From skill to value: isolating the influence of end-user behaviour on seasonal forecast assessment

Matteo Giuliani[1], Louise Crochemore[2], Ilias Pechlivanidis[2], and Andrea Castelletti[1]

[1]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Piazza L. da Vinci, 32, I-20133 Milano, Italy

[2]Swedish Meteorological and Hydrological Institute (SMHI), Hydrology Research Unit, Folkborgsvagen 17, SE-601 76, Norrkoping, Sweden

Abstract. Recent improvements in initialization procedures and representation of large scale hydrometeorological processes contributed in advancing the accuracy of hydroclimatic forecasts, which are progressively more skillful over the seasonal and longer timescales. These forecasts are potentially valuable for informing strategic multisector decisions, including irrigated agriculture, where

5   they can improve crop choices and irrigation scheduling. In this operational context, the accuracy associated with the forecast system setup does not necessarily yield proportional marginal benefit, as this is also affected by how forecasts are employed by end-users. This paper contributes a novel framework to quantify the value of hydroclimatic forecasts by extending traditional accuracy assessments with estimates of potential economic benefit to the end-users. We also explore the sensitivity

10 of this benefit to both forecast system setup and end-user behavioral factors. The approach is demonstrated on the Lake Como system (Italy), a regulated lake operated for flood protection and irrigation supply. Our framework relies on an integrated modeling chain composed of three building blocks: bias-adjusted seasonal meteorological forecasts are used as input to the continentally-calibrated E-HYPE hydrological model; predicted lake inflows are used for conditioning the daily lake operations;

15 the resulting lake releases feed an agricultural model to estimate the net profit of the farmers in a downstream irrigation district. Results suggest that despite the gain on average conditions is negligible, during intense drought episodes informing the operations of Lake Como based on seasonal hydrological forecasts allows gaining about 15% of the farmers' profit with respect to a baseline solution not informed by any forecast. Moreover, our analysis suggests that behavioral factors cap-

20 turing different perceptions of risk and uncertainty significantly impact on the quantification of the benefit to the end-users, where the estimated forecast value is potentially undermined by different levels of end-user risk aversion. Lastly, our results show an exponential skill-to-value relation where

large gains in forecast skills are necessary to generate moderate gains in end-user profit, with the ratio that becomes less demanding during extreme drought events.

1 Introduction

Recent advances in initialization procedures (e.g., Ceglar et al., 2018) and representation of large-scale hydro-meteorological processes (e.g., Krysanova et al., 2017) have contributed in greatly advancing the accuracy of hydroclimatic services. State-of-the-art meteorological and hydrological forecast products are increasingly skillful over seasonal and longer time scales and thus become valuable assets for informing strategic decisions contributing to flood protection (e.g., Coughlan de Perez et al., 2017; Neumann et al., 2018), drought management (e.g., Crochemore et al., 2017; Turco et al., 2017), or hydropower production (e.g., Block, 2011; Boucher and Ramos, 2018). Irrigated agriculture is one of the sectors expected to benefit the most from hydroclimatic services to better inform crop choices and irrigation scheduling decisions (e.g., Li et al., 2017; Guimarães Nobre et al., 2019), which strongly depend on the expected hydro-meteorological conditions.

In such operational contexts, forecast accuracy is key to communicate along with hydroclimatic services. Accuracy depends on the forecast system setup, which introduces uncertainties that depend on initial hydro-climatic conditions on the forecast date, scenarios of predicted meteorological conditions (e.g., climate model outputs), and sometimes the adopted impact model. At seasonal time scales, probabilistic forecasts are often used to convey these uncertainties, potentially adding value for decision making (see Georgakakos and Graham, 2008; Cloke and Pappenberger, 2009, and references therein).

The idea of moving from forecast accuracy to value has been explored in a few recent studies that quantify the value generated by informing water system operations with perfect or synthetic forecasts (e.g., Turner et al., 2017; Denaro et al., 2017), or a pre-specified real forecast product (e.g., Anghileri et al., 2016; Nayak et al., 2018), in terms of increased system reliability. Only a few studies (e.g., Li et al., 2017; Delorit and Block, 2019) assess the economic value of existing hydroclimatic services in informing the solution of planning problems, which require making single decisions (e.g., selection of crop to cultivate) without considering how they influence analogous decisions in the future.

However, the value associated with forecast accuracy is often filtered by the way end-users make use of the provided information. Higher forecast accuracy does not necessarily imply better decisions because of the challenges associated to the human interpretation of forecasts as well as to the communication of probabilistic information (Ramos et al., 2010, 2013; Crochemore et al., 2016). The personal interpretation of uncertainty is indeed a subjective process affected by multiple factors, including the way outcomes are framed, the severity of the event being forecasted, and the personal behavioral attitude of the end-users (Gigerenzer et al., 2005; Joslyn et al., 2009). Individual

behaviours and risk perceptions therefore play a key role in influencing the end-user assessment of probabilistic seasonal forecast value (Kirchhoff et al., 2013).

In this paper, we introduce a novel framework that allows (i) the quantification of the value of hydroclimatic services by extending traditional forecast quality assessment methods with estimates of the potential economic benefit of the forecasts in informing operational decisions; (ii) the analysis of the isolated sources of forecast value by assessing the sensitivity of the results on both the model set up and end-user behavioral factors; (iii) the inference of the relation between gains in forecast skill and gains in end-user profit. The approach is demonstrated on the Lake Como system (Italy), a regulated lake primarily operated for flood control and irrigation supply.

Our framework relies on an integrated modeling chain composed of three building blocks: (1) bias-adjusted ECMWF System 4 seasonal meteorological forecasts are used as input to the continentally-calibrated E-HYPE hydrological model (Hundecha et al., 2016); (2) predicted lake inflows are then used for conditioning the daily lake operations; (3) the resulting lake releases finally feed a crop growth model to estimate the forecast value in terms of gain in net profit for the farmers in the downstream irrigation district. In addition, this composite framework is used to isolate the sources of the estimated forecast value by testing alternative forecast products in the first block of the framework and by assuming different interpretations of the probabilistic forecast information by the end-users in the second block, which capture increasing levels of drought risk aversion.

The paper is organized as follows: in the next section we introduce the Lake Como study site, while Section 3 describes the proposed evaluation framework. Results are reported in Section 4, while conclusions and directions for future research are discussed in the last section.

