensity conditioned on rainy days. For the precipitation, the computed perturbation factor is used to scale up (down) the original parameter values. The change of temperature is formulated in Eq. (1d) as an additive term. Specifically, we estimated a monthly perturbing factor for each ensemble member by aggregating the daily hindcast data, and then we considered the average factor across the ensemble’s members. The perturbation parameters are then used to generate synthetic, daily time series for 1 year according to the considered forecast information:

$$F_i^{\text{wet,pert}} = \frac{E[\sum_{j=1}^m \mathbf{1}(P_{i,j} \geq 1 \text{ mm})]}{E[\sum_{j=1}^m \mathbf{1}(P_{i,j}^h \geq 1 \text{ mm})]}, \quad (1a)$$

$$F_i^{\text{dry,pert}} = \frac{E[\sum_{j=1}^m 1(P_{i,j} < 1 \text{ mm})]}{E[\sum_{j=1}^m 1(P_{i,j}^h < 1 \text{ mm})]}, \quad (1b)$$

$$F_i^{P,\text{pert}} = \frac{E[P_i | \text{wet}]}{E[P_i^h | \text{wet}]}, \quad (1c)$$

$$F_i^{T,\text{pert}} = E[T_i] - E[T_i^h]. \quad (1d)$$

### 3.2 The process-based agricultural model

The second step of our procedure (middle block in Fig. 2) aims at estimating expected crop yield at the end of the agricultural season, which is assumed to be decision-relevant information for the considered farmer-agents. For this purpose, we rely on a spatially distributed process-based model of the Muzza irrigation district (Giuliani et al., 2016), which is composed of three interlaced modules: (i) a distributed-parameter water balance module that simulates water resources, conveyance, distribution, and soil–crop water balance (Facchi et al., 2004; Gandolfi et al., 2006); (ii) a heat unit module that simulates the sequence of growth stages as a function of temperature (Neitsch et al., 2011); and (iii) a crop growth module that estimates the optimal and actual yields, accounting for the effects of water stresses due to the insufficient water supply that may have occurred during the agricultural season (Steduto et al., 2009).

The water balance module partitions the irrigation district with a regular mesh of cells with a side length of 250 m, which was selected to properly reproduce the spatial distributions of all the modelled processes, especially in terms of water balance (Vassena et al., 2012). Each individual cell identifies a soil volume which extends from the soil surface to the lower limit of the root zone. This soil volume is subdivided into two layers, modelled as two nonlinear reservoirs in cascade: the upper one (evaporative layer) represents the upper 15 cm of the soil; the bottom one (transpirative layer) represents the root zone and has a time-varying depth. The water percolating out of the bottom layer constitutes the recharge to the groundwater system.

The heat unit module defines the relationships between the temperature and some variables and parameters related to the crop growth stage (e.g. root length, basal coefficient, leaf area index), which also influence the water balance module. According to heat unit theory (Neitsch et al., 2011), crop growth stage at time  $t$  in  $i$ -cell is defined as a function of the cumulated heat units ( $HU_t^{(i)}$ ). A range is defined for each crop: the minimum is the base temperature  $T_b$ , which defines the day of sowing (i.e. when  $HU_t^{(i)} > T_b$ ) and the maximum is the cut-off temperature over which heat units are no longer cumulated.

Finally, the crop growth module first estimates the maximum yield achievable in optimal conditions and then reduces it to take into account the stresses due to insufficient water supply from the precipitation and irrigation that occurred during the agricultural season. The yield response to water stresses is estimated according to the empirical func-

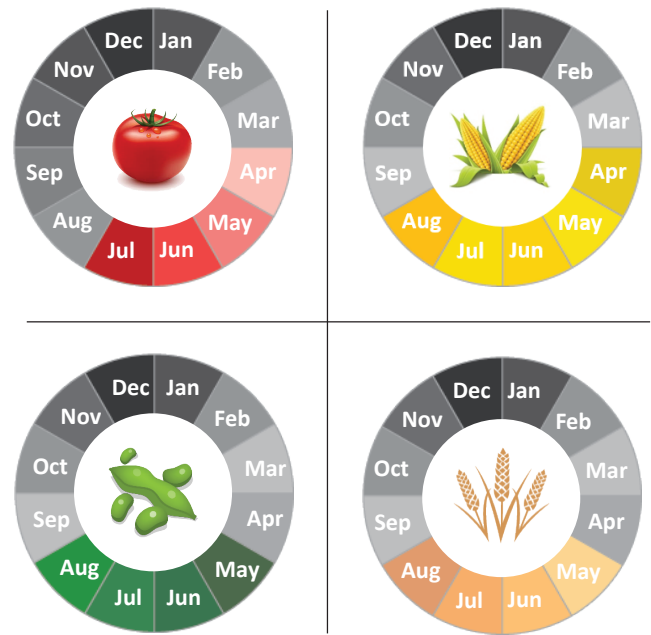


Figure 3. Growing period of the four considered crops.

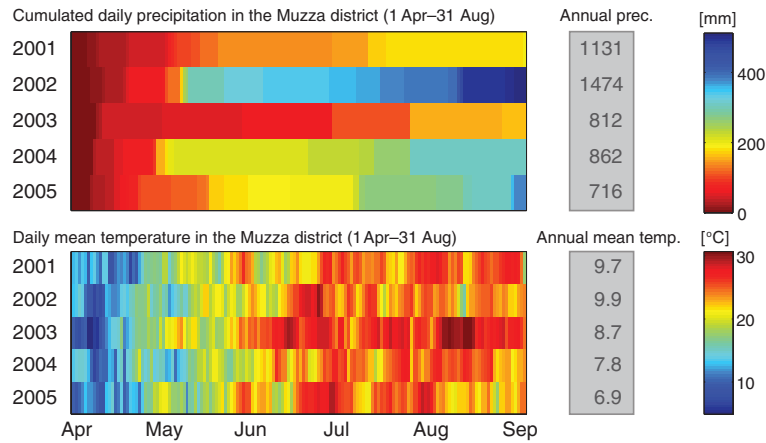
tion proposed in the AquaCrop model (Steduto et al., 2009; Raes et al., 2009) and based on the approach proposed by the FAO (Doorenbos et al., 1979):

$$1 - \frac{Y_{\text{real}}^{(i)}}{Y_{\text{opt}}^{(i)}} = k_y \left( 1 - \frac{\text{Tr}_{\text{real,tot}}^{(i)}}{\text{TrO}_{\text{tot}}^{(i)}} \right), \quad (2)$$

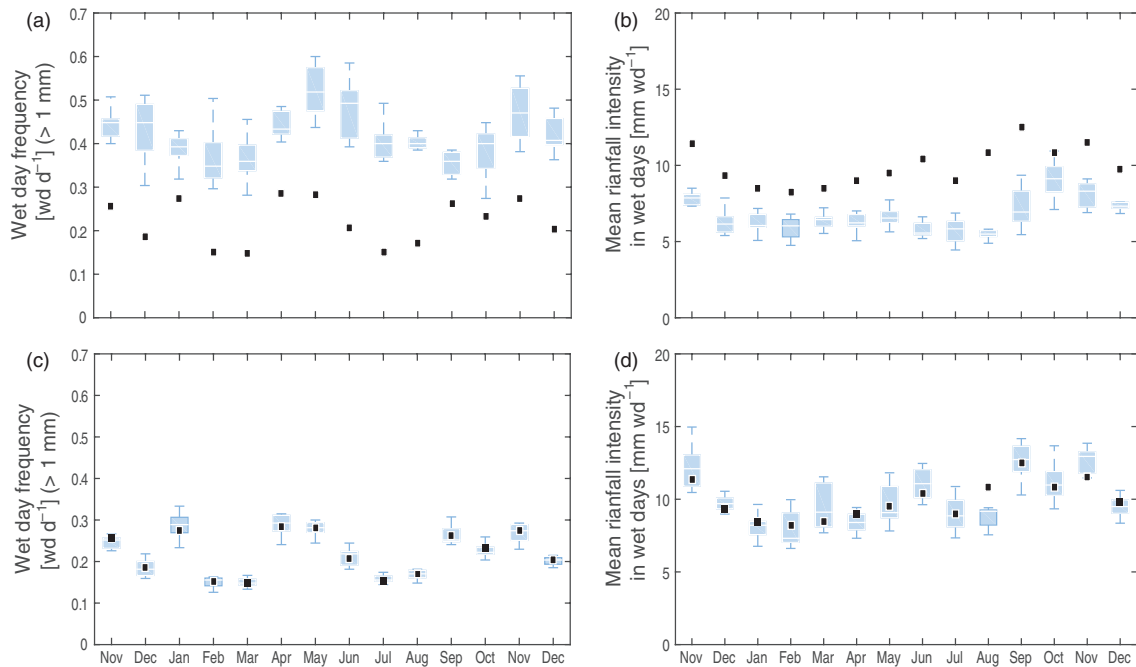
where  $Y_{\text{real}}^{(i)}$  and  $Y_{\text{opt}}^{(i)}$  are the actual and optimal yield in the  $i$ th cell,  $\text{Tr}_{\text{real,tot}}^{(i)}$  and  $\text{TrO}_{\text{tot}}^{(i)}$  the actual and optimal transpiration in the  $i$ th cell during the whole growth period, and  $k_y$  is a crop-specific coefficient relating yield decline and water stress.

