Seasonal hydrological forecasting in Europe: Analysis of skill and its key driving factors

Pechlivanidis I.G.[1*], Crochemore L.[1] And Bosshard T.[1]

[1] Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden

*corresponding author

e-mail: ilias.pechlivanidis@smhi.se

Abstract

Recent advances in understanding and forecasting of climate have led into skillful meteorological predictions, which can consequently increase the confidence of hydrological prognosis. There is currently a need to understand the large European river systems and make practical use of seasonal hydrological forecasts. Here, we analyze the seasonal predictive skill along Europe's hydro-climatic gradient using the pan-European E-HYPE multi-basin hydrological model. Both model state initialization and provision of climatology are based on forcing input derived from the WFDEI product. An ensemble of re-forecast forcing data (daily mean precipitation and temperature) from ECMWF System 4 are firstly bias corrected using a modified version of the DBS method, and further used to drive E-HYPE. The predictive skill of streamflow based on ECMWF and climatology for the European basins is assessed on monthly timescales. Seasonal re-forecasts are evaluated geographically and temporally with respect to their accuracy against perfect forecasts of streamflow. We analyze the skill across 35408 subbasins, which represent various climatologies, soil-types, land uses, altitudes and basin scales within Europe. We finally use the Classification and Regression Trees analysis to link the gain in the seasonal skill to physiographic-hydro-climatic characteristics and meteorological skill, in order to suggest possible improvements.

Keywords: Seasonal hydrological forecasting, E-HYPE, ensemble forecasts, pan-European scale

1. Introduction

Seasonal forecasts hold the potential for being of great value for a wide range of stakeholders/end-users who are affected by the vagaries of the climate and who would benefit from understanding and better managing climate-related risks (Wood and Lettenmaier, 2006). Recent advances in our understanding and forecasting of climate have resulted in skillful and useful meteorological predictions, which can consequently increase the confidence of hydrological prognoses, and awareness from an end-user perspective. Efforts at the Swedish Meteorological and Hydrological Institute complement the "deep" knowledge from basin-based modeling using process-based multi-basin modeling at the continental scale (http://hypeweb.smhi.se), which are capable of representing human influences (i.e. irrigation, reservoirs etc.). Such models can encompass many river basins, cross regional and international boundaries, represent a number of different physiographic and climatic zones, and describe shifts in streamflow regime due to human-impacts; hence advance process understanding by building a numerical background for comparative hydrology. This approach allows further the identification of regions of similar hydrological forecasting skill, whilst it allows a potential link of the skill to physiographic-hydro-climatic characteristics and meteorological skill, in order to suggest possible model improvements. However, understanding processes in large systems is challenging. Indeed, physical properties (e.g. vegetation and soil type) generally exhibit high spatial variability, which results in significant differences in system behavior and predictability. As expected, this spatial heterogeneity introduces further high uncertainty on the categorization of important drivers that influence the predictive hydrological skill. In addition, large river basins are often strongly influenced by human activities (e.g. irrigation, hydropower production, groundwater use) for which information is rarely available and therefore rarely described in hydrological model processes; hence introducing additional uncertainty regarding process understanding and description. Although such modeling type has limitations, which vary in space, here we make the step forward to gain insights in spatial patterns of hydrological skill at the large scale, and link this to the characteristics of the basin system.

2. Methods

2.1 E-HYPE impact model

The Hydrological Predictions for the Environment, HYPE, model is a dynamic, semi-distributed and process-based model capable of describing the hydrological processes at the basin scale. The model represents processes for snow accumulation and melting, evapotranspiration, soil moisture, discharge generation, groundwater recharge, and routing through rivers and lakes. HYPE simulates the water flow paths in soil, whilst the outflow from a lake is determined by a rating curve. Irrigation is simulated based on crop water demands calculated either with the FAO-56 crop coefficient method or relative to a reference flooding level for submerged crops (e.g. rice). The demands are withdrawn from rivers, lakes, reservoirs, and/or groundwater within and/or external to the sub-basin where the demands originated. The HYPE model setup for the pan European region (8.8 million km^2) is named E-HYPE. The model has a spatial resolution of 35408 sub-basins, i.e. in average 215 km^2 and runs at a daily time step. The model requires information on terrain, soil and land use, lakes and reservoirs and irrigation as input, which, in this application, has been obtained from global sources (Hundecha *et al.*, 2016). Forcing data (daily mean precipitation and temperature) based on the WFDEI product (Weedon *et al.*, 2014) has been used to calibrate E-HYPE.

