FutureWater

Prototype Demonstration and Evaluation

Seasonal Hydrological Forecasting for the Segura River Basin, Spain

REPORT 197

PROJECT IMPREX (H2020)

AUTHORS
Raed Hamed
Alberto de Tomás
Sergio Contreras
Johannes Hunink

DATE December 2019

Seasonal Hydrological Forecasting for the Segura River Basin, Spain

Prototype Demonstration and Evaluation

FutureWater Report 197

Client
European Union H2020 research programme

Project
IMPREX (www.imprex.eu)

Authors
Raed Hamed
Alberto de Tomás
Sergio Contreras
Johannes Hunink - j.hunink@futurewater.es

Date
December 2019

ADDRESS
FutureWater
Calle Comedias 1, 4B
30201 Cartagena
Spain

TELEPHONE +34 968 209 834
WEBSITE www.futurewater.eu

Summary

For the Segura River Basin, a prototype of a drought Decision Support Systems (dDSS) was developed and evaluated. A preliminary design based on the user-requirements is presented of the drought Decision Support System (dDSS). Then an in-depth evaluation of the system performance is presented and a discussion on how the users evaluated the system based on an pilot that was done during three months.

The dDSS of the Segura case study (SE-Spain) targets the River Basin Authority and forecasts several hydrological drought indicators for the Segura river basin, and the connected upper Tagus basin. The dynamic climate model-based seasonal forecasting system was developed and tested comparing it with the currently used forecasting model (a simple statistical approach). The Segura system was tested during winter 2018/2019 in an operational setting.

The developed system has been compared to the already established planning and management operations applied by the river basin authority. The analysis undertaken in this report show that improvements when using dynamic seasonal forecasts are feasible but and not yet impactful enough to entice the user to change current practices. The latter is reinforced by two major factors: (1) current practices are the result of decades of stakeholder processes and institutional settings, and (2) the probabilistic output of seasonal forecasts shows large spread and therefore doesn't convey a simple decision-framework to the end user.

Still, for the first time the basin authority has come to get acquainted with the climate model-based seasonal forecasts and although uncertain, they are now more aware of where the challenges are and have shown to be very willing to further study how the techniques and tools can complement the existing toolset and drought management procedures. They have also gained or increased trust in the partners involved and the science behind. To further build on the awareness and trust that was generated within IMPREX, follow-up work should target an even closer collaboration, for example by having researchers working on-site within the premises of the user (e.g. by secondments), so the tools can be embedded efficiently within the institutional and legislative setting.

Content

Tables

Figures

1 Introduction

The Segura River Basin (SRB) is located in the south-eastern corner of the Iberian Peninsula, bordering the Jucar basin, and has an extension of 20,234 km^2 (Figure 1-1). This semi-arid Mediterranean region is characterized by a mean annual temperature of 18 °C and mean annual precipitation ranging between 300 mm in the downstream areas to 600 mm in the upstream area. Most of the rainfall occurs in a few intensive events that take place in spring and autumn. The temperature conditions make the area suitable for profitable agriculture (fruit trees, horticultural, etc.) despite having the lowest percentage of renewable water resources of Spain and recurrent drought episodes.

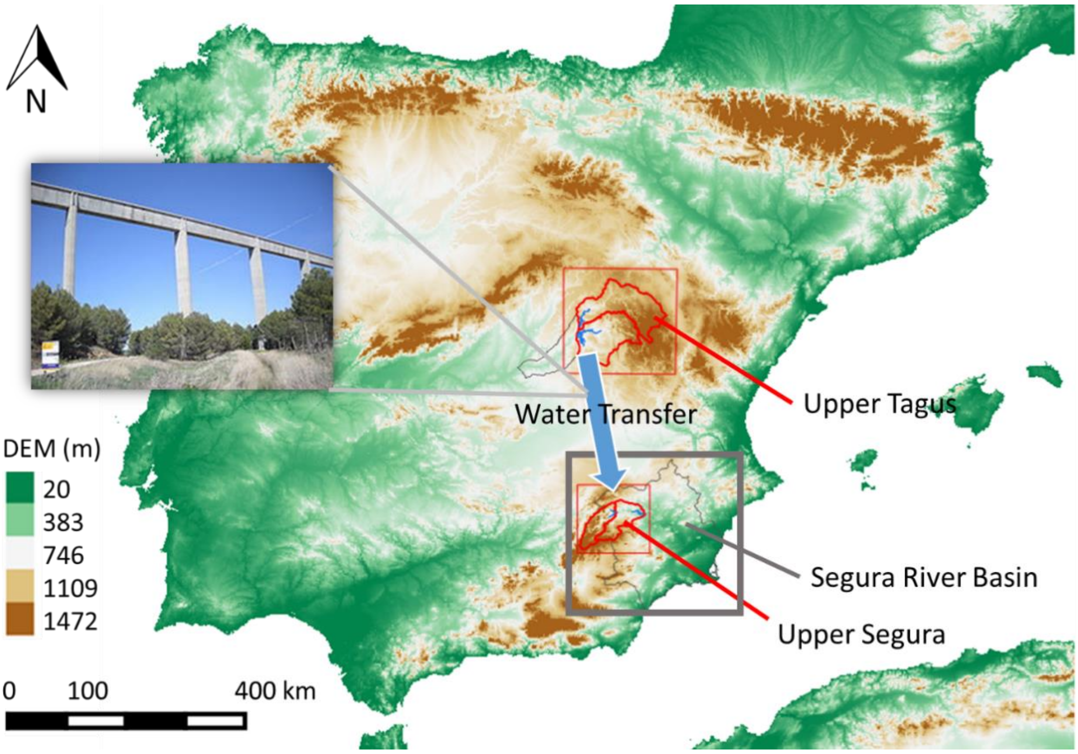


Figure 1-1: Study case basin location indicating the water transfer. Red squares delimit upper Tagus and upper Segura river basins. Grey square delimits the Segura river basin.

The impacts of these events are considerable: for example, the severe drought in 1994 led to an 11-19% reduction of production and a 14% reduction of irrigated area compared to average (CHS, 2007). Production losses due to the lack of irrigation water, in the last year of the extended drought period, amounted to 120 million euros. This reflects the direct economic effects of long-lasting drought periods in the region and therefore the importance of mitigative actions and drought early warning systems to reduce such impacts.

The River Basin Authority of the Segura Basin (Confederación Hidrográfica del Segura – CHS) is responsible for managing most of the water-related infrastructure and distribution of water resources in the basin. During drought periods, CHS must take decisions on the use of alternative water resources (emergency groundwater wells, costly desalination water) and water saving measures. Besides water from the SRB itself, an external transfer provides water for domestic and irrigation in the basin, coming from the Upper Tagus (central Spain): The Tajo-Segura Water Transfer (TSWT) (Figure 1-1) system. The latter connects the Tagus basin in Central Spain to reservoirs in the Segura river basin through a large, ~300 km long infrastructure.

The Drought Management Plan uses drought indices that are based on the resources stored in the reservoirs in the Segura river basin and in the Upper Tagus, to trigger drought mitigation actions. Currently these drought indices are based on observations (reservoir level) only (Table 1 shows drought levels used for the Tajo-Segura transfer). During meetings with CHS it has become clear that so far, the user did not have the resources to explore the potential of integrating forecasts in their decision-making. Also, CHS showed to be quite skeptical about the skill of these type of systems in this basin, given the extremely high inter-annual and intra-annual variability of rainfall and water resources in this basin.

For the connected Upper Tagus-basin however, forecasts are made of reservoir inflows based on a statistical (linear regression) relationship with measured flows in the previous months to estimate flows in the next three months. They use these forecasts to decide on the amount of water that can be transferred to the Segura river basin (see Table 1). Meetings with the main irrigator´s association took place in which they showed interest in having a more skillful forecasting method. Seasonal forecasts based on dynamical modelling approaches have not been used so far in the Segura river basin. The currently used statistical method is considered the benchmark for the Segura river basin. For both river basins, a dynamic hydrological forecasting system was built. For the Upper Taguas, it was compared with the currently used forecasting method (hereafter referred to as “user forecast”).

Table 1. Tajo-Segura Water Transfer exploitation rule.

Levels	Decision rules*	Max. transf. vol. (hm^3/month)
Level 1 – ordinary situation	$V_i > 1500$ hm^3 or $Q_{acc} > 1000$ hm^3	68
Level 2 – abnormal situation	$V_i < 1500$ hm^3 and $Q_{acc} < 1000$ hm3	38
Level 3 – abnormal situation (Council of Ministers approval)	V_i < tabulated values per month that the limiting values	23
Level 4 – absence of surplus	$V_i < 240$ hm^3	0

* Vi = current storage volume; Qacc = three-month accumulated inflows

2 Prototype specifications

2.1 General approach

The dynamic forecasting system was developed in two steps: (1) calibrating a hydrological model with historic data to generate a reference run, and (2) force the model with an ensemble of meteorological hindcasts to analyse performance.