### 3.3 The agent-based decision models

In the last step of our procedure (bottom block in Fig. 2), the process-based model described in the previous section is combined with an agent-based model representing the decisions made by the farmers in the 39 irrigation units of the Muzza irrigation district. In particular, each irrigation unit is modelled as a single agent and the decision of each agent is limited to a single crop in each agricultural season. The possible crop choices include tomato, corn, soybean, and rice, which are the most common crops in the considered study area. The crop-growing period slightly varies from one crop to another, with maize being the crop with longest growing period (see Fig. 3). Note that the modelled agents do not represent individual farmers in the system, but rather a group of farmers located in one of the 39 irrigation units. This hypothesis is tantamount to describing the median behaviour of the ensemble of farmers aggregated at the irrigation unit level



**Figure 4.** Observed precipitation and temperature from 2001 to 2005 in the Muzza district. The time series cover only the crop-growing period (April–August). The grey boxes report the annual statistics (i.e. total annual precipitation and annual mean temperature) of each year over the entire Lake Como basin.



**Figure 5.** Comparison between the mean precipitation intensity and wet day frequency for IFS/HOPE model from ECMWF annual product in 2002 before (a, b) and after (c, d) the bias correction.

under the assumption of rational behaviours, and provides a simple and effective way to capture the inter-annual dynamics of land use at the district scale (Giuliani et al., 2016).

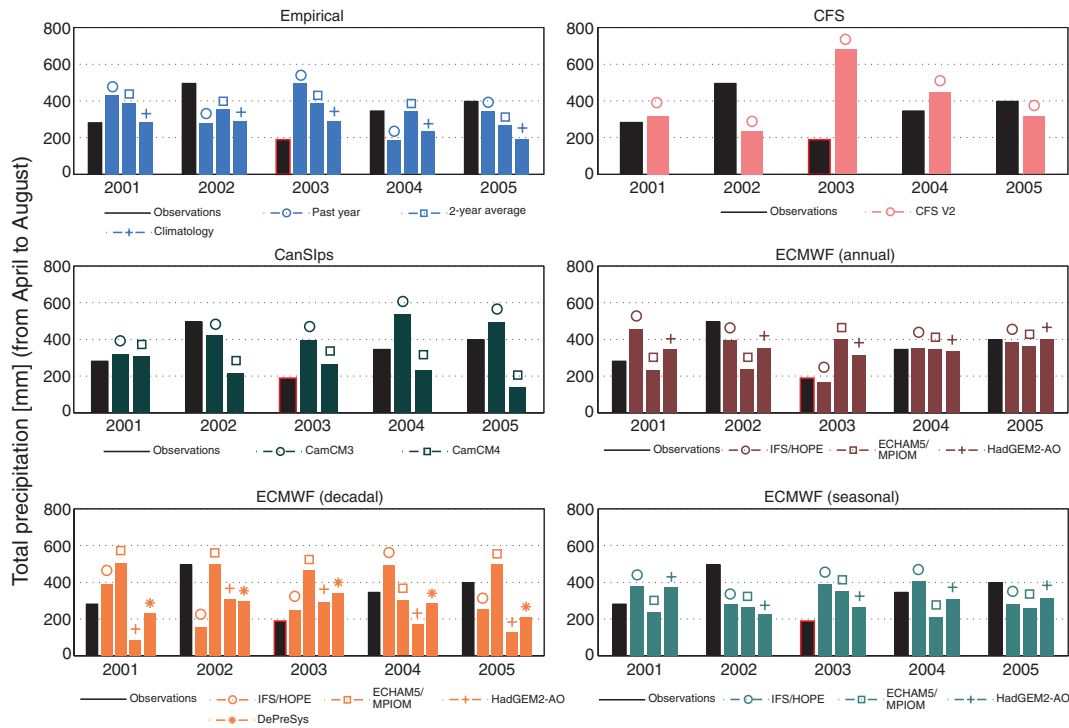
The agent’s decision problem is hence formalized as follows:

$$\gamma_k^* = \arg \max_{\gamma_k} \Psi_{\hat{\varepsilon}} \left[ \mathcal{P} \left( Y_{\text{real}}(\gamma^k), p(\gamma^k), c(\gamma^k), \sigma(A_k) \right) \right]$$

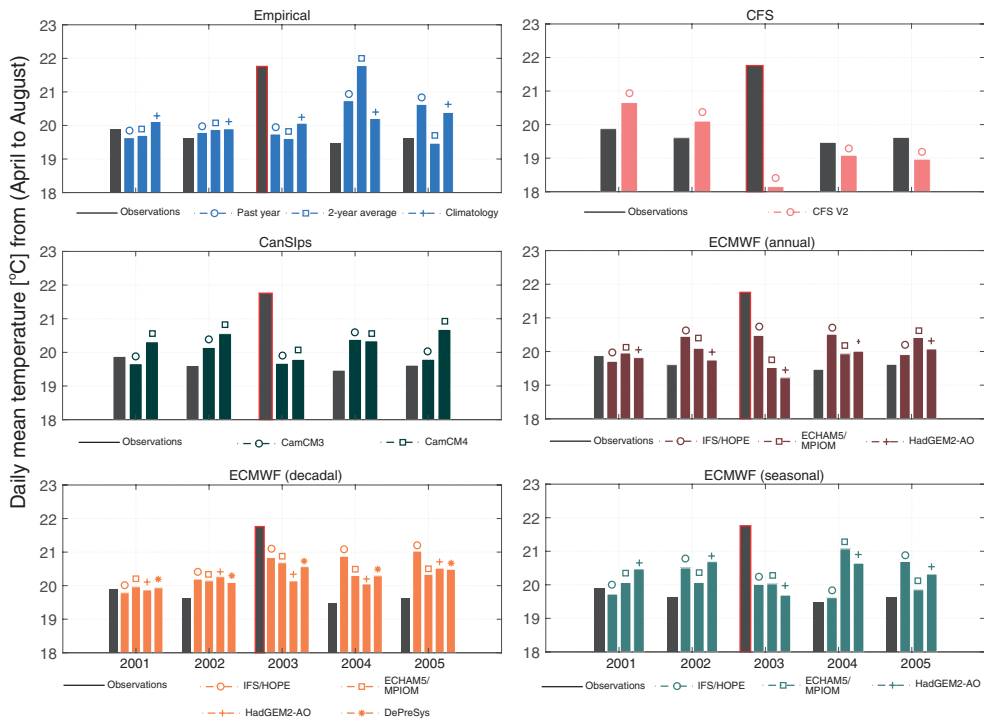
$$k = 1, \dots, 39, \tag{3}$$

where  $\mathcal{P}(\cdot)$  is the net profit obtained at the end of the agricultural season from the yield  $Y_{\text{real}}(\gamma^k)$  of crop  $\gamma^k$  (estimated

from Eq. 2),  $p(\gamma^k)$  and  $c(\gamma^k)$  are the corresponding price and cost, respectively, and  $\sigma(A_k)$  the subsidies for the  $k$ th agent (with  $k = 1, \dots, 39$ ). The subsidies, which depend on the cultivated area  $A_k$  and not on the selected type of crop (Gandolfi et al., 2014), derive from the EU’s Common Agricultural Policy (CAP), which complements a system of direct payments to farmers with measures to help rural areas in facing a wide range of economic, environmental, and social challenges (Britz et al., 2003).