2 Study site

Located in the Italian Alps, the Lake Como basin (Figure 1) is a heavily man-overworked water system, including a large regulated lake (active capacity 247 Mm^3) serving a wide irrigation-fed cultivated area (1,320 km^2), where maize is the most widely grown and productive crop (52% of the area and 1.5 Mton/year). The hydro-meteorological regime is typical of sub-alpine regions, characterized by dry periods in winter and summer, and peaks in late spring and autumn fed by snowmelt and rainfall, respectively. Snowmelt from May to July is the most important contribution to the creation of the seasonal storage, which is then used for irrigation supply in the summer during the peak demand period.

The regulation of the lake has been actively studied since the 1980s (e.g., Guariso et al., 1984, 1986) and is driven by two primary competing objectives: water supply, mainly for irrigation, and flood control in the city of Como, which sits at the lowest point on the lake shoreline. The agricultural districts downstream prefer to store snowmelt in the lake to satisfy the peak summer water demands, when the natural inflow is insufficient to meet irrigation requirements. Yet, storing such
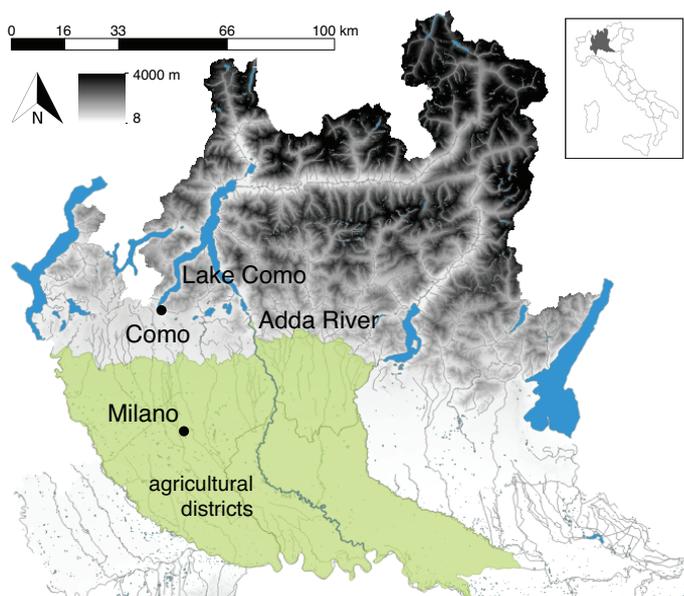
Figure 1. Map of the Lake Como basin. The map was generated via Q-GIS using layers from the Geoportal of Regione Lombardia (www.geoportale.regione.lombardia.it).

water increases the lake level and, consequently, the flood risk. On the basis of previous works (e.g., Castelletti et al., 2010; Giuliani and Castelletti, 2016; Giuliani et al., 2016a; Denaro et al., 2017), these two objectives (both to be minimized) can be formulated as follows:

95   – Flood control (J^F): the average annual number of flooding days in the simulation horizon, defined as days when the lake level is higher than the flooding threshold of 1.24 m;

     – Water supply deficit (J^D): the daily average quadratic water deficit between the lake release and the daily water demand of the downstream system, subject to the minimum environmental flow constraint to ensure adequate environmental conditions in the Adda River.

100   3   Evaluation framework

The overall workflow of our evaluation framework relies on an integrated modeling chain composed of the three building blocks illustrated in Figure 2: (i) the E-HYPE hydrological model produces seasonal forecasts of the Lake Como inflows driven by ECMWF System 4; (ii) the Lake Como operational model designs the optimal lake regulation including the inflow forecasts as additional

105   input in the operating policy that determines the water released by the dam; (iii) the agricultural
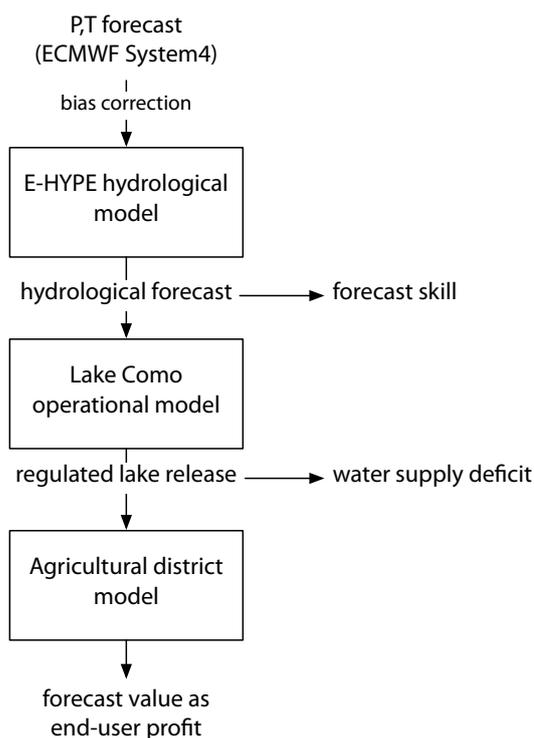
Figure 2. Overview of the integrated modeling chain used in the evaluation framework.

district model estimates the profit of the farmers in the Muzza district, which is the largest among the irrigation districts served solely by the Adda River (about 700 km^2) as well as the one with the largest water concession (2370 Mm^3/yr). A detailed description of each component of the evaluation framework is provided in the next subsections.

3.1 E-HYPE hydrological model

The European setup of the HYPE hydrological model (E-HYPE; Hundecha et al. (2016)) was used to generate dynamical streamflow forecasts. E-HYPE is a process-based model that reproduces streamflow and water balance over the entire European continent. Its parameters were calibrated based on a set of 115 catchments representing the diversity of land-use and soil characteristics, as well as human impacts, and over the 1980–1999 period. The model was validated in about 550 catchments for which streamflow observations are available (see details in Hundecha et al., 2016). Here, precipitation and temperature data from the WFDEI reanalysis (Weedon et al., 2014) were used as reference and streamflow simulations were generated by forcing the E-HYPE model with WFDEI meteoro-