The hydrological component is simulated using the Spatial Processes in Hydrology model (SPHY) (Terink et al., 2015) that runs at a resolution of 5x5 km for the Upper Tagus (UT) basin on a daily timestep. For the reference run, precipitation and temperature values from the Spain02, v4.0 dataset (Herrera et al., 2016) are used for the period of 1980-2010. This simulation creates a best estimate for the spatial variability of hydrological variables which are then used to bias-adjust initial conditions for the hindcasts model run.

The hindcasts consist of seasonal prediction for precipitation and temperature variables taken from the ECMWF's seasonal forecast SEAS5 (ECMWF_SEAS5 hereafter) at daily resolution. These inputs are bias adjusted and downscaled first before being used as an input for the SPHY model. The hindcasts consist of a 25-member ensemble covering the period from 1980 to 2010 and represent 3-month simulations initialized for the month of January, April, July and October (start of the hydrological year). A graphic summary of the methodology is provided in Figure 2-1.

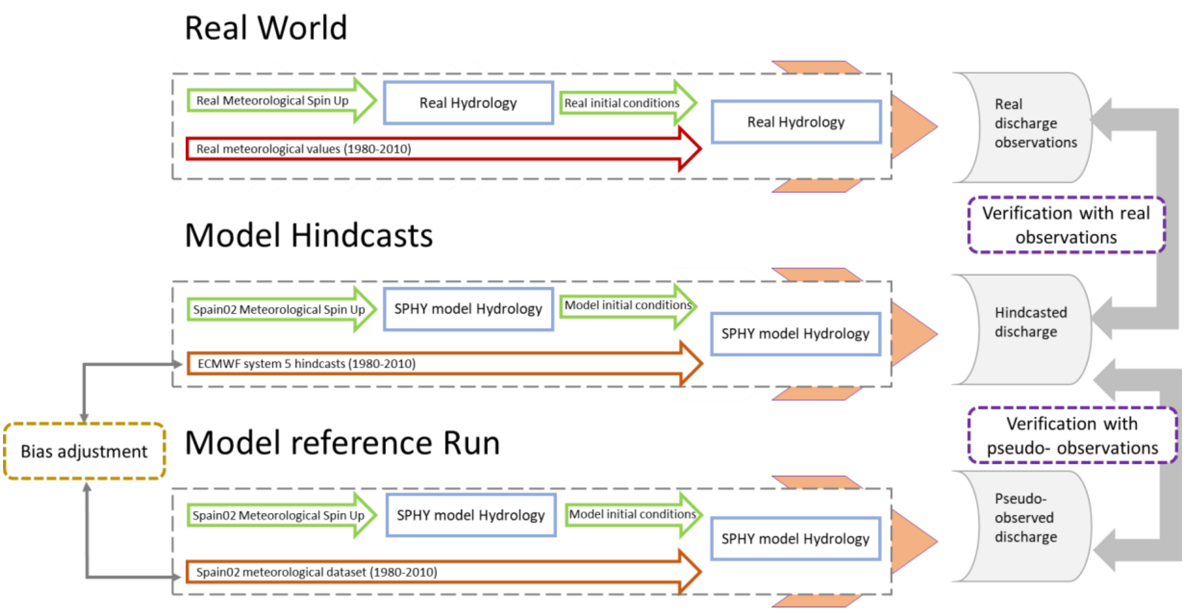


Figure 2-1: Modelling structure for the process-based forecasting system including the "real-world" conditions used for verification. In each box, the flow from left to right represented by green hollow arrows depict the creation of subsequent initial conditions whereas the single red arrow represents the meteorological forcing. Figure adapted from (Greuell et al., 2018)

2.2 Data

2.2.1 Historic meteorological data

The climate dataset called Spain02, v4.0 (Herrera et al., 2016) is a series of high-resolution daily precipitation and (maximum and minimum) temperature gridded datasets developed for peninsular Spain and the Balearic Islands. The dataset uses a dense network of ~2500 quality-controlled stations (~250 for temperatures) and provides information for the period 1971-2010 at a 0.11º (~11km) grid resolution. The dataset is used to force the hydrological model in order to generate the reference simulation in addition to deriving a best estimate for initial conditions used for hindcast simulation. Also, the dataset is used for bias adjustment and downscaling of the hindcasts.

2.2.2 Seasonal forecasts

The ensemble meteorological hindcasts are obtained from the ECMWF SEAS5. Seasonal forecasts are derived from process-based climate models combining ocean and atmosphere dynamics (Weisheimer and Palmer, 2013). Global predictability within these coupled models is most importantly derived from the adequate representation of large-scale phenomena's such as El Nino-Southern Oscillation (ENSO). For example, the latter can offer a predictive signal six months ahead in time in particular areas around the globe (Doblas-Reyes et al., 2013; Jin et al., 2008). Of interest here are a set of hindcasts (re-forecasts) used to run skill diagnosis (1980-2016). These are constituted of a 25-member ensemble forecasts of key variables such as precipitation and temperature at a 36km grid resolution.

2.2.3 Bias adjustment and downscaling of the meteorological hindcasts

Systematic errors in climate models have made the usability of resultant data in hydrological impact assessments challenging (Hagemann et al., 2011). A common methodology is to apply a statistical bias adjustment method to reduce the propagation of errors into the hydrological modelling, herein, influencing the meteorological input to resemble better observed measurements. The adjustment for precipitation is made here with a gridded monthly correction factor, calculated by dividing the monthly climatology of Spain02 with ECMWF-SEAS5 25 ensembles mean. For temperature, the forcing is corrected by applying a lapse rate to the Digital Elevation Model. The data also underwent statistical downscaling, another pre-processing feature in which both the ECMWF-SEAS5 and Spain02, v4.0 datasets were regridded to fit the resolution required for the basin level case study. The interpolation rearranged the data from their initial grid point resolution to a 5 km grid resolution for The Upper Tagus (UT) basin.

2.2.4 Hydrological modelling

The Spatial Processes in Hydrology (SPHY) model (Terink et al., 2015) is grid-based and often used for high-mountain hydrology. For this application, the land component is most important (see Figure 2-2). This is divided in two upper soil stores and a third groundwater store, with their corresponding drainage components: surface runoff, lateral flow and base flow. SPHY simulates for each cell the soil water balance, evapotranspiration, infiltration, and other fluxes. The cell-specific runoff, which becomes available for routing, is the sum of surface runoff, lateral flow, base flow, snow melt and glacier melt.