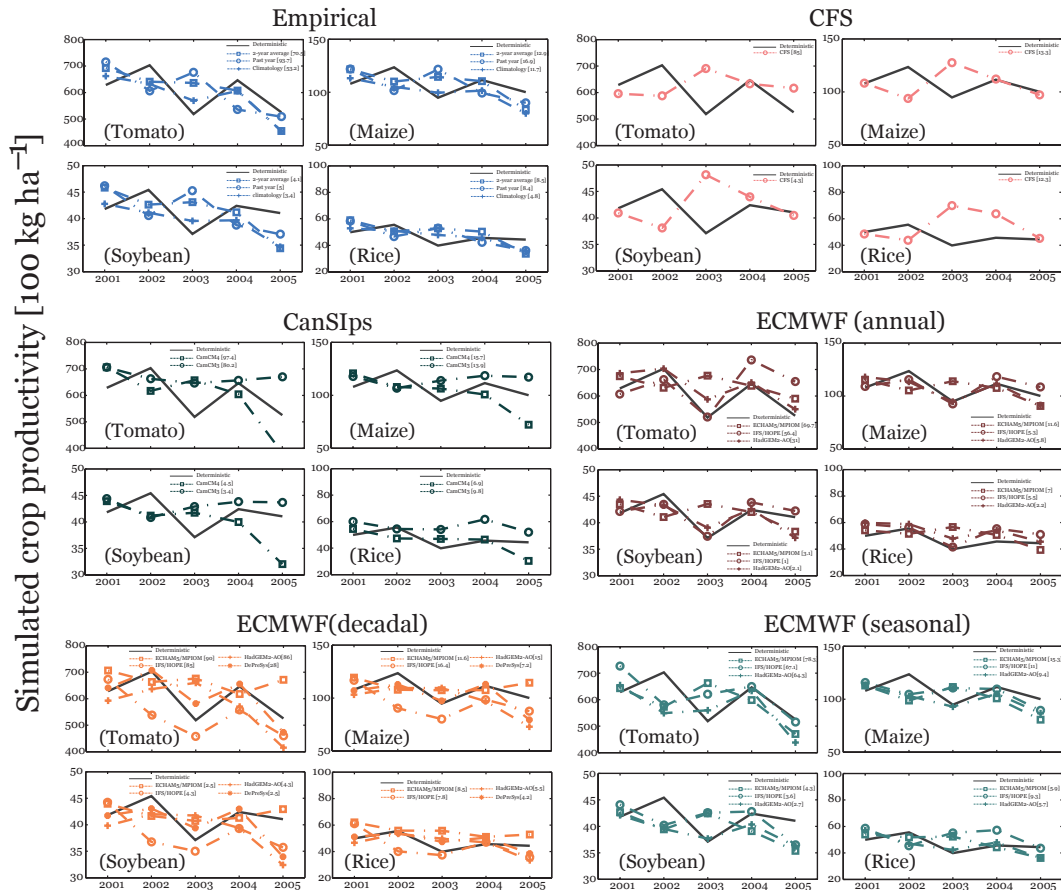


**Figure 6.** Comparison of predicted and observed total precipitation during the crop-growing season across different forecast products. The colour is associated with the family of forecast products and the markers on the top of the bars represent different models within a single family. The red-edge marked bars highlight the drought of 2003.



**Figure 7.** Comparison of predicted and observed daily mean temperature during the crop-growing season across different forecast products. The colour is associated with the family of forecast products and the markers on the top of the bars represent different models within a single family. The red-edge marked bars highlight the drought of 2003.





**Figure 8.** Simulated crop productivity from the period 2001–2005. The arrangement of simulated crop types are tomato (upper left), maize (upper right), soybean (lower left), and rice (lower right). The mean absolute error of each forecast product over the 5 years is reported within square brackets in the legend of each plot.

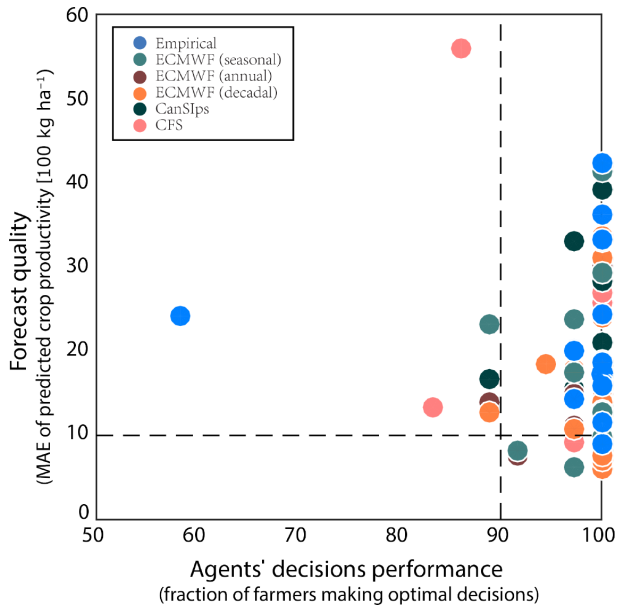
In problem (3), the optimal cropping pattern decision  $\gamma_k^*$  is conditioned on the forecast information  $\hat{\varepsilon}$ , with the statistic  $\Psi$  filtering the uncertainty in the forecast products and capturing the personal risk aversion of each farmer-agent (Giuliani and Castelletti, 2016). In fact, depending on its personal behavioural attitude and on its level of trust in the forecast products, an agent can use the forecast information in different ways, particularly when it is provided in the form of prediction ensembles. In this work, we explore three different levels of farmers’ risk aversion creating a spectrum of behavioural attitudes, namely risk averse, risk neutral, or risk prone (e.g. Rogers, 1975; Mosley and Verschoor, 2005; Koundouri et al., 2006; Djanibekov and Villamor, 2017). A risk-averse, pessimistic behaviour (or a low level of trust in the forecast products) implies that agents decide on the basis of the worst-case realization, which means they will select the cropping patterns able to ensure the highest profit in the most adverse conditions. Yet these decisions may prove to be overly conservative if the actual realization is different from the worst possible one. Conversely, a risk-prone, optimistic behaviour produces decisions that rely on the best possible situation.

This choice increases the chance of cultivating crops that are highly productive under favourable weather conditions, but might also be highly vulnerable under more adverse conditions. Finally, risk-neutral agents with a sufficient level of trust in the forecast products ground their decisions on the expected profitability of the crops using the probability of realizations derived from the forecast information.

These alternative behaviours are formalized by means of the following three statistics  $\Psi$  which are used in Eq. (3) to filter the uncertainty in the forecast products:

- Risk-averse behaviours are modelled using the *minimax regret* metric (Savage, 1951), where decisions are based on the regret, defined as the difference between the performance resulting from the best alternative given that the predicted  $\hat{\varepsilon}_j$  is the true realization of precipitation and temperature and the performance of a given cropping pattern  $\gamma$  under the same weather conditions  $\hat{\varepsilon}_j$ , i.e.

$$r(\gamma, \hat{\varepsilon}_j) = \max_{\gamma'} (\mathcal{P}(\gamma', \hat{\varepsilon}_j) - \mathcal{P}(\gamma, \hat{\varepsilon}_j)). \quad (4)$$



**Figure 9.** Scatterplot of forecast quality of predicted crop productivity ( $y$  axis) and farmers' crop decisions performance ( $x$  axis) under different forecast products.