(a) Annual mean cumulated flow

(b) Monthly mean cumulated flow
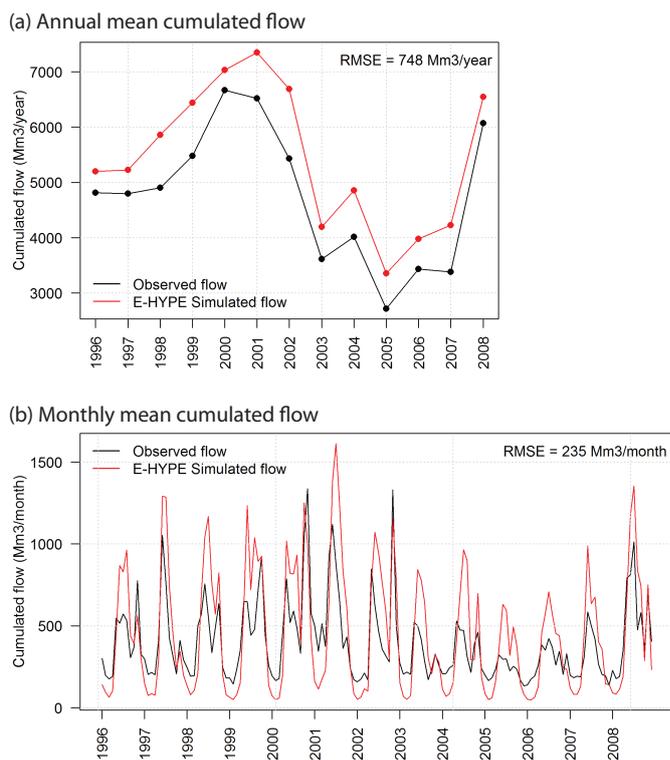
Figure 3. Annual mean accumulated flow (panel a) and monthly mean accumulated flow from observations and E-HYPE simulations from 1996 to 2008 (panel b).

logical inputs. In the Lake Como basin, E-HYPE exhibits good overall performance in simulating

120   yearly streamflow, though a distinct bias can be seen (Figure 3a). E-HYPE achieves an average yearly root-mean squared error (RMSE) of 748 Mm³/year in the Lake Como basin. This yearly performance hides an underestimation of winter flows, and an overestimation of summer flows at the monthly time step (Figure 3b), which is potentially due to an inaccurate representation of snowmelt dynamics in E-HYPE along with the alterations of the natural hydrologic processes introduced by

125   the operations of the Alpine hydropower reservoirs in the upstream part of the basin.

3.2   Operational model of the lake

As mentioned in the previous section, Lake Como is primarily operated looking at two competing objectives, namely water supply and flood control in the city of Como. The operational model of the

lake is focused on reproducing the controlled dynamics of the lake, which is described by a mass
130 balance equation assuming a modeling and decision-making time-step of 24 hours, i.e.

$$s_{t+1} = s_t + q_{t+1} - r_{t+1} \tag{1}$$

where s_t is the lake storage [m^3], while q_{t+1} and r_{t+1} are the net inflow (i.e., inflow minus
evaporation losses) and the outflow volumes in the time interval $[t, t+1)$, respectively. The release
volume r_{t+1} is determined by a nonlinear, stochastic function that depends on the release decision
135 u_t (Soncini-Sessa et al., 2007). This function allows representing the effect of the uncertain inflows
between the time t (at which the decision is taken) and the time $t + 1$ (at which the release is com-
pleted). The actual release might not be equal to the decision due to existing legal and physical
constraints on the reservoir level and release, including spills when the reservoir level exceeds the
maximum capacity.

140 The lake operations is determined by a closed-loop operating policy p that computes the release
decision u_t at each time step t as a function of the day of the year d_t, the lake level h_t and the inflow
forecast $\hat{q}_{t+\tau}$ over the lead time τ. The Pareto optimal operating policies are computed by solving a
multi-objective optimal control problem (Castelletti et al., 2008) formulated as follows:

$$p^* = \arg\min_p \mathbf{J}(p) = |J^F, J^D| \tag{2}$$

145 Note that the resolution of this problem does not yield a unique optimal solution but a set of op-
timal solutions exploring different tradeoffs between flood control and irrigation supply. A solution
is defined as Pareto optimal (or nondominated) if no other solution gives a better value for one ob-
jective without degrading the performance in at least one other objective. The image in the objective
space of the Pareto-optimal solutions is the Pareto front.

150 ### 3.3 Agricultural district model

The agricultural district model simulates the dynamic processes in the Muzza irrigation district. The
model is composed of three distinct modules devoted to specific tasks: (i) a distributed-parameter
water balance module that simulates water sources, conveyance, distribution, and soil-crop water
balance (Facchi et al., 2004); (ii) a heat unit module that computes the sequence of growth stages
155 as a function of the temperature (Neitsch et al., 2011); (iii) a crop yield module that estimates the
optimal and actual yields, accounting for the effects of stresses due to 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 allows
the representation of the space variability of crops, soil types, meteorological inputs, and irrigation
160 distribution. Further details about the different model components are provided in Giuliani et al.
(2016c) and Li et al. (2017).

3.4 Data and Experimental Settings

The assessment of the forecast operational value is performed over the time period from January 1, 1996 to December 31, 2008. This period was selected because it shows good variability in the local hydrological conditions including some intense droughts events that negatively impacted the agricultural production of the system.

For the purpose of this study, we consider two ensemble streamflow forecasts produced by E-HYPE. The first one is named *ESP* (Ensemble Streamflow Prediction; Day (1985)) and is generated by forcing E-HYPE with WFDEI historical scenarios of precipitation and temperature that correspond to the time period of the forecast. The second one is named *SYS4* and uses dynamical precipitation and temperature forecasts from the European Centre for Medium-range Weather Forecasts (Molteni et al., 2011) as input to the E-HYPE model. These forecast inputs are bias adjusted against the WFDEI reference with the Distribution-Based Scaling method (Yang et al., 2010) prior to running the hydrological model. Both ESP and SYS4 forecasts are delivered once a month in the form of a 15-member ensemble with a 7-month lead time. The ensemble means of both ESP and SYS4 are then accumulated over a lead-time of 51 days. This timeframe was demonstrated by Denaro et al. (2017) to be valuable for improving Lake Como operations. In addition to considering the ensemble means, we investigate the sensitivity of the overall assessment framework with respect to end-user behavioral factors. Specifically, we replace the ensemble mean with the 25th and 10th percentiles as well as with the ensemble minimum, which capture increasing levels of drought risk aversion. Lastly, the operational value of these two forecast systems is benchmarked against a set of baseline solutions that rely on the local observed climatology and two sets of upper bound solutions using perfect forecasts corresponding to either E-HYPE simulations forced with meteorological observations or the observed lake inflows.

The comparative analysis of results obtained using different forecast products allows isolating the sources of forecast value as illustrated in Table 1. The sources of forecast value include the initial hydrologic conditions, the hydrologic model, the predictions of precipitation and temperature, and the behavioral factors (i.e., the different percentiles of the forecast ensemble considered). In this matrix, each cell identifies the specific forecasting component that is responsible for the differences in farmers' profit using the forecast system indicated on the columns with respect to the benchmark indicated on the rows.