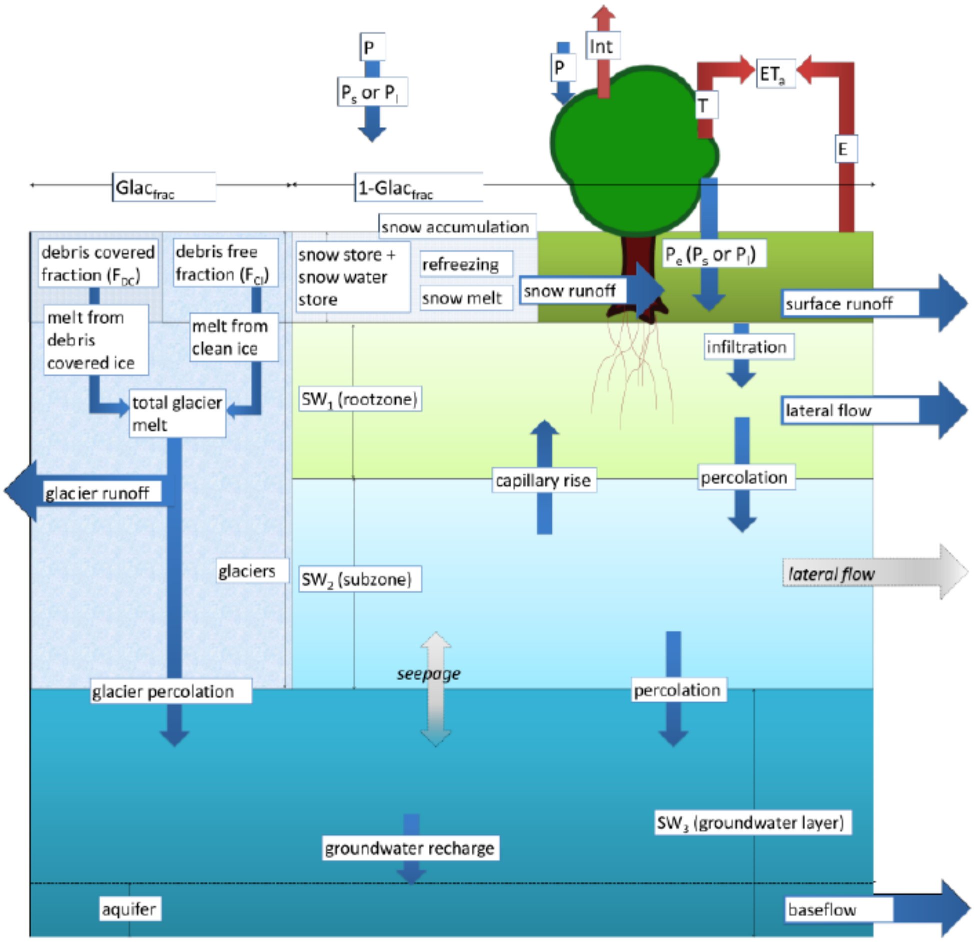


Figure 2-2: Conceptual scheme of the hydrological model SPHY

The SPHY model was calibrated for the 1990-2000 period (using 1988-1989 as spin-up years) against discharge observations for the Segura and Upper Tagus basin. The calibration took place using the Statistical Parameter Optimization Tool for Python - SPOTPY (Houska et al., 2015) and the Maximum Likelihood Estimation algorithm on the most sensitive parameters which were previously diagnosed. RMSE of monthly discharge is chosen as the statistic for the objective function. The following parameters were selected for calibration: caprisemax: initial capillary rise [mm]; deltagw: delay in groundwater recharge [days]; kx: flow recession coefficient or routing coefficient [-]; rootdepthflat: thickness of rootzone [mm]; root_ksat: saturated hydraulic conductivity rootzone [mm/day]; sub_ksat: saturated hydraulic conductivity subsoil [mm/day]. The model performance statistics is assessed using the following metrics: (Root Mean Square Error, Percent Bias, Nash-Sutcliffe Efficiency and the Correlation Coefficient) for both the calibration and validation periods.

2.2.5 Performance indicators for skill diagnosis

The analysis focused in a first step on validating discharge predictions for the user “benchmark” system, as well as for the dynamic system, against observed measurements. Performance was analysed per target season grouping a three-month forecast initialized in January, April, July and October.

A second step focused on validating the drought index of interest to the user comparing values derived from model output and user-forecasts. The multitude of governing attributes qualifying the forecasts implies that no single score measurement on its own would suffice to describe the quality of the hindcasts (Mason and Stephenson, 2008).

The correlation coefficient (CC), the mean absolute error (MAE) and the mean error (ME) was used for deterministic test scores and the continuous Ranked Probability score (CRPS) and the Relative operating curve (ROC) to assess probabilistic test scores. A more detailed documentation of those statistical methods is presented in deliverable 4.2 of IMPREX.

In order to compare forecasts issued by the dynamic seasonal forecasting model to the user benchmark forecasting system, we calculate the skill score to highlight the relevant improvement of a forecast over a considered reference forecast where 0 indicates no improvement and 1 highlights a perfect score.

3 System performance

This section evaluates the performance of the dDSS by using re-forecasts. Besides, also an operational implementation was done with the dDSS using operational forecasts, which was evaluated with the user for the 2018/2019 winter period as a 3-month pilot (section 3.4).

3.1 Bias adjustment of meteorological forecasts

Figure 3-1 shows bias adjustment results for the Upper Tagus basin. The mean bias adjustment ratio was of 1.1 for the Upper Tagus. There is a general underestimation of precipitation when comparing the ECMWF SEAS5 ensemble mean with Spain02. This underestimation is especially large in October, November and December. On the contrary, an overestimation is observed for March and June.

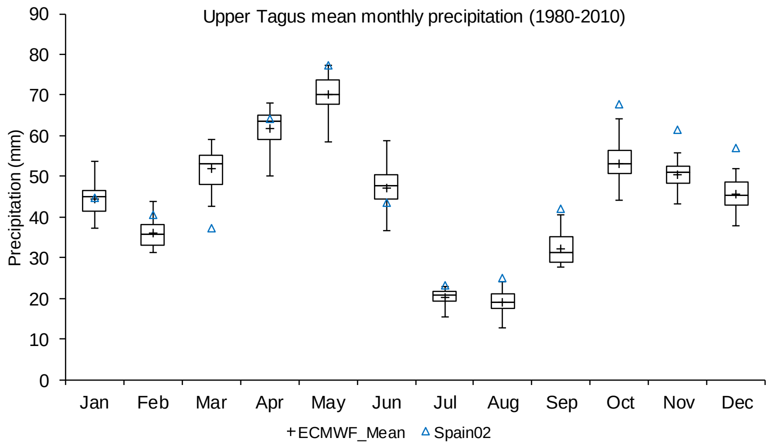
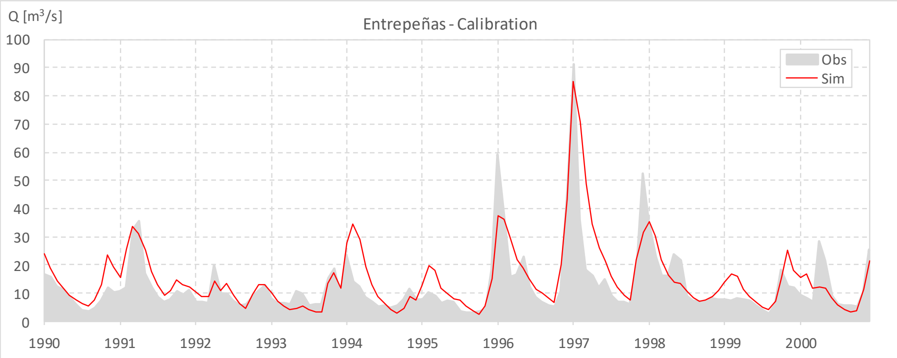


Figure 3-1: Upper Tagus monthly precipitation climatology (period 1980-2010). Boxplots represent ECMWF-SEAS5 25 ensembles mean, while triangles represent the Spain02 observational reference.

3.2 Calibration and verification of the hydrological model

In addition to the performance statistics for both the calibration (1990-2000) and validation (2001-2010) periods, hydrographs for the Entrepeñas of the Upper Tagus basin are shown in Figure 3-2. For other stations, similar outputs were generated (not shown here).

Table 2 provides a summary of the performance statistics. Overall, given the values obtained for the calibration performance statistics, the calibration and validation can be considered satisfactory, although certainly also showing that there is scope for improvement, most likely related to forcing (rainfall) and limitations in the model structure, for example due to the complex and unknown groundwater dynamics in the upper Jucar and Segura.



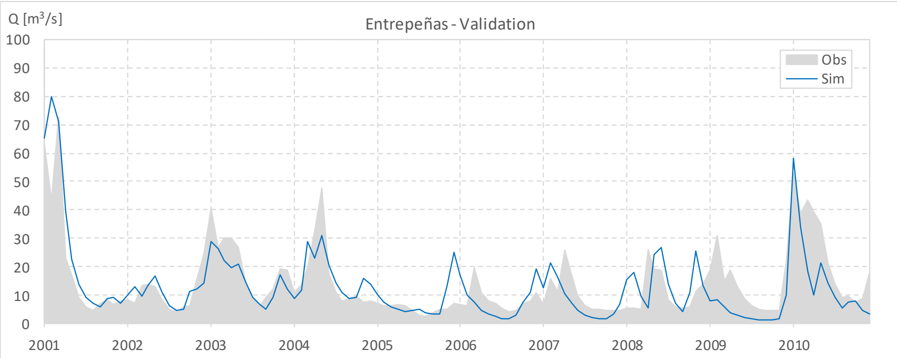


Figure 3-2: Hydrographs for station 1 (Entrepeñas), calibration (1990-2000) and validation (2001-2010) periods.

Table 2: SPHY model performance for two Upper Tagus stations for the calibration (1980-2000) and validation periods (2001-2010).

	Station 1 Upper Tagus		Station 2 Upper Tagus	
	Calibration	Validation	Calibration	Validation
RMSE (m^3/s)	7.2	8.5	6.5	7.8
PBIAS (%)	4.8	-8.9	16.5	-10.0
NSE	0.6	0.6	0.6	0.3
CC	0.8	0.8	0.8	0.6

3.3 Forecasting system performance

In this section, discharge output for the reference run, the SPHY-ECMWF coupled model system and the user forecast are compared to observed measurements. Forecasts initialized at the beginning of January, April, July and October are considered with a forecast lead time of three months. These are grouped as average values into 4 categories highlighting different periods throughout the year: (January-February-March) JFM, (April-May-June) AMJ, (July-August-September) JJA and (October-November-December) OND. The analysis covers the period from 1982 to 2010 where data is available consistently for all the different setups considered.