Then, this metric selects the best cropping pattern  $\gamma^*$  adopting a pessimistic approach, namely by minimizing the maximum regret across all the members of the forecast ensemble  $\hat{\varepsilon} \in \Xi$ , i.e.

$$\gamma^* = \arg \min_{\gamma} \left( \max_{\hat{\varepsilon} \in \Xi} r(\gamma, \hat{\varepsilon}) \right). \quad (5)$$

- Risk-neutral behaviours are modelled using the *principle of insufficient reason* (Laplace, 1951), where decisions are made by assigning equal probability to each forecast ensemble member. Then, the best cropping pattern  $\gamma^*$  is selected as the one associated with the maximum expected performance, i.e.

$$\gamma^* = \arg \max_{\gamma} \left( \frac{1}{n} \sum_{j=1}^n \mathcal{P}_j(\gamma, \hat{\varepsilon}_j) \right), \quad (6)$$

where  $n$  is the number of members in the ensemble.

- Risk-prone behaviours are modelled using the *maximax* metric (French, 1988), where decisions are made by looking at the best possible performance of each decision and selecting the cropping pattern  $\gamma^*$  such that

$$\gamma^* = \arg \max_{\gamma} \left( \max_{\hat{\varepsilon} \in \Xi} \mathcal{P}(\gamma, \hat{\varepsilon}) \right). \quad (7)$$

This metric is generally associated with an optimistic point of view as it assumes that the best state of the world will be realized.

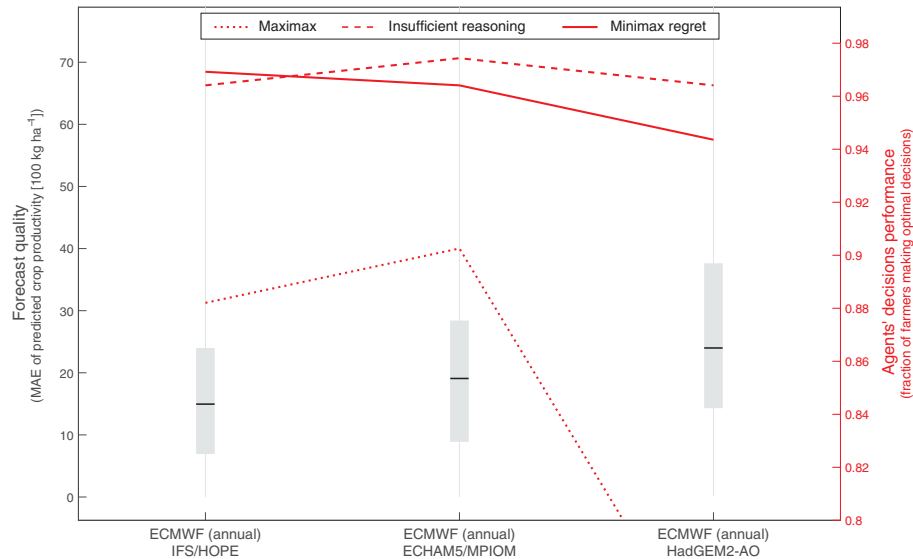
#### 4 Experiment settings

Hindcast of precipitation and surface temperature data are collected from the ECMWF ENSEMBLE project, NCEP, and the Canadian Centre for CCCma, respectively. Table 1 reports some general information about the considered forecast products.

The ECMWF hindcast consists of a comprehensive set of seasonal, annual, and decadal products. The Climate Forecast System version 2 (CFS v2) from NCEP is similar to ECMWF products, generated using fully coupled models representing the interactions between the Earth's atmosphere, oceans, land, and sea-ice (Saha et al., 2014). The Canadian Seasonal to Inter-annual Prediction System (CanSIPS) is a long-range multi-model prediction system whose objective is to forecast the evolution of global climate conditions (Merryfield et al., 2011). There are two versions of coupled climate models inside the CanSIPS system, namely the CamCM3 model (Arora et al., 2011) and the CamCM4 model (Scinocca et al., 2008). To tackle the impact of uncertainties in the initial conditions, most models run a number of simulations with slightly different atmospheric and oceanic initial states to generate ensemble outputs.

In addition to the institutional forecast products, we also include in the analysis three simple empirical models representing farmers' prior knowledge based on past observation. Specifically, EmpPast refers to the empirical forecast obtained by duplicating the past year's observations. The Emp2Ave stands for the simple forecast averaging the past 2 years' observations which is analogous to the climatology forecast with a 2-year memory basis as reflective of farmers' best possible capacity. Lastly, the EmpClima is simply the climatology forecast over past observations.

A total of 13 years of observations (1993–2005) are available, with the last 5 years (i.e. 2001–2005) used for running the model simulations and estimate the operational value of the considered forecast products. This horizon was selected to include a fairly balanced number of normal, wet, and dry agricultural seasons with variable temperature patterns, as shown in Fig. 4. For each simulation, all observations available at the beginning of the year are used for bias correction and downscaling. For example, if the simulation starts in 2003, then the control dataset from 1993 to 2002 is used to calibrate the transfer function and the parameters in the LARS-WG, which will be used to bias-correct the hindcast data in 2003 and generate the time series of precipitation and temperature required by the model. As an illustrative example, Fig. 5 shows the performance of bias correction for the IFS/HOPE model from ECMWF annual product in 2002. The hindcast data (top panels) clearly overestimate the pre-



**Figure 10.** Forecast quality uncertainty of predicted crop productivity (left y axis) and farmers' crop decisions performance for different levels of risk aversion (right y axis).

precipitation frequency and underestimate the precipitation intensity, particularly during the summer. The bias correction successfully solves these issues and the post-processed values match the observations (bottom panels).

## 5 Numerical results

The first step of our framework (Fig. 2) aims at evaluating the forecast quality in terms of the difference between the post-processed forecast variables and the observed ones. Figure 6 shows the post-processed forecast of precipitation against the observed one during the crop-growing season across different forecast products. The empirical memory-based forecast (blue bars in Fig. 6) assumes the weather in the incoming year to be similar to previous conditions. This mechanism leads to significant forecast errors, such as for 2002, which was predicted as a normal year but was wet, or 2003, which was predicted as wet but was extremely dry. The climatology-based forecast assumes the realization of average conditions determined from historical observations. This strategy works in normal years, such as 2001 or 2004, when extreme weather conditions tend to be filtered out during averaging, and year-to-year variations are less significant. For the institutional forecast products, the CFS product seems to work well in normal years while being less accurate in wet and dry years. In particular, it is not able to capture the variation from high to low precipitation in 2002–2003. Similar results can be observed for the Canada CanSIPs products, with CamCM4 generally underestimating the precipitation compared with CamCM3. Estimating the total precipitation for wet/dry years is challenging also for ECMWF products, which involve multiple forecast systems with various lead

times. Nevertheless, there are some exceptions, such as the IFS/HOPE model from among the annual forecast products, which seems to be able to predict quite well the variability from the period 2002–2004. Similarly to the results in Fig. 6, the comparison of the forecast quality evaluated with respect to the daily mean temperature reported in Fig. 7 shows similar patterns among all products, with most of the forecast values close to the observed ones, except for 2003, which was an extremely hot summer due to a European heat wave (García-Herrera et al., 2010).

The second step of our framework (Fig. 2) moves the focus from climatic variables to decision-relevant variables obtained via simulation of the process-based agricultural model. In our problem, this implies looking at the accuracy of estimating the crop yield given the post-processed forecast information. The agricultural model described in Sect. 3.2 provides a mean to transform the climatic variables of interest into the crop yield, which is considered one of the most relevant variables involved in farmers' decision-making process. The comparison between the predicted and the observed (i.e. simulated using the observed weather data for removing possible model biases and focus only on the forecast errors) crop productions is reported in Fig. 8, with the mean absolute error (MAE) of each forecast product over the 5 years reported within square brackets in the legend of each plot. In general, the fluctuations of the production follow the fluctuations of climate variables, especially the precipitation, with the highest productions in 2002 and the lowest in 2003. For most institutional forecast products, the predicted crop productivity in wet/dry years is significantly different from the ones obtained with the observed climate. In many cases, several products tend to overestimate crop

**Table 1.** Summary of the W&C services used in this work. The entries from 1 to 10 are all obtained from ECMWF ENSEMBLE project with different lead time.