To optimize the operating policy (see eq. 2), we used the evolutionary multi-objective direct policy search method (Giuliani et al., 2016b), a Reinforcement Learning approach that combines direct policy search, nonlinear approximating networks, and multi-objective evolutionary algorithms. The policies are defined as Gaussian radial basis functions (Busoniu et al., 2011) and the policy parameters are optimized using the self-adaptive Borg MOEA (Hadka and Reed, 2013), a combination that has been demonstrated to be effective in solving these types of multi-objective policy design problems featuring the possibility of enlarging the information used for conditioning operational de-

Table 1. Benchmarking matrix to isolate the sources of forecast value; baseline is observed climatology, ESP E-HYPE Ensemble Streamflow Prediction, SYS4 E-HYPE driven by dynamical precipitation and temperature forecasts, SYS4* replaces the ensemble mean used in SYS4 with different statistics capturing increasing levels of drought risk aversion.

	ESP	**SYS4**	**SYS4***
baseline	hydrological model + initial conditions	hydrological model + initial conditions + P,T forecast	hydrological model + initial conditions + P,T forecast
ESP		P,T forecast	P,T forecast + behavioral factors
SYS4			behavioral factors

cisions (Giuliani et al., 2015; Zatarain-Salazar et al., 2016; Giuliani et al., 2018). Each optimization

200   was run for 2 million function evaluations over the simulation horizon 2007-2015. To improve solution diversity and avoid dependence on randomness, the solution set from each formulation is the result of 20 random optimization trials. The final set of Pareto optimal policies for each experiment is defined as the set of non-dominated solutions from the results of all the optimization trials. In total, the analysis comprises 320 million simulations that required approximately 42,670 computing

205   hours on an Intel Xeon E5-2660 2.20 GHz with 32 processing cores and 96 GB Ram.

4   Results

4.1   Forecast value for irrigated agriculture

Following the proposed evaluation framework (Figure 2), the operational value of alternative forecast systems can be first assessed in terms of reduction of water supply deficit when Lake Como is

210   operated with a policy informed by forecast information. Then, the simulation of the agricultural district model will provide a more tangible measure of the forecast operational value by converting the water supply deficit J^D into monetary values of farmers' profit.

The performance of different sets of solutions obtained by solving the Problem in eq. (2) is shown in Figure 4a, where each circle represents a different operating policy of Lake Como. The two axes

215   of the figure represent the two operating objectives (to be minimized) and the arrows indicate the direction of increasing preference, with the best solution located in the bottom-left corner of the figure. The comparison of the different Pareto-optimal sets shows large differences in performance that determine a clear ranking of the generated solutions. Not surprisingly, the use of perfect forecasts, either in the form of local observations (black circles) or of E-HYPE simulation (blue circles),

220   allow designing (ideal) policies that largely outperform the other solutions. The policies using ESP and SYS4 forecasts are also superior to the baseline solutions, particularly in terms of water supply deficit values. The considered 51-days lead time is indeed too long to provide valuable information
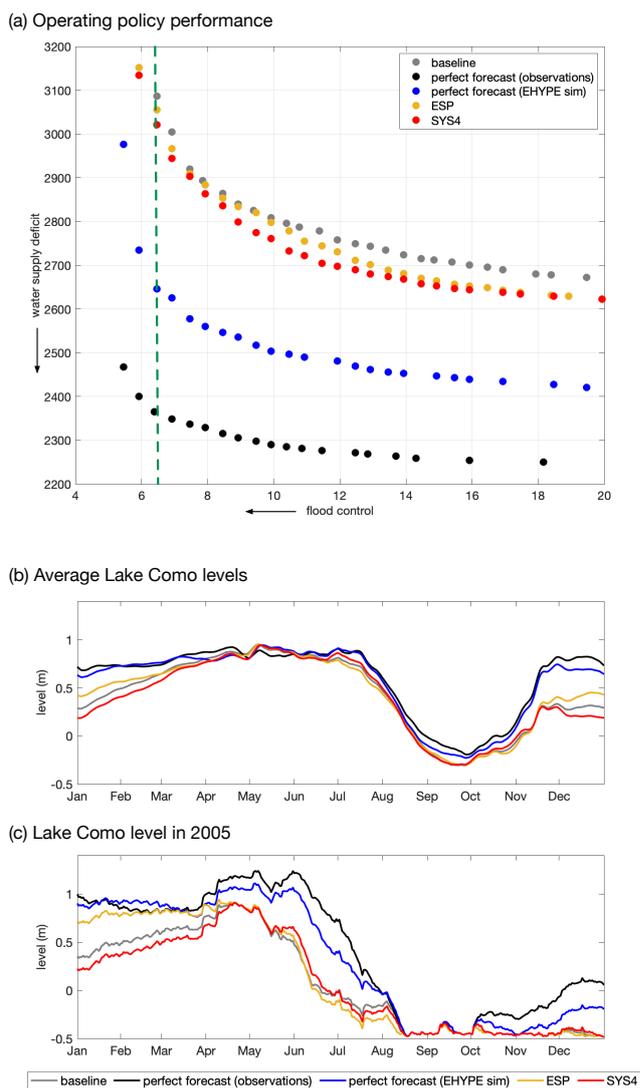
Figure 4. Performance obtained by different Lake Como operating policies (panel a) informed with ESP and SYS4 forecasts, along with the upper bound of the system performance (perfect inflow forecasts from observations or E-HYPE simulation) and the baseline operating policies based on observed climatology. The green dashed line marks the performance of the historical lake regulation in terms of flood control. Analysis of average Lake Como levels (measured with respect to the Malgrate reference level at 197.37 m.a.s.l.) under different operating policies (panel b) and during the extreme drought recorded in 2005 (panel c).

to control the fast flood dynamics, which is on the order of few days and would therefore require much shorter lead times.

225   To better understand the contribution of the different forecast information to the Lake Como operations, we analyze the dynamic behavior of the system under operating policies that use distinct information. This analysis focuses on the solutions located along the green dashed line in Figure 4a, which marks the performance of the historical lake regulation in terms of flood control. The rationale of this choice is to look at solutions that reduce the water supply deficit J^D without degrading the

230   performance in J^F. The historical regulation cannot be used as a reference since it also includes additional objectives not accounted for in our model (e.g., navigation, fishing, tourism). All the simulated trajectories of the Lake Como level under each considered policy show a clear annual pattern, with the highest levels observed in late spring due to the snowmelt contribution (Figure 4b). In this period, maximizing the storage while avoiding floods is crucial to support the summer drawdown cy-

235   cle driven by high irrigation demands. The policies conditioned on perfect forecast (black and blue lines) are able to maintain the highest level and to delay the drawdown. Conversely, the baseline solution (gray line), which has no information about future inflows, reaches the highest level at the beginning of May and, subsequently, the level is maintained about 10 cm below the perfect forecast trajectory to have space for buffering potential floods. A similar trajectory is followed by the policy

240   informed by ESP and SYS4 forecasts (orange and red lines, respectively), with the latter being able to further delay the drawdown to the beginning of July. In addition to the average levels, it is interesting to investigate how the different solutions operate the lake during the extreme drought recorded in 2005 (Figure 4c). The low inflows experienced during this drought event produced an early drawdown of the lake level starting at the beginning of June, when the downstream water demand is at its

245   maximum, with the levels reaching the lower limit of -0.50 m around middle August. This extreme event confirms and emphasizes the differences observed on the average lake levels; the policies conditioned on perfect forecast maintain the highest level from April to mid-August thus delaying the drawdown. ESP and SYS4 forecasts, although less efficient than the perfect forecast solutions, are able to keep higher lake levels than the baseline solution from mid-May to the beginning of July,

250   thus reducing the water supply deficit.