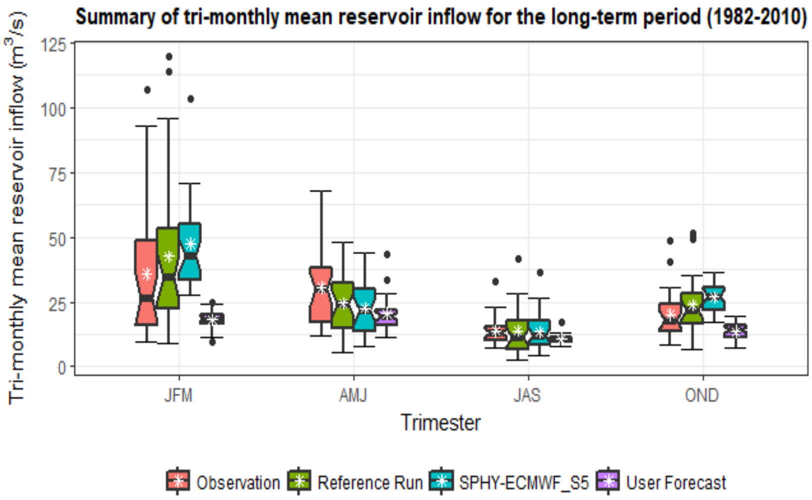


Figure 3-3: Comparison of SPHY-ECMWF_SEAS5 mean output, the reference run, the user forecast and the observation long term (1982-2010) 3-month average streamflow. The white star represents mean values. Distribution of the timeseries are represented for the four periods as box-and-whisker plots.

Figure 3-3 illustrates the different simulated and observed long term 3-month streamflow averages for the period 1982-2010. The distribution of each timeseries within the trimester are represented as box-and-whisker plots where the boxes refer to the 25-75% inter-quantile range and the band inside the boxes to the median value. The whiskers represent 1.5 times the inter-quantile range and values beyond are plotted as single data points. The reference run distribution and mean value indicates a minor overestimation for JFM, JAS and OND periods and an underestimation for the AMJ period compared to observations. The SPHY-ECMWF_SEAS5 mean ensemble distribution and mean values tend to resemble very much the one resulting from the reference run, showing similar albeit more pronounced overestimation/underestimation behaviour compared to observations. The user forecast mean values and distributions show reduced variability along the years and among trimesters compared to the other methods. Streamflow values in that case and for all periods are underestimated compared to observations.

Table 3 reports on correlation coefficients for user forecast and mean SPHY-ECMWF SEAS5 output against observation data. The results show a better performance for SPHY-ECMWF SEAS5 data in correlating with observation for JFM, JAS and OND periods. The higher correlation during seasons of larger meteorological variability indicates an added value of seasonal meteorological forecasts during those periods. Assessments such as ESP and reverse ESP experiments allow detection of sources of predictability, but were not performed for this study. Nevertheless, the sensitivity analysis on sources of skill in Europe developed in deliverable 4.2 of IMPREX showed that seasonal meteorological forecasts provide the dominant source of predictability for the area of study over considered periods and lead times.

Table 3: Pearson correlation coefficients for streamflow forecasts against observation data with associated P-values (ns ≥ 0.1, * < 0.1, ** ≤ 0.01, * ≤ 0.001, **** ≤ 0.0001)**

CC	User Forecast	SPHY-ECMWF_SEAS5
JFM	0.33*	0.65***
AMJ	0.34*	0.34*
JAS	0.52**	0.75****
OND	0.15 **ns**	0.33*

Table 4 differentiates between real and theoretical skill. Real skill is measured by the correlation coefficient between the SPHY-ECMWF SEAS5 mean output and observations, whereas theoretical skill measures the performance relative to the reference run. This computation is usually adopted to mask the error coming from the hydrological model itself when calculating levels of skill (Greuell et al., 2018). The ratio illustrates the level of agreement between those two setups. Degradation of skill for this case study is reported to be largely dependent on the season with very little differences noted for JFM and JAS periods whereas mild ones for OND and particularly large ones for AMJ. The real skill is lowest in spring and autumn which are the rainiest seasons: most likely due to a combination of data deficiencies, especially in rainfall data as well as the complex system response not fully captured by the hydrological model.

Table 4: Comparing Correlation coefficients between SPHY-ECMWF SEAS5 mean verified with observations (Real Skill) and Reference run (Theoretical Skill)

CC	Real Skill	Theoretical Skill	Ratio
JFM	0.65	0.72	0.90
AMJ	0.34	0.83	0.41
JAS	0.75	0.97	0.77
OND	0.33	0.5	0.66

Table 5 highlights the forecast bias, the mean error between observations and the two forecasting systems. The user forecast underestimates discharge for all periods, in particular for JFM and AMJ periods. The model-based system SPHY-ECMWF_SEAS5 reports lower bias for all considered seasons in particular JAS although that period is known to be extremely dry with very low meteorological and hydrological activity. Positive bias is reported in JFM and OND, which is possibly due to the hydrological model underestimating groundwater storage in addition to the ECMWF data overestimating precipitation levels as reported in Figure 3-1 and Figure 3-3.

Table 5: Mean error between observations and streamflow forecasts for the considered periods.

ME	SPHY-ECMWF_SEAS5_Mean	User Forecast
JFM	11.7	-17.7
AMJ	-7.9	-10.0
JAS	-0.4	-3.2
OND	6.6	-6.6

The CRPS is reduced to MAE if the forecast is deterministic. This makes it possible to compare the probabilistic forecast of SPHY-ECMWF SEAS5 to the user forecast both verified against observations in Table 6. Errors among forecasting systems tend to be similar and of the same magnitude with largest differences in JFM season and smallest in JAS season. The latter can be explained by the particularities of seasonality as low hydrological and meteorological activity is experienced during JAS and the opposite is true for JFM. Computing the skill score based on MAE/CRPS in Table 7 shows that differences are minor. The user forecast performs slightly better during the dry season JAS whereas the probabilistic forecast provides better predictability for JFM and AMJ.

Table 6: MAE and CRPS between observations and user forecast in addition to both SPHY-ECMWFSEAS5 deterministic and probabilistic form.

MAE/CRPS	SPHY-ECMWF SEAS5_Mean	User Forecast	SPHY-ECMWF SEAS5_Prob
JFM	18.7	20.0	18.4
AMJ	11.9	12.4	8.9
JAS	3.5	3.7	4.2
OND	9.1	7.9	8.2

Table 7: Skill score based on MAE/CRPS for both deterministic and probabilistic SPHY-ECMWFSEAS5 coupled systems using the user forecast as reference

Skill score based on MAE/CRPS	SPHY-ECMWF SEAS5_Mean	SPHY-ECMWF SEAS5_Ens
JFM	0.05	0.1
AMJ	0	0.26
JAS	0.054	-0.13
OND	-0.15	-0.037

3.4 Forecasting the drought index for Tajo-Segura Water transfer system

The Tajo-Segura water transfer system is managed by the Segura river basin authority and is governed by a drought index derived from discharge. The value of the drought index categorises the drought risk: No drought, Pre-Alert, Alert and Emergency, upon which different operational decisions are made. For that purpose, userthe drought index based on the user forecast model (currently used as input for the water transfer system) is compared to output from the coupled SPHY-ECMWF forecasting system. Results are shown in Figure 3-4 where both forecast derived indices are plotted against the actual drought index computed using observed discharge values.

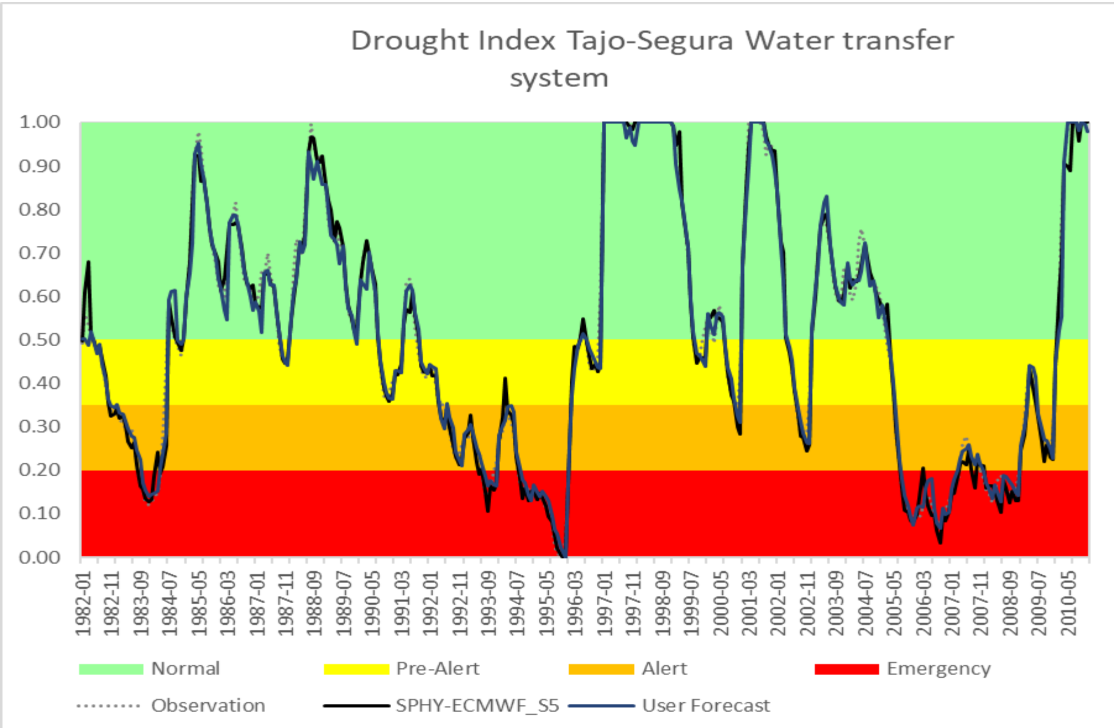


Figure 3-4: Drought index for the Tajo-Segura Water Transfer system (Period 1981-2010)

From a general point of view, both timeseries reflect a high performance in predicting the real drought index. This high level of predictability is somewhat surprising given the evaluation presented in the previous section. However, it can be explained by the fact that the index hardly depends on forecasted discharge: other (observed) variables such as current reservoirs levels are dominant and lead to large agreement between the two selected methods. Nevertheless, some differences between the systems are noted leading in some cases to different conclusions regarding predicted drought category.