Institutes	Model name	Spatial resolution (°)	Temporal resolution	Lead time	Ensemble members	Reference
ECMWF	IFS/HOPE	2.5 × 2.5	daily	7 months	9	Facchi et al. (2004)
ECMWF	ECHAM5/MPIOM	2.5 × 2.5	daily	7 months	9	Roeckner et al. (2003)
ECMWF	HadGEM2-AO	2.5 × 2.5	daily	7 months	9	Martin et al. (2011)
ECMWF	IFS/HOPE	2.5 × 2.5	daily	14 months	9	Facchi et al. (2004)
ECMWF	ECHAM5/MPIOM	2.5 × 2.5	daily	14 months	9	Roeckner et al. (2003)
ECMWF	HadGEM2-AO	2.5 × 2.5	daily	14 months	9	Martin et al. (2011)
ECMWF	IFS/HOPE	2.5 × 2.5	monthly	decadal	3	Facchi et al. (2004)
ECMWF	ECHAM5/MPIOM	2.5 × 2.5	monthly	decadal	3	Roeckner et al. (2003)
ECMWF	HadGEM2-AO	2.5 × 2.5	monthly	decadal	3	Martin et al. (2011)
ECMWF	DePreSys	2.5 × 2.5	monthly	decadal	3	Liu et al. (2012)
NCEP	CFS v2	0.9375 × 0.9375 (i.e. T126 Gaussian)	6-hourly	9 months	4	Kim et al. (2012)
CCCma	CanSIPS CamCM3	2.5 × 2.5	monthly	12 months	10	Kharin et al. (2009)
CCCma	CanSIPS CamCM4	2.5 × 2.5	monthly	12 months	10	Kharin et al. (2009)
Empirical	EmpPast	–	–	–	–	–
Empirical	Emp2Ave	–	–	–	–	–
Empirical	EmpClima	–	–	–	–	–

yield in dry years and to underestimate it in favourable wet years. One exception is again represented by the IFS/HOPE model, which is able to provide quite accurate forecast of crop yield (i.e. average MAE across the four crops equal to 17.1 kg ha<sup>-1</sup>). These results suggest that, as expected, forecasting crop yields is a more complex task than forecasting precipitation and temperature. This is further confirmed by the poor performance attained by the empirical products when their forecast quality is evaluated in terms of crop yield (i.e. average MAE across the four crops and the three empirical forecast products equal to 24.4 kg ha<sup>-1</sup>). Especially for a water-demanding profitable crop like tomato, erroneously forecasting a wet year causes an over-optimistic expectation, which significantly differs from the actual outcome. Similarly, some products (e.g. decadal forecast from ECMWF ECHAM5/MPIOM models) may forecast a wet year which instead turns out to be dry, such as 2005, thus producing a large overestimation of crops' productivity. Finally, these results also show the emergence of some differences in the accuracy of precipitation and temperature forecasts with respect to the corresponding prediction of crop yield. A clear example is 2001, for which CFS V2 exhibits a significantly higher accuracy in predicting the precipitation than the IFS/HOPE model with ECMWF annual product (see Fig. 6). Yet, this superiority does not imply a better forecast of crop production and both the products indeed have similar levels of accuracy across all the four simulated crops (see Fig. 8).

Looking at the accuracy of the predicted precipitation and temperature as well as the predicted crop yields provides a measure of the forecast quality without exploring the potential benefit of using W&C services to inform farmers' decisions. The quantification of the operational value of W&C

services is performed in the third step of our framework (Fig. 2), where we use our agent-based model to simulate farmers' decision-making process and estimate the profit obtained from the cultivation of the selected crops. This is contrasted with the profit obtained under the assumption of perfect foresight to estimate the opportunity cost of using W&C services. It is worth pointing out that, although perfect forecast accuracy can hardly be achieved, farmers' decisions under forecast information may coincide with that selected with perfect foresight. Figure 9 illustrates the relationship between the performance of agents decisions ( $x$  axis), measured in terms of fraction of farmers making optimal decisions (i.e. selecting the same cropping pattern as in the perfect foresight case) and attaining an opportunity cost equal zero, and the associated forecast quality ( $y$  axis), evaluated in terms of MAE of the selected crops. The scatterplot is divided into four zones, where the bottom right corner indicates that a good prediction skill leads to better decision outcomes, while the upper left corner corresponds to the situation where forecast errors induce a large opportunity cost. Both the empirical forecast and the institutional forecast products are spread along the  $y$  axis, confirming the variability of forecast quality in predicting crops' productivity. Numerical results show that most of the points characterized by a good forecast quality, defined as MAE below 1000 kg ha<sup>-1</sup>, correspond to institutional products. These high-quality forecast products provide valuable information to support agents' decisions, as demonstrated by the fact that all the points below the 1000 kg ha<sup>-1</sup> line successfully inform a large fraction of agents (i.e. 90–100%), who are able to make optimal decisions. However, many empirical products are also able to achieve zero opportunity cost, even though their forecast

quality is generally worse than that of institutional forecasts. This can be explained by considering that agents are deciding by looking at the ranking of crops' profitability rather than on absolute expected profitability. As a consequence, an overall under/overestimation of the profitability of all the crops (e.g. the profit of each crop is predicted to be 10 % lower than reality) results in a poor forecast quality but, at the same time, this forecast error does not generate a rank reversal and the agents select the optimal cropping pattern anyway.

## 6 Impacts of farmers' behavioural attitudes

The results presented in the previous section are obtained assuming risk-neutral agents, where the most profitable cropping pattern is selected by the modelled agents on the basis of the crops' profitability predicted by the agricultural model when simulated under a single synthetic time series of post-processed precipitation and temperature (see Fig. 2). Yet, in a more realistic setting, farmers are exposed to uncertain forecasts and, moreover, their behavioural factors may influence the use of W&C services. In this section, we explore how different levels of risk aversion affect the agents' decisions and, consequently, on the estimated operational value of W&C services. In particular, we focus on the ECMWF annual product, which attained both high forecast quality and high operational value, and we generate 100 synthetic time series of precipitation and temperature over the evaluation horizon (2001–2005) by means of the weather generator. These simulations yield 100 uncertain values of crop profitability at the end of the agricultural season. This uncertainty is then filtered by agents through a proper statistic capturing their personal risk aversion, including risk-neutral, risk-prone, and risk-averse behaviours (see Sect. 3.3).

The results obtained by adopting these different levels of risk aversion are reported in Fig. 10, where the left y axis shows the distributions of the forecast quality for the three considered models, while the right y axis shows the fraction of farmer-agents making optimal decisions. The figure shows that, although we are considering a single forecast product, the forecast quality varies according to the model used for producing the forecast, with IFS/HOPE characterized by the lowest MAE, both in terms of median and variance, and outperforming both ECHAM5/MPIOM and HadGEM2-AO. Interestingly, these differences in terms of forecast quality are not linearly transferred to the performance of agents' decisions. Our results show that the level of agents' risk aversion significantly affect their use of forecast products. Risk-averse behaviours (i.e. agents deciding on the basis of minimax regret, represented by the solid red line) attain a performance that decreases when moving from high- to low-quality forecast. However, this does not hold for risk-neutral or risk-prone behaviours, simulated as agents deciding according to the principle of insufficient reason (red dashed line) and the maximax metric (red dotted line), respectively. In both cases,

the highest fraction of agents making optimal decisions is obtained by using the ECHAM5/MPIOM forecast despite this product having a lower quality than IFS/HOPE.