    This analysis can be translated into economic terms via simulation of the Agricultural district model, which estimates the crop production and the associated net profit (i.e., gross revenue minus production costs, also accounting for the EU Common Agricultural Policy subsides (Gandolfi et al., 2014)) for the farmers in the Muzza irrigation district served by the Lake Como releases under

255   different operating policies. Figure 5 shows the same ranking of solutions obtained in the space of the operating objective (Figure 4a), with the use of forecast information that allows gaining, on average, from 1% (ESP forecast) to 3.8% (perfect forecasts from observations) of annual farmers' profit (i.e., from 300,000 €/year to 900,000 €/year) in comparison to the baseline solution. Interestingly, these values are much larger when evaluated over the 2005 drought, when the baseline annual profit is
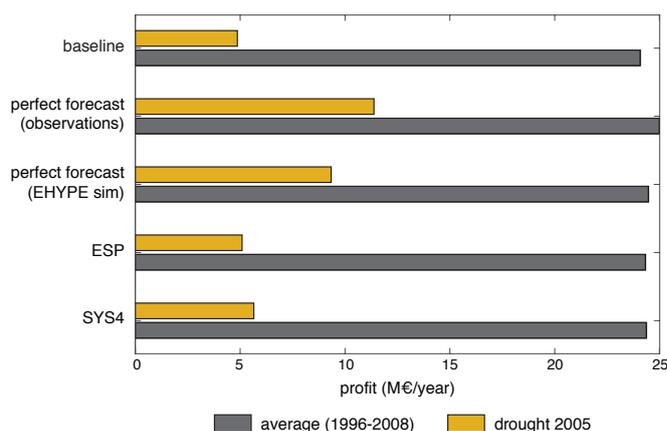
Figure 5. Comparison of farmers' profit under different Lake Como operating policies informed with ESP and SYS4 forecasts, along with the upper bound of the system performance (perfect inflow forecasts from observations or E-HYPE simulation) and the baseline operating policies based on observed climatology.

only 20% of the 1996-2008 average value. In this case, the perfect forecasts generate a profit that is 134% (observations) and 92% (E-HYPE simulation) higher than the baseline; the value of ESP and SYS4 also grows, producing a 5% and 16% increase in farmers' profit, respectively. These results suggest a large potential for using E-HYPE forecast in the management of extreme droughts.

4.2 Impact of forecast system setup and behavioral factors on forecast value

Following the benchmarking analysis in Table 1, we investigate the isolated sources of forecast value by assessing the sensitivity of the farmers' profit on both forecast system set up and end-user behavioral factors. For the former aspect, we compare our baseline solution against the operating policies informed by ESP and SYS4 forecasts (using the ensemble means). For the latter, we explore increasing levels of risk aversion in the use of SYS4 forecasts by informing the operating policy with the 25th and 10th percentiles as well as the minimum of the forecast ensemble.

The results are reported in the comparative matrix in Table 2, which shows again the superiority of ESP and SYS4 over the baseline. Interestingly, the role of predicted precipitation and temperature in drought conditions differs from the average conditions. The use of SYS4 instead of ESP in 2005 generates a 11% gain in farmers' profit, while this difference drops to 0.2% in average conditions. Over the full period, the most important components of the forecast system are the hydrological model and the initial conditions, which together produce more than 1% increase in farmers' profit. Hydrological initial conditions provide the most similar gains between the entire period and the 2005 dry conditions, suggesting that this component is the least sensitive to hydrological conditions. The

Table 2. Results of benchmarking analysis to isolate the sources of forecast value. The matrix reports the percentage change in farmers' profit for the forecast systems on the columns with respect to the benchmarks on the rows, estimated as average over the 1996-2008 period and for the 2005 drought (in parenthesis).

	ESP	SYS4 - mean	SYS4 - min	SYS4 - p10	SYS4 - p25
baseline	1.04 (4.65)	1.21 (16.13)	1.35 (36.26)	1.16 (40.80)	0.90 (22.32)
ESP		0.17 (10.97)	0.31 (30.20)	0.12 (34.54)	-0.14 (16.88)
SYS4 - mean			0.14 (17.34)	-0.05 (21.25)	-0.31 (5.33)
SYS4 - min				-0.19 (3.33)	-0.45 (-10.23)
SYS4 - p10					-0.26 (-13.13)

analysis of the behavioral factors shows that the potential operational value of SYS4 depends on the

280   level of risk aversion used in interpreting the information provided by the forecast ensemble. The average 1.2% increase in farmers' profit with respect to the baseline using the ensemble average grows to 1.35% when the policy is informed by the ensemble minimum, probably because E-HYPE generally overestimates observed inflows (Figure 3a) and predictions of winter low flows are more interesting for managing drought risk. However, results do not demonstrate a linear relationship

285   between forecast value and risk aversion, with the average gain over the baseline being 1.16% when using the 10th percentile of the ensemble (which is equal to the gain produced by the ensemble mean) and 0.9% when using the 25th percentile of the ensemble.

In addition, our results show that the average contribution to the forecast value of predicted precipitation and temperature (+0.12%) is comparable to the one of the isolated behavioral factors. A

290   solution that uses the ensemble minimum produces a profit 0.14% higher than using the ensemble mean (+0.31% with respect to ESP), whereas the 25th percentile of the ensemble generates a 0.31% reduction (-0.14% with respect to ESP). This means that the added value of SYS4 meteorological forecasts can be potentially undermined if end-users are not able to properly extract the most valuable information from the forecast ensemble. However, it should be noted that our results also show

295   that there is not a single best statistic that consistently provides the most valuable information for improving the Lake Como operations. In average conditions, using the ensemble minimum marginally improves the farmers' profit with respect to all the other solutions informed by SYS4 forecasts; conversely, during the 2005 drought, the 10th percentile results to be more valuable than the minimum. The use of risk averse statistics in interpreting the forecast ensemble is therefore recommended for

300   water supply operations exposed to drought risk, but more extensive investigations over multiple extreme events and, possibly, across different case studies is necessary to provide general recommendations.