A closer look at the user forecast and SPHY-ECMWF_SEAS5 coupled forecasting system performance in predicting the drought can be seen in Figure 3-5 and Figure 3-6. Several ROC curves are plotted together with AUC (Area under the curve) values showing the ability of the different methods to successfully discriminate between events and non-events, here correctly predicting a particular drought category. A high success rate for both methods is noted amongst all categories, in particular for normal and emergency levels with slightly worse performances for the mid alert and pre-alert categories.

Table 8: Skill score based on the ROC metric for probabilistic SPHY-ECMWFSEAS5 coupled system using the user forecast as reference

ROCSS	LD1/LD2/LD3	LD3
JFM	0.81	0.77
AMJ	0.68	0.67
JAS	0.77	0.20
OND	0.89	0.64

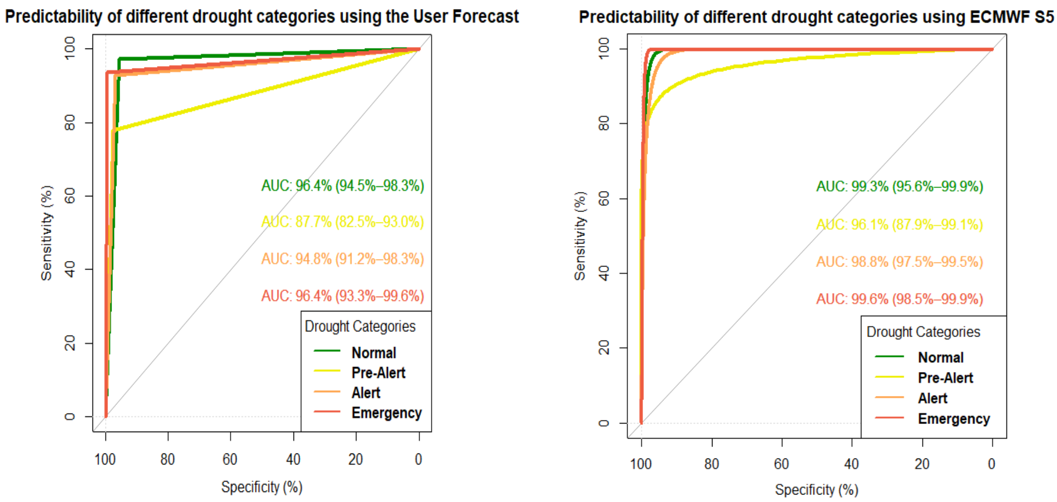


Figure 3-5:ROC curves for drought index predictability using the User Forecast on the left and the SPHY-ECMWF SEAS5 probabilistic system on the right for the long period (1982-2010)

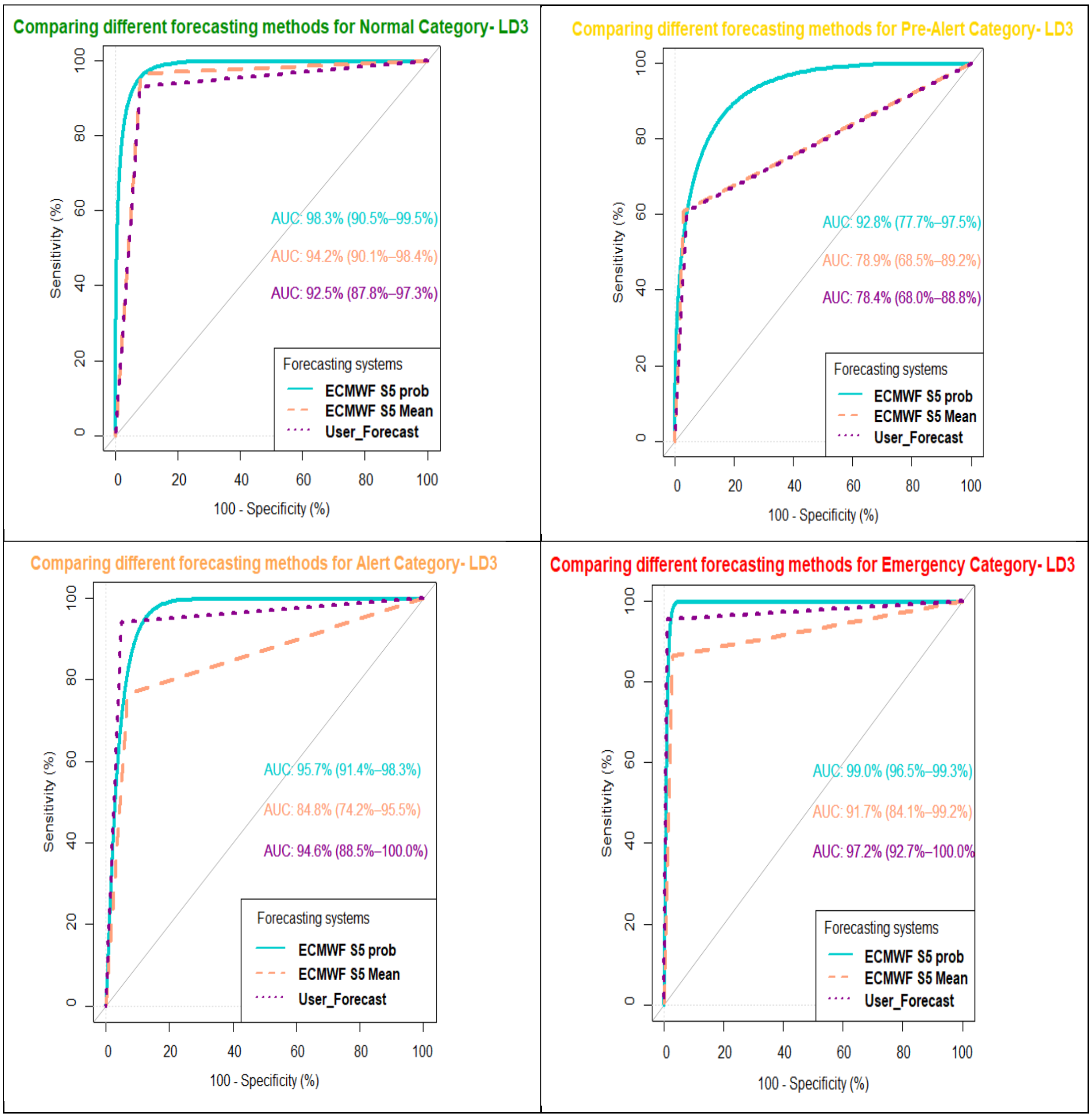


Figure 3-6: ROC curves for drought index predictability comparing all forecasting methods for lead time 3 (Mars, June, September and December) over the long period 1982-2010

3.5 Conclusion

Seasonal hydrological forecasting systems are increasingly tested in scientific research and operational case studies for Europe (Meißner et al., 2017). Despite showing relatively low skill relative to other continents, funding policies over the last two decades have encouraged efforts in improving such systems over statistical methods of prediction (Barnston et al., 2012). The work presented here fits well into this progress, comparing a user statistical forecasting method to a dynamic model-based forecasting method in generating streamflow value for a particularly selected basin. The distributions of 3-month average streamflow outputs are better represented by the dynamic model as user forecast tend to show low variability between and among seasons (see Figure 3-3).