This unexpected finding can be explained by the fact that forecast accuracy metrics quantify the error in predicting the agricultural production, while the operational value estimated through the decision model relies on the ranking of the available options (i.e. cropping patterns). Sub-optimal decisions are made when the forecast productivity of the crops produces a different ranking with respect to the one resulting at the end of the agricultural season. However, such rank reversals are not linearly related to the forecast accuracy: large but consistent (e.g. systematic over/underestimation) errors for all the crops may produce the same ranking and result in optimal decisions, while smaller and variable errors can produce sub-optimal decisions. For example, the values of forecast accuracy reported in Fig. 8 show that in 2001 ECHAM5/MPIOM (which in Fig. 10 attains the highest decision performance) systematically overestimates the productivity of all the crops, while IFS/HOPE underestimates the productivity of tomato and overestimates that of rice, potentially reversing the ranking of these crops and producing sub-optimal decisions.

Finally, it is worth noting that the criterion associated with the largest fraction of agents making optimal decisions, which might be considered as the "best" way for taking advantage of W&C services, varies across the models. The minimax regret is the best when applied to the IFS/HOPE forecast, while the principle of insufficient reason is superior when used for ECHAM5/MPIOM and HadGEM2-AO products. A misdefinition of the stakeholders' perception of W&C services, here explored in terms of risk aversion, may hence represent a strong bias in the analysis of W&C services operational value. For example, the opportunity cost of using ECHAM5/MPIOM simulated assuming the principle of insufficient reason is equal to 3 %, meaning that one agent over the 39 considered in our model is selecting a sub-optimal cropping pattern. The opportunity cost for the same product simulated assuming the minimax regret is instead equal to 4 %, meaning that two agents over 39 select sub-optimal cropping patterns. Finally, the simulation of risk-prone agents adopting the maximax criterion produces an opportunity cost of 10 %, meaning that four agents select sub-optimal cropping patterns. These results provide strong evidence supporting the importance of considering personal, behavioural attributes to produce a proper assessment of W&C services operational value.

## 7 Conclusions

In this work, we propose a novel framework for assessing the operational value of several weather and climate services. This approach, which relies on an integrated model of a coupled human–natural system, is applied in the Muzza irri-

gation district (Italy), a complex agricultural system where farmer-agents select the crops to cultivate by maximizing the expected net profit at the end of the agricultural season. Our framework allows us to quantify the quality of the considered forecast products both in terms of climatic and decision-relevant variables as well as to estimate the associated payoff for the farmers, and also to explore the impact of behavioural attributes on the uptake and use of W&C services.

Our study shows that, at present, the accuracy of most state-of-the-art weather forecast products is still limited, especially in regard to the prediction of precipitation with a lead time of 7 months or longer. The ECMWF annual forecasts simulated by the IFS/HOPE model displayed the maximum forecast skill among the considered products and they were able to also predict some extreme events, including the intense drought of 2003. The predictions of crop yield obtained via simulations of process-based models using the predicted values of precipitation and temperature as climate forcing show similar performance in terms of forecast quality.

Numerical results on the use of these forecast to inform agents' decisions show that the accuracy of estimating crop yield and the probability of making optimal decisions are not necessarily linearly correlated. The assessment of the operational value of W&C services should therefore include a decision model reproducing the actual users' adoption of forecast products within their decision-making process. Some institutional forecast products (e.g. ECMWF products) attain both high forecast quality and high agent decision performance. However, our results also show that in many cases the agents' decisions are still optimal even though informed by products with low forecast quality (e.g. CFS products). From the farmers' point of view, the operational values of ECMWF and CFS products are therefore equivalent despite ECMWF largely outperforming CFS in terms of forecast quality. Finally, we provide numerical evidence of the impact of different farmers' behavioural attributes (i.e. levels of risk aversion) on the quantification of W&C services operational value. The exploration of this behavioural uncertainty further amplifies the key role of the decision model in the assessment procedure. Our results show that the opportunity cost of the same forecast product increases from 3 to 10 % while moving from risk-neutral to risk-prone decisions, potentially producing rank reversals in the quantification of the W&C services' operational value.

To generalize the results obtained in this work, future research efforts should focus on the following directions: extending the evaluation horizon, using large multi-model ensembles, exploring the socio-economic dimension of the problem, and simulating dynamic attitudes of farmers. Our analysis is limited to the time period 2001–2005 because the historical observations available for running the model covers the period 1993–2005, which was divided into two periods, with the first period used for post-processing the forecast products and the second one for performing the analysis.

Moreover, ECMWF forecast products are obtained from the ENSEMBLES project, which provides hindcasts over the period (1960–2005). Although this time period includes a fairly balanced number of normal, wet, and dry agricultural seasons with variable temperature patterns, a longer time horizon including more extreme events would produce more robust findings. The forecast products considered in this work are characterized by a relatively small ensemble size. The use of larger ensembles (possibly multi-model ensembles) has the potential to attain a better performance in terms of forecast quality and, possibly, also in terms of operational value, especially if the analysis is performed on each single ensemble member to better represent the extreme events. However, usage of such large ensembles opens up a number of additional challenges, such as how to limit the smoothing effect on the extreme events or how to combine multiple products with different levels of accuracy, which is beyond the scope of this paper and will be explored in a future analysis. Our model assumes that the predicted water availability is the main factor influencing farmers' decisions, while additional drivers (e.g. expected crop prices, use of nutrients and fertilizers), are assumed as deterministically known. Another next step in our research will be to explore the role of the socio-economic dimension of the problem and its impacts on farmers' decisions. The behavioural attitudes considered in our analysis include diverse levels of farmers' risk aversion with respect to forecast uncertainty. Future research will focus on capturing dynamic behavioural dependences, where the attitude of the farmers in making decisions for the incoming agricultural season is affected by the yield in the previous one. However, the calibration of a decision model implementing such behavioural dependency requires long behavioural time series to identify the proper lag time as well as the magnitude of the effect for different levels of drought intensity.

*Data availability.* The forecast data used in this study are available at the following websites: ECMWF ENSEMBLE project (<http://chfps.cima.fcen.uba.ar/ensemble.html>), NCEP ([http://www.cpc.ncep.noaa.gov/products/CFSv2/CFSv2\\_body.html](http://www.cpc.ncep.noaa.gov/products/CFSv2/CFSv2_body.html)), CanSIPS ([https://weather.gc.ca/grib/grib2\\_cansips\\_e.html](https://weather.gc.ca/grib/grib2_cansips_e.html)).

*Competing interests.* The authors declare that they have no conflict of interest.

*Special issue statement.* This article is part of the special issue "Sub-seasonal to seasonal hydrological forecasting".

*Acknowledgements.* The work has been partially funded by the IMPREX project funded by the European Commission under the Horizon 2020 framework programme (grant no. 641811) and the SOWATCH project funded by Fondazione Cariplo (grant no. 2015-0220). The authors would like to thank Claudio Gandolfi

and his group for their contribution in developing the process-based agricultural model.