4.3 From forecast skills to end-user value

Lastly, we aim to identify a relation between the increase in forecast skill and the resulting gain in
farmer profit from isolated forecast system components. The general assumption is that a gain in
forecast performance should result in a gain in profits; however, the gain in farmer profit might be
particular sensitive to having good forecast skill in specific period of the year (see Appendix A for
details about the selected time periods for the computation of the forecast skill). Figure 6 shows that
the most skillful hydrological forecasts are able to provide the maximum higher conversion rate of
skill into end-user value (i.e., farmers' profits), with the overall skill-profit relation well aligned over
an exponential function (i.e., the function $y = 0.184 \cdot e^{0.116x}$, where x is the gain in skill and y the
gain in profit, attains a $R^2 = 0.965$).

Results also show that the correct assimilation of hydrological conditions on the forecast issue date
(i.e., ESP) yields the greatest and only significant gain in skill over the 1996-2008 period (squares).
This 10.7% gain in skill obtained by initializing the hydrological model is associated with a 1.04%
gain in average farmers' profit. SYS4 forecasts yield a 2% gain in skill which leads to a 0.17% gain
in profit. These results suggest a 10 to 1 relation between skill and profit when the entire period is
considered. In this case, the behavioral factors considering low percentiles of the forecast distribu-
tion lead to losses in the skill of the (deterministic) information extracted from the forecast ensemble
(white squares). These forecasts are associated to small losses and gains in profit that are not system-
atic and hardly interpretable, suggesting that risk averse behaviors are likely not relevant in average
hydrological conditions.

In the case of 2005 (circles), behavioral factors yield the greatest gains in skill. Focusing on
the 10th percentile of the forecast distribution yields gains in profit and skill of 21.2% and 40.9%,
respectively, whereas focusing on the minimum of the forecast distribution yields gains of 17.7%
and 39.3%, respectively. In these cases, the skill to profit relation becomes 2 to 1, while this relation
decreases into a 3 to 1 for the SYS4 forecast and into a 6 to 1 for the ESP forecast. Overall, these
results confirm that improving the skills of seasonal forecast is expected to be particularly valuable
to inform the management of extreme events.

5 Conclusions

In this paper we showcase a novel evaluation framework relying on an integrated modeling chain to
quantify the value of hydroclimatic services in terms of added economic benefit of the forecasts in
informing end-user decisions. Moreover, we analyze the isolated sources of forecast value in terms
of both forecast system set up and end-user behavioral factors. The framework is applied to the
operations of Lake Como in the Italian lake district.

Numerical results showcase the potential of the E-HYPE hydrological forecast to inform the op-
erations of Lake Como, generating an average 290,000 €/year gain in the net profit of the farmers
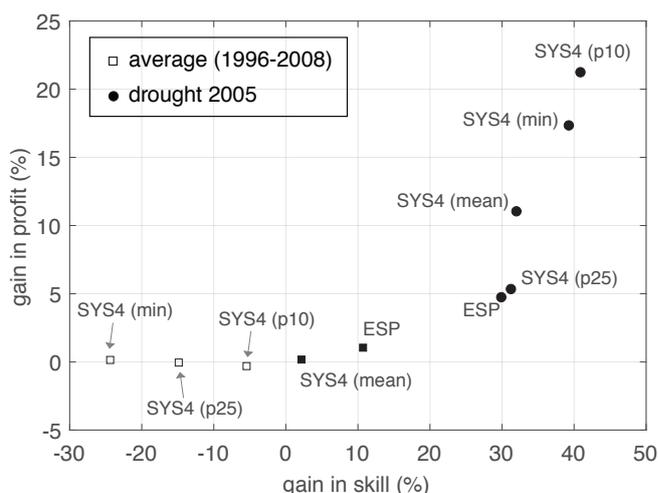
Figure 6. Exponential relationship between forecast skill and operational value. Each marker represents either an isolated component of the forecast system or a behavioral factor evaluated over the 1996-2008 period (circles) or the 2005 drought (squares); filled markers identify a positive gain in forecast skill.

served by the lake releases (about 1% of the average profit obtained by a baseline solution without forecast information). This gain rises up to 16% (i.e., 800,000 € against a baseline profit equal to

340 4.9 M€) during the extreme drought experienced in 2005.

The analysis of the isolated sources of the estimated forecast value attributes the largest share of value to the initialization of the hydrological forecasts with conditions relevant to the forecast issue date. For the extreme drought of 2005, the forecast value is instead mostly attributable to the use of precipitation and temperature predictions and to risk averse decisions focused on the lowest part of

345 the forecast ensemble.

Overall, these results suggest the need of transitioning from forecast skill assessment to integrated frameworks that include decision models and account for end-user behavioral factors capturing different perception of risk and uncertainty. Investing in advanced training for decision makers and reservoir operators is expected to be crucial for maximizing the uptake of forecast information and

350 their operational value (Crochemore et al., 2016). Conversely, the added value of hydroclimatic services might be undermined if end-users are not able to adequately interpret the uncertainty associated to the forecast ensemble. Lastly, our results show an exponential skill-to-value relation where large gains in forecast skills are necessary to generate moderate gains in end-user profit. However, during the 2005 drought, this relationship is less demanding, suggesting that a 10% increase in profit can

355 be obtained with a 30% improvement in forecast skill.

Our study suggests new possibilities for further research directions. To generalize the results, the proposed evaluation framework should be tested over a longer time horizon, possibly including several extreme droughts. Extending the analysis to other irrigated agricultural systems as well as other economic sectors (e.g., hydropower companies, flood protection) is also warranted. Finally, it would
360 be interesting to assess the value of hydroclimatic services under a changing climate characterized by more frequent and intense extreme events that can make forecast information more valuable in the future.

Data availability: The seasonal meteorological forecasts SEAS5 from the European Centre for
365 Medium-Range Weather Forecasts are freely accessible from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). The HYPE model code is available from the HYPEweb portal (http://hypeweb.smhi.se/model-water/). Real-time seasonal forecasts obtained through E-HYPE are openly available on the HYPEweb portal (http://hypeweb.smhi.se/explore-water/forecasts/seasonal-forecasts-europe/). Local observations of lake inflows along with the other meteorological variables
370 used by the agricultural district model were provided by Consorzio dell'Adda (http://www.addaconsorzio.it) and by Agenzia Regionale per la Protezione dell'Ambiente (http://ita.arpalombardia.it).