Considered metrics and in particular for MAE/CRPS and ME (scores of bias) have shown that the difference between the two forecasting systems is relatively small with the dynamic process-based model performing slightly better during winter periods JFM and AMJ (see Table 6 and Table 7). Although no single metric can represent a comprehensive comparison between two forecasting methods, case-to-case discrimination as indicated by correlation skill reflects a fundamental ability of a prediction model (Murphy and Epstein, 1989). It is understood that bias-related problems are more easily correctable through calibration efforts (Barnston et al., 2012). To that end, the seasonal dynamic model has surpassed (Table 3) the user forecast in correlation skill for all considered periods except AMJ where it reported similar predictive skill. Furthermore, more sophisticated techniques for statistical post-processing of seasonal flow forecasts based on re-forecasts could also increase forecasting skill. Considering forecast of drought indices, both systems report a very high ROC score, but with the dynamic process-based performing slightly but consistently better than the user forecast among all categories and lead times.

With performance relatively comparable between the two systems (the dynamic system only slightly outperforms the benchmark), it is important to point out the fundamental differences between the two selected methods. The user forecast model is based on a statistical relationship established between predictand and predictors whereas the mechanistic model relies on estimating equations that describes complex relationships between biophysical processes. The added value of the first system relates to its computational easiness and speed of processing when compared to the more demanding resources of mechanistic models. Nevertheless, the skill derived from empirical models relies heavily on the assumption of climatological stationarity. To build up a statistical model, a long series of data is necessary in order to derive a robust relationship between predictand and predictors. This is increasingly problematic in future climate change scenarios estimated to impact drastically the current climatology and hence the predictive capacity of empirical models (Ogutu et al., 2017).

Although dynamic models require large historic datasets for skill verification in addition to a reference for defining variable anomalies, such information is not relevant to their basic functioning (Barnston et al., 2012). Skill derived with coupled seasonal forecasting models is usually dependent on the presence of slow and predictable variations in soil moisture, snow cover, sea-ice, ocean surface temperature in addition to the behaviour of the atmosphere within those boundary conditions (Turco et al., 2017). Currently, the limiting factor in reference to skill relates to initialization errors in addition to bias drifts introduced by the imperfect numerical representation of the different physical processes occurring on a multitude of temporal and spatial scales (Barnston et al., 2012). It is likely that these shortcomings will continuously be improved in time increasing the quality gap between statistical and dynamic methods (Chen and Cane, 2008).

Overall, the performance assessment showed that the developed dynamic forecasting system can in principle provide valuable information for drought-related decision-making in the Segura river basin. However, in practice, the use of these systems may still be limited due to various non-technical factors. To obtain information on the usefulness from a users´ perspective, the Segura forecasting system was put in operational mode during one winter season. The next section summarizes this pilot effort.

4 Operational evaluation

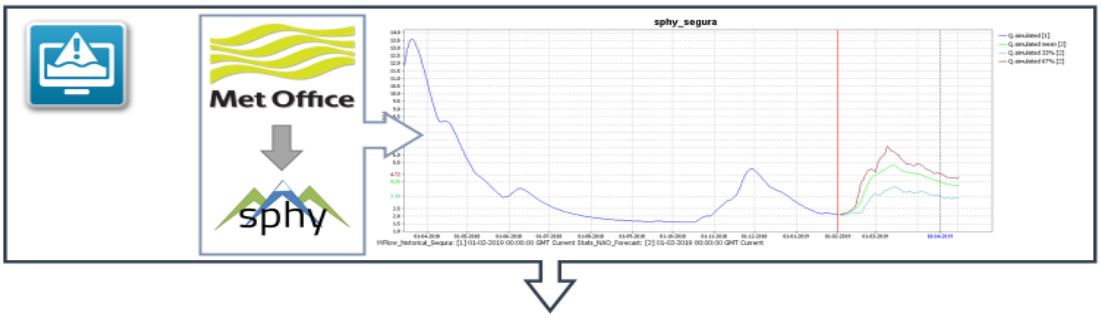
4.1 Specifications of Seasonal Hydrological Outlook

For an operational evaluation, the model was implemented in an operational mode. For this, close collaboration took place with two IMPREX partners: UK MetOffice and Deltares. The SPHY model presented before was used to provide the hydrological predictions for a forecast period of three months ahead. Delft-FEWS was used to put the system in operational mode. Figure 4-1 shows a schematic of the setup.

Key specifications are:

Glosea5 NAO index predictions as proxy predictor for rainfall values (doi:10.1175/JAMC-D-15-0102.1, 2016)

- The Spatial Processes in HYdrology model (SPHY) for streamflow simulation, set up for the upper Segura basin (2x2 km). (doi:10.5194/gmd-8-2009-2015, 2015).
- The SPHY model is calibrated against discharge observations using the SPOTPY tool (doi:10.1371/journal.pone.0145180,2015.) (Simulated Annealing algorithm, RMSE objective function).
- Initial conditions used for the operational forecast are retrieved from the ECMWF EFAS system. (doi:10.2760/806324, 2019)
- Historical runs in order to generate the climatological distributions are obtained using the E-OBS gridded dataset (doi:10.1029/2017JD028200, 2018)
- The whole coupled model-based system runs on the FEWS system (doi:10.1016/j.envsoft.2012.07.010, 2013)
- Output of interest are 3 month predictions initialized in Dec, Jan and Feb



Rainfall	Dec-Feb	Jan-Mar	Feb-Apr
Above-normal	47%	44%	56%
Near-Normal	37%	38%	36%
Below-normal	16%	18%	8%

Streamflow	Dec-Feb	Jan-Mar	Feb-Apr
Above-normal	82%	44%	17%
Near-Normal	18%	52%	34%
Below-normal		4%	49%

Figure 4-1. Schematic of the setup of data flows of the operational setup

4.2 Description of pilot during winter 2018-2019

The dynamic forecasting system for the Segura river basin was used for producing operational forecasts during the winter 2018/2019, targeting the Segura River Basin Authority as user. A tailored risk outlook was developed and co-designed with the user – meaning that there were several iterations. The principal aim of this pilot was to receive feedback on the possibilities and limitations to integrate the tool into their decision-making.

The user received the seasonal risk outlook in the form a monthly bulletin, using seasonal forecast data, during three consecutive months. The bulletin was sent by email to the Water Resources Planning department as soon as the bulletin was released, in December 2018, January 2019 and February 2019. The bulletins included a forecast about rainfall and flows three months ahead. During the service provision, the information in the bulletins was evaluated by the user, by means of informal meetings and feedback by email. More details on the risk outlook and the bulletin can be found in Deliverable 14.5 of the IMPREX project (www.imprex.eu).

The bulletins included:

- Probabilistic values for expected precipitation and streamflow conditions for the upcoming 3 month period.
- Information on how to interpret those values against a climatological forecast where each category is predicted with a 1/3 chance of occurrence.
- Explanation on possible sources of predictability and main mechanisms behind the forecasting system to make sure that the decision maker can value the displayed results.

4.3 Example of a released bulletin

On the four next pages, one of the Seasonal Hydrrological Outlooks is presented. This particular one was produced for the forecast period of Feb-April 2019.

Pronóstico estacional de la Cuenca del Segura

Titulares

Periodo Pronóstico	Fecha previsión	Fecha documento
Febrero 2019 a Abril 2019	01-Febrero-2019	11-Marzo-2018

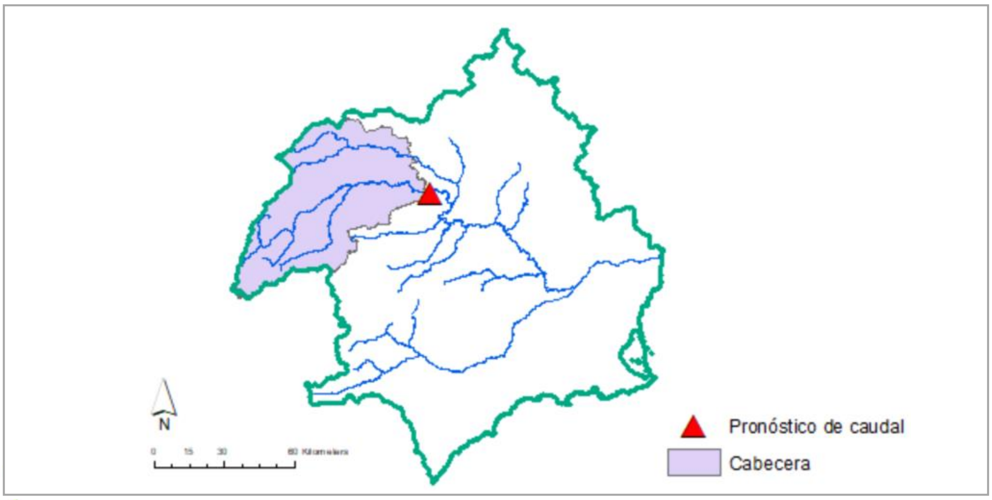
Resumen

Para el periodo desde el 1 febrero 2019 hasta el 30 de abril 2019 :

- Los pronósticos de precipitación para las dos regiones de interés se prevén superiores (alta probabilidad) al valor normal observado en condiciones promedio.