Edited by: Quan J. Wang

Reviewed by: three anonymous referees

## References

- Anghileri, D., Castelletti, A., Pianosi, F., Soncini-Sessa, R., and Weber, E.: Optimizing watershed management by coordinated operation of storing facilities, *J. Water Res. Pl.-ASCE*, 139, 492–500, 2013.
- Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V., Lee, W. G., and Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophys. Res. Lett.*, 38, 2011.
- Baigorria, G., Jones, J., and O'Brien, J.: Potential predictability of crop yield using an ensemble climate forecast by a regional circulation model, *Agr. Forest Meteorol.*, 148, 1353–1361, 2008.
- Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, *Nature*, 525, 47–55, 2015.
- Britz, W., Perez, I., and Wieck, C.: Mid-term review proposal impact analysis with the CAPRI modelling system, in: Directorate-General for agriculture mid-term review of the common agricultural policy, proposals impact analyses, European Commission, Brussels, 111–140, 2003.
- Brunet, G., Jones, S., and Ruti, P.: Seamless Prediction of the Earth System: From Minutes to Months, World Meteorological Organization, 2015.
- Calanca, P., Bolius, D., Weigel, A. P., and Liniger, M. A.: Application of long-range weather forecasts to agricultural decision problems in Europe, *J. Agr. Sci.*, 149, 15–22, <https://doi.org/10.1017/S0021859610000729>, 2011.
- Ceballos, A., Martinez-Fernandez, J., and Luengo-Ugidos, M.: Analysis of rainfall trends and dry periods on a pluviometric gradient representative of Mediterranean climate in the Duero Basin, Spain, *J. Arid Environ.*, 58, 215–233, 2004.
- Challinor, A., Slingo, J., Wheeler, T., and Doblaz-Reyes, F.: Probabilistic simulations of crop yield over western India using the DEMETER seasonal hindcast ensembles, *Tellus A*, 57, 498–512, 2005.
- Cloke, H. and Pappenberger, F.: Ensemble flood forecasting: a review, *J. Hydrol.*, 375, 613–626, 2009.
- Cloke, H., Pappenberger, F., Smith, P., and Wetterhall, F.: How do I know if I've improved my continental scale flood early warning system, *Environ. Res. Lett.*, 12, 044006, <https://doi.org/10.1088/1748-9326/aa625a>, 2017.
- Coulibaly, Y., Kundhlande, G., Tall, A., Kaur, H., and Hansen, J.: Which climate services do farmers and pastoralists need in Malawi? Baseline Study for the GFCS Adaptation Program in Africa, CCAFS Working Paper no. 112. CGIAR Research Program on Climate Change, Agriculture and Food Security (CAAFS), 2015.
- Crane, T. A., Roncoli, C., Paz, J., Breuer, N., Broad, K., Ingram, K. T., and Hoogenboom, G.: Forecast Skill and Farmers' Skills: Seasonal Climate Forecasts and Agricultural Risk Management in the Southeastern United States, *Weather, Climate, and Society*, 2, 44–59, <https://doi.org/10.1175/2009WCAS1006.1>, 2010.
- Crochemore, L., Ramos, M.-H., and Pappenberger, F.: Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts, *Hydrol. Earth Syst. Sci.*, 20, 3601–3618, <https://doi.org/10.5194/hess-20-3601-2016>, 2016.
- Dai, A.: Drought under global warming: a review, *Wiley Interdisciplinary Reviews, Climate Change*, 2, 45–65, 2011.
- Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J., and Ramos, M.: Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting, *Environmental Hazards and Risk Communication*, 7, 115127, <https://doi.org/10.1016/j.envhaz.2007.05.001>, 2007.
- Denaro, S., Anghileri, D., Giuliani, M., and Castelletti, A.: Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data, *Adv. Water Resour.*, 103, 51–63, 2017.
- Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, *Glob. Planet. Change*, 57, 16–26, 2007.
- Djanibekov, U. and Villamor, G. B.: Market-based instruments for risk-averse farmers: rubber agroforest conservation in Jambi Province, Indonesia, *Environ. Dev. Econ.*, 22, 133–155, 2017.
- Doblaz-Reyes, F., Garcia-Serrano, J., Lienert, F., Biescas, A., and Rodrigues, L.: Seasonal climate predictability and forecasting: status and prospects, *Wiley Interdisciplinary Reviews, Climate Change*, 4, 245–268, 2013.
- Doorenbos, J., Kassam, A., and Bentvelsen, C.: Yield response to water, irrigation and drainage. Paper no. 33, Tech. Rep., Food and Agriculture Organization, Rome, Italy, 1979.
- Dutra, E., Di Giuseppe, F., Wetterhall, F., and Pappenberger, F.: Seasonal forecasts of droughts in African basins using the Standardized Precipitation Index, *Hydrol. Earth Syst. Sci.*, 17, 2359–2373, <https://doi.org/10.5194/hess-17-2359-2013>, 2013.
- Facchi, A., Ortunani, B., Maggi, D., and Gandolfi, C.: Coupled SVAT-groundwater model for water resources simulation in irrigated alluvial plains, *Environ. Modell. Softw.*, 19, 1053–1063, 2004.
- French, S.: Decision theory: an introduction to the mathematics of rationality, Halsted Press, 1988.
- Frick, J. and Hegg, C.: Can end-users' flood management decision making be improved by information about forecast uncertainty?, *Atmos. Res.*, 100, 296–303, 2011.
- Gandolfi, C., Facchi, A., and Maggi, D.: Comparison of 1d models of water flow in unsaturated soils, *Environ. Modell. Softw.*, 21, 1759–1764, 2006.
- Gandolfi, C., Sali, G., Facchi, A., Tediosi, A., Bulgheroni, C., Rienzner, M., and Weber, E.: Integrated modelling for agricultural policies and water resources planning coordination, *Biosyst. Eng.*, 128, 100–112, 2014.
- García-Herrera, R., Díaz, J., Trigo, R., Luterbacher, J., and Fischer, E.: A review of the European summer heat wave of 2003, *Crit. Rev. Env. Sci. Tec.*, 40, 267–306, 2010.
- Garcia-Morales, M. and Dubus, L.: Forecasting precipitation for hydroelectric power management: how to exploit GCM's seasonal ensemble forecasts, *Int. J. Climatol.*, 27, 1691–1705, 2007.
- GFCS: Implementation Plan of the Global Framework for Climate Services, World Meteorological Organization, 2014.