Acknowledgements. The work has been partially funded by the European Commission under the IMPREX project belonging to Horizon 2020 framework programme (grant n. 641811).

Appendix A: Analysis of sensitive intra-annual periods for forecast skill

375 Despite the general assumption that a gain in forecast performance should result in a gain in profits, when relating profits to performance averaged over all months of the year we observed that a loss in skill sometimes resulted in a gain in profit. This suggests that the profit is sensitive to different periods of the year, with the critical intra-annual period that may vary if we focus on the entire study period (1996-2008) or on dry years such as in the example of 2005. A simple sensitivity analysis
380 was thus carried out to identify the months of the year that explain and impact the calculated profits the most. All possible continuous combinations of months were successively tested to compute the forecast skill, which was then related to the estimated profits. When relating profit and skill over the 1996-2008 period, the profit is mostly related to the skill computed over the April to December period (Figure 7a). When relating profit and skill for 2005, the profit well aligned with the skill
385 averaged from January to September (Figure 7b).

These results are consistent with the strategies adopted in the operation of Lake Como, where the period from April to September corresponds to the agricultural season. Forecasting and managing that period correctly will always play an important role on the yearly profit. In addition, the fall sea-
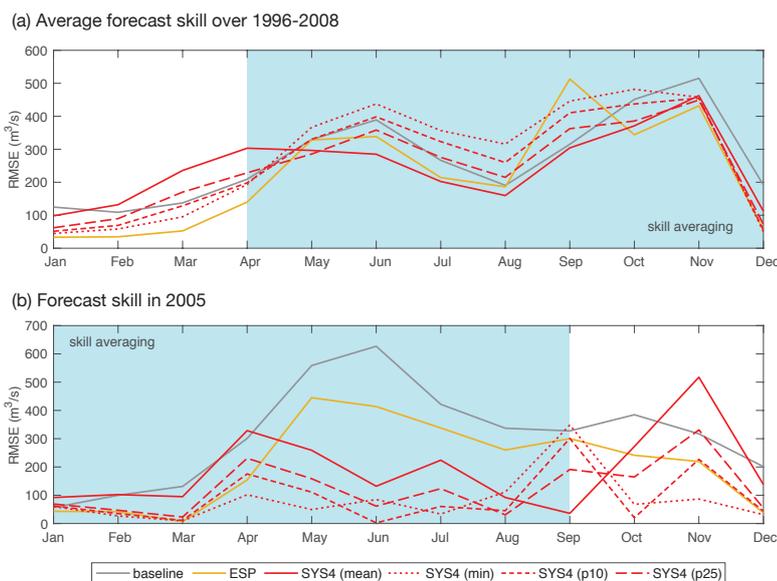
Figure 7. Performance of the forecasts investigated in terms of root mean squared error and over (a) 1996-2008 and (b) 2005. The cyan shaded area represents the time of the year when the skill can be related to the profit evaluated over the corresponding period.

son also plays an important role for the multipurpose operation of the lake, since intense precipitation

390 events cause generate high risk of flooding. Conversely, in dry years, predicting the filling up of the lake at the end of the winter season is more crucial than predicting winter flooding events, since the latter have low probability of occurrence in dry conditions. In the considered 2005 drought, the lake operations benefit from skillful forecast also during the period from January to March (Figure 7b).

References

395 Anghileri, D., Voisin, N., Castelletti, A., Pianosi, F., Nijssen, B., and Lettenmaier, D.: Value of long-term stream-flow forecasts to reservoir operations for water supply in snow-dominated river catchments, Water Resources Research, pp. –, 2016.

Block, P.: Tailoring seasonal climate forecasts for hydropower operations, Hydrology and Earth System Sciences, 15, 1355–1368, doi:10.5194/hess-15-1355-2011, https://www.hydrol-earth-syst-sci.net/15/1355/

400 2011/, 2011.

Boucher, M. and Ramos, M.: Ensemble Streamflow Forecasts for Hydropower Systems, Handbook of Hydrom-eteorological Ensemble Forecasting, pp. 1–19, 2018.

Busoniu, L., Ernst, D., De Schutter, B., and Babuska, R.: Cross–Entropy Optimization of Control Policies With Adaptive Basis Functions, IEEE Transactions on systems, man and cybernetics–Part B: cybernetics,

405 41, 196–209, doi:10.1109/TSMCB.2010.2050586, 2011.

Castelletti, A., Pianosi, F., and Soncini-Sessa, R.: Water reservoir control under economic, social and environ-mental constraints, Automatica, 44, 1595–1607, 2008.

Castelletti, A., Galelli, S., Restelli, M., and Soncini-Sessa, R.: Tree-based reinforcement learning for optimal water reservoir operation, Water Resources Research, 46, 2010.

410 Ceglar, A., Toreti, A., Prodhomme, C., Zampieri, M., Turco, M., and Doblas-Reyes, F.: Land-surface initialisa-tion improves seasonal climate prediction skill for maize yield forecast, Scientific reports, 8, 1322, 2018.

Cloke, H. and Pappenberger, F.: Ensemble flood forecasting: a review, Journal of Hydrology, 375, 613–626, 2009.

Coughlan de Perez, E., Stephens, E., Bischiniotis, K., van Aalst, M., van den Hurk, B., Mason, S., Nissan,

415 H., and Pappenberger, F.: Should seasonal rainfall forecasts be used for flood preparedness?, Hydrology and Earth System Sciences, 21, 4517–4524, 2017.

Crochemore, L., Ramos, M., Pappenberger, F., van Andel, S., and Wood, A.: An experiment on risk-based decision-making in water management using monthly probabilistic forecasts, Bulletin of the American Me-teorological Society, 97, 541–551, 2016.

420 Crochemore, L., Ramos, M., Pappenberger, F., and Perrin, C.: Seasonal streamflow forecasting by conditioning climatology with precipitation indices, Hydrology and Earth System Sciences, 21, 1573–1591, 2017.

Day, G.: Extended streamflow forecasting using NWSRFS, Journal of Water Resources Planning and Manage-ment, 111, 157–170, 1985.

Delorit, J. and Block, P.: Using Seasonal Forecasts to Inform Water Market-Scale Option Contracts, Journal of

425 Water Resources Planning and Management, 145, 04019 018, 2019.

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, Advances in Water Resources, 103, 51–63, doi:10.1016/j.advwatres.2017.02.012, 2017.