- Existe una baja probabilidad de padecer condiciones más secas que la media en términos de precipitaciones. Como referencia, durante el mismo periodo del año anterior, se registraron una situación anómala con precipitaciones inferiores al valor normal.

- El pronóstico de caudal arroja una predicción inferior (alta probabilidad) al valor normal observado en el mismo periodo. El pronóstico de precipitación para el periodo de estudio, con valores superiores a a lo normal, no compensa el efecto de bajo caudal observado al inicio del periodo de simulación.



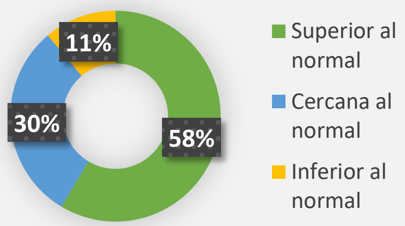
@imprex_eu FutureWater Deltares Met Office

Pronóstico estacional de la Cuenca del Segura

Precipitación

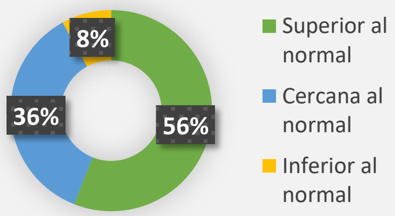
Febrero 2019 – Abril 2019	Titular	Prob. Superior al valor normal (invierno típico = 33%)	Prob. cercana al valor normal (invierno típico = 33%)	Prob. Inferior al valor normal (invierno típico = 33%)
Precipitación (toda la Cuenca del Segura)	**Probabilidad aumentada de precipitación por encima de lo normal**	**58%**	**30%**	**11%**
Precipitación en la Cabecera	**Probabilidad aumentada de precipitación por encima de lo normal**	**56%**	**36%**	**8%**

Predicción de probabilidades de precipitación (toda la cuenca)



Ejemplo de años con valores por encima de la normalidad: 2004, 2005, 2010

Predicción de probabilidad de precipitación (solo Cabecera)



Ejemplo de años con valores por encima de la normalidad: 2003, 2010, 2014, 2017

¿Qué significan estos valores?

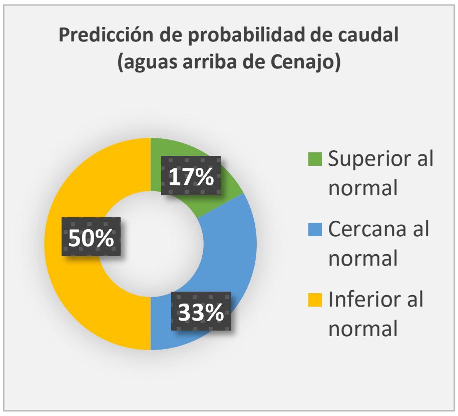
La predicción de precipitación se agrupa en tres categorías según los valores observados en los periodos (trimestres invernales) anteriores. Las figuras de arriba muestran la probabilidad de ocurrencia de cada categoría según la predicción realizada para el periodo simulado del año anterior. Bajo condiciones normales-promedio, la probabilidad de ocurrencia de cada una de las categorías es del 33%.

Para una categoría concreta, una predicción **superior** al 33% significa un **aumento** de la probabilidad de ocurrencia; una predicción **inferior** al umbral del 33% implica una **reducción** en la probabilidad de ocurrencia de esa categoría. Es importante subrayar que la precipitación total dentro de una categoría no es necesariamente idéntica o representa condiciones de abundancia o escasez igualmente severas.

Pronóstico estacional de la Cuenca del Segura

Caudal

Enero 2019 – Marzo 2019	Titular	Prob. Superior al valor normal (invierno típico = 33%)	Prob. cercana al valor normal (invierno típico = 33%)	Prob. Inferior al valor normal (invierno típico = 33%)
Caudal (aguas arriba de Cenajo)	**Probabilidad aumentada de precipitación por encima de lo normal**	**17%**	**34%**	**49%**



¿Qué significan estos valores?

La predicción de caudal se agrupa en tres categorías según los valores observados en los periodos (trimestres invernales) anteriores. Las figuras de arriban muestran la probabilidad de ocurrencia de cada categoría según la predicción realizada para el periodo simulado del año anterior. Bajo condiciones normales-promedio, la probabilidad de ocurrencia de cada una de las categorías es del 33%.

Para una categoría concreta, una predicción **superior** al 33% significa un **aumento** de la probabilidad de ocurrencia; una predicción **inferior** al umbral del 33% implica una **reducción** en la probabilidad de ocurrencia de esa categoría. Es importante subrayar que el caudal total dentro de una categoría no es necesariamente idéntica o representa condiciones de abundancia o escasez igualmente severas.

Pronóstico estacional de la Cuenca del Segura

Metodología para realizar el pronóstico estacional

En los últimos años, los modelos informáticos de pronóstico meteorológico han sido mejorados y empleados para hacer predicciones estacionales. El pronóstico estacional es más difícil que el pronóstico diario por lo que estos modelos solo pueden aportar indicaciones sobre la probabilidad general de que una estación sea húmeda o seca, templada o fría. Esta información permite evaluar la probabilidad de que ciertos eventos, por ejemplo sequías, ocurran. En la Cuenca del Segura, se ha comprobado que estos modelos pueden pronosticar la precipitación con cierta efectividad.

Adicionalmente, los pronósticos de caudal se han obtenido mediante modelización del sistema hidrológico. Para una resolución temporal estacional, los pronósticos derivados de estos modelos dependen del volumen de agua almacenada en el sistema (acuíferos, embalses) y de la cantidad de agua de lluvia precipitada.

El pronóstico de caudal del Segura es el resultado de la combinación de modelos de pronóstico meteorológico y modelos hidrológicos que han sido integrados bajo un mismo marco metodológico desarrollado dentro del proyecto IMPREX (www.imprex.eu). Posibles escenarios de precipitación derivados de un modelo de pronóstico meteorológico pueden ser usados con un modelo hidrológico para evaluar las probabilidad estacional de caudal y el volumen de agua potencialmente disponible para satisfacer las demandas de agua para los diferentes sectores productivos.

¿Cómo de fiable es el pronóstico?

Las probabilidades calculadas representan los niveles de confianza estimados para el rango de posibles resultados o categorías consideradas.

¿Qué genera un aumento en la probabilidad de ocurrencia de evento húmedo?

Los patrones meteorológicos de invierno observados en Europa están influenciados por los patrones climáticos globales. La ocurrencia de los fenómenos de El Niño/La Niña en el Océano Pacífico Tropical, o el patrón de temperaturas en el Océano Atlántico Norte pueden desencadenar cambios en los patrones meteorológicos promedio en la estación invernal. Este invierno, hay un evento Niño de intensidad moderada, que tiende a favorecer un patrón de reducción y aumento de la precipitación en el Norte y Sur de Europa, respectivamente. Este fenómeno se desencadena a través de un "puente" entre el Pacífico y Europa que se caracteriza por fuertes vientos en altura, y que pronostica una mayor probabilidad de ocurrencia de condiciones húmedas en la cuenca del Segura al final del trimestre invernal.

Créditos

Este pronóstico usa pronósticos estacionales del Servicio Meteorológico Nacional del Reino Unido. Se agradece el soporte brindado por el proyecto IMPREX financiado por la Unión Europea a través del Programa H2020 de Investigación e Innovación

@imprex_eu

4.4 Evaluation by the user

During the evaluation period, several interactions took place with the technical staff and head of the planning department of The Segura River Basin Authority. During initial meetings, they indicated that a dynamic seasonal forecast of drought indices could potentially improve their decision-making. The drought indices they use currently perform well and in principle they see no need to change them. However, they see scope in studying the usefulness and skill of seasonal forecasts of these indices over the following months or next year. In some occasions during drought periods, some stakeholders of the Basin Authority had requested a seasonal outlook of water resources availability, but so far the authority had not been able to respond to this sort of requests, as no forecasting system is available nowadays to them (only the statistical method for the connected Upper Tagus).

Figure 4-2. Intermediate evaluation meeting with Jesús Garcia (head Planning department, right) and Jaime Fraile Jiménez de Muñana (drought expert, left) of Segura River Basin Authority

The user evaluation during the three-month pilot was performed by means of a combination of physical meetings and interaction through email and phone calls. After finalizing the pilot, a final reflection was requested from the technical staff and the head of the Water Resources Planning department. The feedback received can be summarized in the following points:

- The information presented in the bulletins was considered clear and understandable for the technical staff working with this type of information. Key in this aspect was considering the feedback from user in the design of the bulletin.
- The bulletins were received with high expectations, each time they were delivered. The user is well aware of the probabilistic nature of the forecasts and the inherent limitations in skill. Therefore, during this pilot the information was not used to modify any decisions: decisions followed the stipulated procedures as in the Drought Management Plan.
- The performance of the system for the testing period was limited (the forecasts were relatively optimistic in terms of rainfall while rainfall amounts turned out to be relatively low). This made the user reluctant in going forward with the integration of dynamic seasonal forecast information in their decision-making, given the current state-of-art. Also, they highlighted that this would require a large effort in engaging with and convincing of all the water-users that depend on the allocation decisions of the user.