- Giuliani, M. and Castelletti, A.: Is robustness really robust? How different definitions of robustness impact decision-making under climate change, *Climatic Change*, 135, 409–424, <https://doi.org/10.1007/s10584-015-1586-9>, 2016.
- Giuliani, M., Pianosi, F., and Castelletti, A.: Making the most of data: an information selection and assessment framework to improve water systems operations, *Water Resour. Res.*, 51, 9073–9093, <https://doi.org/10.1002/2015WR017044>, 2015.
- Giuliani, M., Li, Y., Castelletti, A., and Gandolfi, C.: A coupled human-natural systems analysis of irrigated agriculture under changing climate, *Water Resour. Res.*, 52, 6928–6947, <https://doi.org/10.1002/2016WR019363>, 2016.
- Hammer, G., Hansen, J., Phillips, J., Mjelde, J., Hill, H., Love, a., and Potgieter, a.: Advances in application of climate prediction in agriculture, *Agr. Syst.*, 70, 515–553, 2001.
- Hansen, J.: Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges, *Agr. Syst.*, 74, 309–330, 2002.
- Hansen, J.: Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya, *Agr. Forest Meteorol.*, 125, 143–157, 2004.
- Hartmann, H. C., Pagano, T. C., Sorooshian, S., and Bales, R.: Confidence builders: Evaluating seasonal climate forecasts from user perspectives, *Bull. Am. Meteorol. Soc.*, 83, 683–698, 2002.
- Hashmi, M. Z., Shamseldin, A. Y., and Melville, B. W.: Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed, *Stoch. Env. Res. Risk A.*, 25, 475–484, 2011.
- Hourdin, F., Grandpeix, J., Rio, C., Bony, S., Jam, A., Cheruy, F., Rochetin, N., Fairhead, L., Idelkadi, A., Musat, I., Dufresne, J.-L., Lahellec, A., Lefebvre, M.-P., and Roehrig, R.: LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection, *Clim. Dynam.*, 40, 2193–2222, 2013.
- Ines, A. and Hansen, J.: Bias correction of daily GCM rainfall for crop simulation studies, *Agr. Forest Meteorol.*, 138, 44–53, 2006.
- Kharin, V. V., Teng, Q., Zwiers, F. W., Boer, G. J., Derome, J., and Fontecilla, J. S.: Skill assessment of seasonal hindcasts from the Canadian Historical Forecast Project, *Atmos.-Ocean*, 47, 204–223, 2009.
- Kim, H. M., Webster, P. J., and Curry, J. A.: Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter, *Clim. Dynam.*, 39, 2957–2973, 2012.
- Koundouri, P., Nauges, C., and Tzouvelekas, V.: Technology adoption under production uncertainty: theory and application to irrigation technology, *Am. J. Agr. Econ.*, 88, 657–670, 2006.
- Laplace, P.: *A Philosophical Essays on Probabilities*, Dover, 1951.
- Lee, S.-S., Lee, J.-Y., Ha, K.-J., Wang, B., and Schemm, J.: Deficiencies and possibilities for long-lead coupled climate prediction of the Western North Pacific-East Asian summer monsoon, *Clim. Dynam.*, 36, 1173–1188, 2011.
- Lehner, B., Döll, P., Alcamo, J., Henrichs, T., and Kaspar, F.: Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis, *Climatic Change*, 75, 273–299, 2006.
- Liu, C., Haines, K., Iwi, A., and Smith, D.: Comparing the UK Met Office climate prediction system DePreSys with idealized predictability in the HadCM3 model, *Q. J. R. Meteor. Soc.*, 138, 81–90, 2012.
- Liu, J., Dietz, T., Carpenter, S., Alberti, M., Folke, C., Moran, E., Pell, A., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C., Schneider, S., and Taylor, W.: Complexity of coupled human and natural systems, *Science*, 317, 1513–1516, 2007.
- Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L., and Wiltshire, A.: The HadGEM2 family of Met Office Unified Model climate configurations, *Geosci. Model Dev.*, 4, 723–757, <https://doi.org/10.5194/gmd-4-723-2011>, 2011.
- Merryfield, W. J., B. Denis, J.-S. Fontecilla, W.-S. Lee, an J. Hodgson, S., and B. Archambault: The Canadian Seasonal to Interannual Prediction System (CanSIPS), Environment Canada, Development and Operations divisions at CMC and Climate Research division, Tech. Rep., 141, 2910–2945, 2011.
- Mosley, P. and Verschoor, A.: Risk attitudes and the “vicious circle of poverty”, *Eur. J. Dev. Res.*, 17, 59–88, 2005.
- Murphy, A. H.: What is a good forecast? An essay on the nature of goodness in weather forecasting, *Weather Forecast.*, 8, 281–293, 1993.
- Mwangi, E., Wetterhall, F., Dutra, E., Di Giuseppe, F., and Pappenberger, F.: Forecasting droughts in East Africa, *Hydrol. Earth Syst. Sci.*, 18, 611–620, <https://doi.org/10.5194/hess-18-611-2014>, 2014.
- Mylne, K. R.: Decision-making from probability forecasts based on forecast value, *Meteorol. Appl.*, 9, 307–315, 2002.
- Neitsch, S., Arnold, J., Kiniry, J., and Williams, J.: Soil and Water Assessment Tool Theoretical Documentation Version 2009, Tech. Rep. 406, Grassland, Soil and Water Research Laboratory–Agricultural Research Service Blackland Research Center–Texas AgriLife Research, College Station, Texas, 2011.
- Palmer, T. and Hagedorn, R.: *Predictability of weather and climate*, Cambridge University Press, 2006.
- Palmer, T., Shutts, G., Hagedorn, R., D.-R., F., Jung, T., and Leutbecher, M.: Representing model uncertainty in weather and climate prediction, *Annu. Rev. Earth Pl. Sc.*, 33, 163–193, 2005.
- Prodhomme, C., Batté, L., Massonnet, F., Davini, P., Bellprat, O., Guemas, V., and Doblas-Reyes, F.: Benefits of Increasing the Model Resolution for the Seasonal Forecast Quality in EC-Earth, *J. Clim.*, 29, 9141–9162, 2016a.
- Prodhomme, C., Doblas-Reyes, F., Bellprat, O., and Dutra, E.: Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe, *Clim. Dynam.*, 47, 919–935, 2016b.
- Raes, D., Steduto, P., Hsiao, T., and Fereres, E.: AquaCrop–The FAO Crop Model to Simulate Yield Response to Water: II. Main Algorithms and Software Description, *Agron. J.*, 101, 438–447, <https://doi.org/10.2134/agronj2008.0140s>, 2009.



- Ramos, M.-H., Mathevet, T., Thielen, J., and Pappenberger, F.: Communicating uncertainty in hydro-meteorological forecasts: mission impossible?, *Meteorol. Appl.*, 17, 223–235, 2010.
- Ramos, M. H., van Andel, S. J., and Pappenberger, F.: Do probabilistic forecasts lead to better decisions?, *Hydrol. Earth Syst. Sci.*, 17, 2219–2232, <https://doi.org/10.5194/hess-17-2219-2013>, 2013.
- Rayner, S., Lach, D., and Ingram, H.: Weather forecasts are for wimps: why water resource managers do not use climate forecasts, *Climatic Change*, 69, 197–227, 2005.
- Ritchie, J., Zammit, C., and Beal, D.: Can seasonal climate forecasting assist in catchment water management decision-making?, *Agr. Ecosyst. Environ.*, 104, 553–565, 2004.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornbluh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM 5. PART I: Model description, Max Planck Institute for Meteorology, Tech. Rep., 2003.
- Rogers, R. W.: A protection motivation theory of fear appeals and attitude change, *J. Psychol.*, 91, 93–114, 1975.
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P. n., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: The NCEP Climate Forecast System Version 2, *J. Clim.*, 27, 2185–2208, 2014.
- Savage, L.: The theory of statistical decision, *J. Am. Stat. Assoc.*, 46, 55–67, 1951.
- Scinocca, J. F., McFarlane, N. A., Lazare, M., Li, J., and Plummer, D.: Technical Note: The CCCma third generation AGCM and its extension into the middle atmosphere, *Atmos. Chem. Phys.*, 8, 7055–7074, <https://doi.org/10.5194/acp-8-7055-2008>, 2008.
- Semenov, M. A. and Barrow, E. M.: Use of a stochastic weather generator in the development of climate change scenarios, *Climatic Change*, 35, 397–414, 1997.
- Steduto, P., Hsiao, T., Raes, D., and Fereres, E.: AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles, *Agron. J.*, 101, 426–437, <https://doi.org/10.2134/agronj2008.0139s>, 2009.
- Suarez, P. and Patt, A. G.: Cognition, caution, and credibility: the risks of climate forecast application, *Risk, Decision and Policy*, 9, 75–89, <https://doi.org/10.1080/14664530490429968>, 2004.
- Thomson, M., Doblas-Reyes, F., Mason, S., Hagedorn, R., Connor, S., Phindela, T., Morse, A., and Palmer, T.: Malaria early warnings based on seasonal climate forecasts from multi-model ensembles, *Nature*, 439, 576–579, 2006.
- Tippett, M. K., Barnston, A. G., and Robertson, A. W.: Estimation of seasonal precipitation tercile-based categorical probabilities from ensembles, *J. Clim.*, 20, 2210–2228, 2007.
- Vassena, C., Rienzner, M., Ponzini, G., Giudici, M., Gandolfi, C., Durante, C., and Agostani, D.: Modeling water resources of a highly irrigated alluvial plain (Italy): calibrating soil and groundwater models, *Hydrogeol. J.*, 20, 449–467, 2012.
- Wang, D. and Cai, X.: Irrigation scheduling Role of weather forecasting and farmers' behavior, *J. Water Res. Pl.-ASCE*, 135, 364–372, 2009.
- Wetterhall, F., Winsemius, H. C., Dutra, E., Werner, M., and Pappenberger, E.: Seasonal predictions of agro-meteorological drought indicators for the Limpopo basin, *Hydrol. Earth Syst. Sci.*, 19, 2577–2586, <https://doi.org/10.5194/hess-19-2577-2015>, 2015.
- Winsemius, H. C., Dutra, E., Engelbrecht, F. A., Archer Van Garderen, E., Wetterhall, F., Pappenberger, F., and Werner, M. G. F.: The potential value of seasonal forecasts in a changing climate in southern Africa, *Hydrol. Earth Syst. Sci.*, 18, 1525–1538, <https://doi.org/10.5194/hess-18-1525-2014>, 2014.
- Zhu, Y., Toth, Z., Wobus, R., Richardson, D., and Mylne, K.: The economic value of ensemble-based weather forecasts, *Bull. Am. Meteorol. Soc.*, 83, 73–83, 2002.