Facchi, A., Ortuani, B., Maggi, D., and Gandolfi, C.: Coupled SVAT–groundwater model for water resources

430 simulation in irrigated alluvial plains, Environmental Modeling & Software, 19, 1053–1063, 2004.

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, Biosystems Engineering, 128, 100–112, 2014.

Georgakakos, K. and Graham, N.: Potential benefits of seasonal inflow prediction uncertainty for reservoir release decisions, Journal of Applied Meteorology and Climatology, 47, 1297–1321, 2008.

Gigerenzer, G., Hertwig, R., Van Den Broek, E., Fasolo, B., and Katsikopoulos, K.: "A 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts?, Risk Analysis, 25, 623–629, 2005.

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, 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 Resources Research, 51, 9073–9093, doi:10.1002/2015WR017044, 2015.

Giuliani, M., Castelletti, A., Fedorov, R., and Fraternali, P.: Using crowdsourced web content for informing water systems operations in snow-dominated catchments, Hydrology and Earth System Sciences, 20, 5049–5062, doi:10.5194/hess-20-5049-2016, 2016a.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., and Reed, P.: Curses, tradeoffs, and scalable management: advancing evolutionary multi-objective direct policy search to improve water reservoir operations, Journal of Water Resources Planning and Management, 142, doi:10.1061/(ASCE)WR.1943-5452.0000570, 2016b.

Giuliani, M., Li, Y., Castelletti, A., and Gandolfi, C.: A coupled human-natural systems analysis of irrigated agriculture under changing climate, Water Resources Research, 52, 6928–6947, 2016c.

Giuliani, M., Quinn, J. D., Herman, J. D., Castelletti, A., and Reed, P. M.: Scalable Multiobjective Control for Large-Scale Water Resources Systems Under Uncertainty, IEEE Transactions on Control Systems Technology, 26, 1492–1499, doi:10.1109/TCST.2017.2705162, 2018.

Guariso, G., Orlovski, S., Rinaldi, S., and Soncini-Sessa, R.: An application of the risk-averse approach to the management of Lake Como, Journal of Applied Systems Analysis, 5, 54–64, 1984.

Guariso, G., Rinaldi, S., and Soncini-Sessa, R.: The Management of Lake Como: A Multiobjective Analysis, Water Resources Research, 22, 109–120, doi:10.1029/WR022i002p00109, 1986.

Guimarães Nobre, G., Hunink, J. E., Baruth, B., Aerts, J. C., and Ward, P. J.: Translating large-scale climate variability into crop production forecast in Europe, Scientific reports, 9, 1277, 2019.

Hadka, D. and Reed, P.: Borg: An Auto–Adaptive Many–Objective Evolutionary Computing Framework, Evolutionary Computation, 21, 231–259, 2013.

Hundecha, Y., Arheimer, B., Donnelly, C., and Pechlivanidis, I.: A regional parameter estimation scheme for a pan-European multi-basin model, Journal of Hydrology: Regional Studies, 6, 90–111, 2016.

Joslyn, S., Nadav-Greenberg, L., Taing, M., and Nichols, R.: The effects of wording on the understanding and use of uncertainty information in a threshold forecasting decision, Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23, 55–72, 2009.

Kirchhoff, C., Lemos, M., and Engle, N.: What influences climate information use in water management? The role of boundary organizations and governance regimes in Brazil and the US, Environmental Science & Policy, 26, 6–18, 2013.

Krysanova, V., Vetter, T., Eisner, S., Huang, S., Pechlivanidis, I., Strauch, M., Gelfan, A., Kumar, R., Aich, V., Arheimer, B., Chamorro, A., van Griensven, A., Kundu, D., Lobanova, A., Mishra, V., Plotner, S., Reinhardt, J., Seidou, O., Wang, X., Wortmann, M., Zeng, X., and Hattermann, F.: Intercomparison of regional-scale

hydrological models and climate change impacts projected for 12 large river basins worldwide - a synthesis,
Environmental Research Letters, 12, 105 002, 2017.

Li, Y., Giuliani, M., and Castelletti, A.: A coupled human–natural system to assess the operational value of weather and climate services for agriculture, Hydrology and Earth System Sciences, 21, 4693–4709, doi:10.5194/hess-21-4693-2017, 2017.

Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T., and Vitart, F.: The new ECMWF seasonal forecast system (System 4), vol. 49, European Centre for Medium-Range Weather Forecasts Reading, U. K, 2011.

Nayak, M., Herman, J., and Steinschneider, S.: Balancing Flood Risk and Water Supply in California: Policy Search Integrating Short-Term Forecast Ensembles With Conjunctive Use, Water Resources Research, 54, 7557–7576, doi:10.1029/2018WR023177, 2018.

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.

Neumann, J., Arnal, L., Magnusson, L., and Cloke, H.: The 2013/14 Thames Basin Floods: Do Improved Meteorological Forecasts Lead to More Skillful Hydrological Forecasts at Seasonal Time Scales?, Journal of Hydrometeorology, 19, 1059–1075, 2018.

Ramos, M., Mathevet, T., Thielen, J., and Pappenberger, F.: Communicating uncertainty in hydro-meteorological forecasts: mission impossible?, Meteorological Applications, 17, 223–235, 2010.

Ramos, M., van Andel, S., and Pappenberger, F.: Do probabilistic forecasts lead to better decisions?, Hydrology and Earth System Sciences, 17, 2219–2232, doi:10.5194/hess-17-2219-2013, 2013.

Soncini-Sessa, R., Castelletti, A., and Weber, E.: Integrated and participatory water resources management: Theory, Elsevier, Amsterdam, NL, 2007.

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, Agronomy Journal, 101, 426–437, 2009.

Turco, M., Ceglar, A., Prodhomme, C., Soret, A., Toreti, A., and Doblas-Reyes, F.: Summer drought predictability over Europe: empirical versus dynamical forecasts, Environmental Research Letters, 12, 084 006, 2017.

Turner, S., Bennett, J., Robertson, D., and Galelli, S.: Complex relationship between seasonal streamflow forecast skill and value in reservoir operations, Hydrology and Earth System Sciences, 21, 4841–4859, 2017.

Weedon, G., Balsamo, G., Bellouin, N., Gomes, S., Best, M., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505–7514, 2014.

Yang, W., Andréasson, J., Phil Graham, L., Olsson, J., Rosberg, J., and Wetterhall, F.: Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies, Hydrology Research, 41, 211–229, 2010.

Zatarain-Salazar, J., Reed, P., Herman, J., Giuliani, M., and Castelletti, A.: A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control, Advances in Water Resources, 92, 172–185, doi:10.1016/j.advwatres.2016.04.006, 2016.