Overall, the water authority emphasized to be very much interested in a longer testing period, at least during one additional winter. Also, the user sees scope in using the three-month forecasts of rainfall for generating meteorological drought index (SPEI) forecasts. The user has just integrated SPEI in their drought monitoring system (and corresponding Drought Management Plan). The possibility of forecasting SPEI is something that the user would like to study in the near future.

5 References

Barnston, A.G., Tippett, M.K., L'Heureux, M.L., Li, S., Dewitt, D.G., 2012. Skill of real-time seasonal ENSO model predictions during 2002-11: Is our capability increasing? Bull. Am. Meteorol. Soc. 93, 631–651. https://doi.org/10.1175/BAMS-D-11-00111.1

Chen, D., Cane, M., 2008. El Niño prediction and predictability. J. Comput. Phys. vol: 227 (, pp: 3625-3640.

Doblas-Reyes, F.J., García-Serrano, J., Lienert, F., Biescas, A.P., Rodrigues, L.R.L., 2013. Seasonal climate predictability and forecasting: Status and prospects. Wiley Interdiscip. Rev. Clim. Chang. 4, 245–268. https://doi.org/10.1002/wcc.217

Greuell, W., Franssen, W.H.P., Biemans, H., Hutjes, R.W.A., 2018. Seasonal streamflow forecasts for Europe-Part I: Hindcast verification with pseudo- A nd real observations. Hydrol. Earth Syst. Sci. 22, 3453–3472. https://doi.org/10.5194/hess-22-3453-2018

Hagemann, S., Chen, C., Haerter, J.O., Heinke, J., Gerten, D., Piani, C., 2011. Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models. J. Hydrometeorol. 12, 556–578. https://doi.org/10.1175/2011JHM1336.1

Herrera, S., Fernández, J., Gutiérrez, J.M., 2016. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. Int. J. Climatol. 36, 900–908. https://doi.org/10.1002/joc.4391

Houska, T., Kraft, P., Chamorro-Chavez, A., Breuer, L., 2015. SPOTting Model Parameters Using a Ready-Made Python Package. PLoS One 10, e0145180. https://doi.org/10.1371/journal.pone.0145180

Jin, E.K., Kinter, J.L., Wang, B., Park, C.K., Kang, I.S., Kirtman, B.P., Kug, J.S., Kumar, A., Luo, J.J., Schemm, J., Shukla, J., Yamagata, T., 2008. Current status of ENSO prediction skill in coupled ocean-atmosphere models. Clim. Dyn. 31, 647–664. https://doi.org/10.1007/s00382-008-0397-3

Mason, S.J., Stephenson, D.B., 2008. How Do We Know Whether Seasonal Climate Forecasts are Any Good? Seas. Clim. Forecast. Manag. Risk 82, 265–296. https://doi.org/10.1007/978-1-4020-6992-5

Meißner, D., Klein, B., Ionita, M., 2017. Development of a monthly to seasonal forecast framework tailored to inland waterway transport in central Europe 6401–6423.

Murphy, A.H., Epstein, E.S., 1989. Skill Scores and Correlation Coefficients in Model Verification. Mon. Weather Rev. https://doi.org/10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2

Ogutu, G.E.O., Franssen, W.H.P., Supit, I., Omondi, P., Hutjes, R.W.A., 2017. Skill of ECMWF system-4 ensemble seasonal climate forecasts for East Africa. Int. J. Climatol. 37, 2734–2756. https://doi.org/10.1002/joc.4876

Terink, W., Lutz, A.F., Simons, G.W.H., Immerzeel, W.W., Droogers, P., 2015. SPHY v2.0: Spatial Processes in HYdrology. Geosci. Model Dev. 8, 2009–2034. https://doi.org/10.5194/gmd-8-2009-2015

Turco, M., Ceglar, A., Prodhomme, C., Soret, A., Toreti, A., Doblas-Reyes Francisco, J., 2017. Summer drought predictability over Europe: empirical versus dynamical forecasts. Environ. Res. Lett. 12, 084006. https://doi.org/10.1088/1748-9326/aa7859

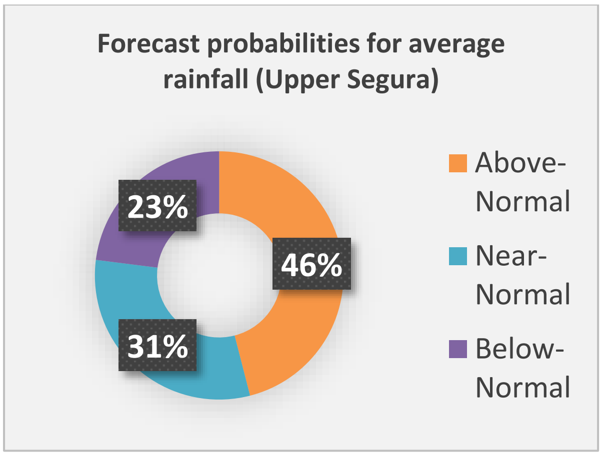
Weisheimer, A., Palmer, T.N., 2013. On the reliability of Seasonal Climate Forecasts.

Annex 1 – Q&A of the Seasonal Hydrological Outlook for the Segura River Basin

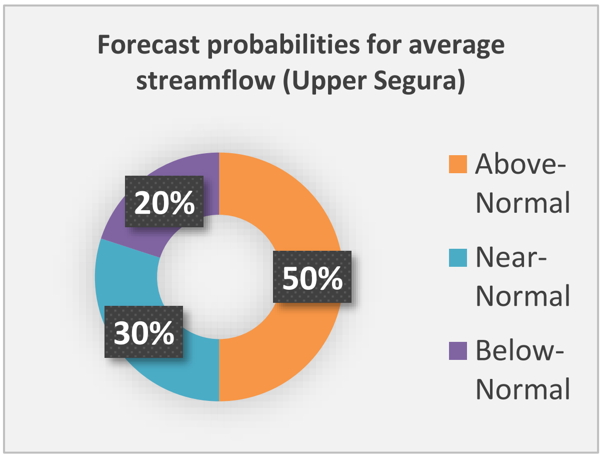
- **Forecast period?** 3-month seasonal outlooks in December, January and February
- **Place?** Segura River Basin (upstream of the Cenajo dam, 2602 km^2)

Orchard in Southeast of Spain, Murcia heavily impacted by drought conditions

- **Decision maker?** The River Basin Authority for the Segura (*Confederación Hidrográfica del Segura* – CHS)
- **Objective?** To guarantee proper water allocation for the different concerned stakeholders based on sustainable water transfer decisions, namely the irrigated agriculture community.
- **Key predicted variables?** Precipitation and discharge forecasts for the next 3-months.
- **How are these forecasts made?** In recent years, computer weather forecast models have been extended to make forecasts for the season ahead. Forecasting for this range is more difficult than for the next few days, so we can only indicate the likelihood of the season being wet or dry, mild or cold, overall. Nevertheless, this allows an assessment of how likely certain hazards like drought may be. Predictive skill is found over the Segura basin in rainfall predictions for winter, these are used as input to produce hydrological seasonal forecasts. Streamflow forecasts are obtained by using hydrological models describing the various parts of the hydrological system such as aquifers, rivers etc. For seasonal timescales, forecasts from these models depend on the water stored in the hydrological system on the moment of the forecast, and on the predicted amount of rainfall for the next months. The Segura seasonal outlook benefits from a combination of weather (GloSeas5) and hydrological models (SPHY) into an integrated framework (Delft-FEWS), a technique that has been developed within the EU IMPREX project.
- **How is this information used?** This information is communicated with the Segura river basin authority in the form of informative bulletins. These include probabilistic values for expected precipitation and streamflow conditions in the Segura river basin. It includes an explanation on possible sources of predictability to make sure that the decision maker can value the displayed results.



○



- **What do these figures show?** Rainfall/streamflow is grouped into three categories based on local observations of past winters. The figures above show the probability of each category occurring depending on the forecast for winter 2019. For a typical winter, the chances of a category to occur is of 33%. Consequently, a forecast predicting a value above that threshold (33%) for a particular category implies an increased chance for that category to occur. In the same way, a forecast value below that threshold implies a decreased chance of occurrence.
- **How confident is this outlook?** The probabilities reflect the level of confidence – these can be seen in the above diagrams, which show the level of confidence in the range of possible outcomes.
- **More info?** Contact **Johannes Hunink (j.hunink@futurewater.es)**