2.2. Forcing data – Seasonal forecasts and bias correction

The E-HYPE hydrological model needs initial conditions (level in surface water, i.e. reservoirs, lakes and wetlands, soil moisture, snow depth) that are obtained by driving the model using "observations" for a spin-up period. Mean daily precipitation and temperature are derived from the WFDEI product and drive the hydrological model for the period 1979-2010. Re-forecast forcing data (i.e. daily mean precipitation and temperature) from ECMWF System 4 (15 members initialized every month) are also available for the period 1981-2010. The re-forecast data were bias corrected using a modified version of the Distribution Based Scaling (DBS) method (Yang *et al.*, 2010) to account for drifting conditioning the bias correction on the lead month. It has been adapted from the quantile-mapping method for application in seasonal forecasting. Bias-correction is conducted on all members of System 4 using WFDEI data as reference. After bias correction, the cumulative distribution of daily precipitation and temperature values closely follows that of the WFDEI data.

2.3. Skill metrics

The skill of seasonal hydrological forecasts was analyzed in all 35408 locations across Europe. The objective was to assess the capacity to forecast the monthly average discharge (hence focus mainly on representing the volume of water) for different lead times (i.e. 0, 2, and 4 months after the initialization) and for all 15 ensemble members. The *beta* metric is used here (Gupta *et al.*, 2009):

$$beta = 1 - \sqrt{(\beta - 1)^2} \qquad \text{(Equation 1)}$$

β is the ratio of the monthly mean of the forecasts (the output of the model forced by ECMWF) over the monthly mean of the "observations" (the output of the model forced by WFDEI); the range of values varies between -∞ and 1, with 1 being the optimum.

2.4 Clustering of skill

To better understand the potential controls of skill and identify regions of similarity, we apply classification and regression trees (CART). Here, we explored the spatial runoff patterns across the entire subcontinent by analyzing the skill in all 35408 catchments modeled by the E-HYPE model. CART is a recursive-partitioning algorithm that classifies the space defined by the input variables/descriptors (i.e. physiographic-hydrologic-climatic characteristics, and remaining climatic biases) based on the output variable (i.e. *beta* skill for lead month 2 and month March). In this case, *beta* is divided into five groups – bad (*beta* < 0.2), poor (0.2 < *beta* < 0.4), medium (0.4 < *beta* < 0.6), good (0.6 < *beta* < 0.8) and very good (*beta* > 0.8), which are termed C0, C1, C2, C3 and C4 respectively. A terminal leaf exists at the end of each branch of the tree, where the probability of belonging to any of the five output groups can be inspected. Here, we summarized the basin characteristics into climatology and biases in forcing input, topography, human impacts, and hydrologic signatures (Table 1).

Table 1. Basin characteristics used in the clustering analysis.

Climatology / Forcing biases	Topography	Human impact	Hydrologic signatures
Precipitation (mm/month)	Area (km^2)	Degree of regulation (%)	Mean annual specific runoff (Qm)
Temperature (°C)	Elevation (m)		Normalized high flow (q05)
Snow depth (cm/month)	Relief ratio (-)		Normalized low flow (q95)
Actual evaporation (mm/month)	Slope (%)		Normalized relatively low flow (q70)
Potential evaporation (mm/month)			Slope of flow duration curve (mFDC)
Dryness index (-)			Range of Parde coefficient (DPar)
Evaporative index (-)			Coefficient of variation (CV)
Bias in precipitation (%)			Flashiness (Flash)
Bias in temperature (%)			Normalised peak distribution (PD)
			Rising limb density (RLD)
			Declining limb density (DLD)
			Baseflow index (BFI)

We then calculate the predictors' importance (and rank them) by summing changes in the risk due to splits on every predictor and dividing the sum by the number of branch nodes. In order to avoid the high dimensionality in the CART analysis, the hydrologic signatures were first clustered into 11 groups with each group receiving an ID (named FlowID) with the k-means clustering approach.

3. Results

3.1. Analysis of skill

We first investigate the spatial variability in performance of the seasonal streamflow forecasts. Fig. 1 presents the variability of the *beta* metric across Europe, for the summer and winter seasons, and lead months 0 and 4. The quality of the forecasts in the first lead month, with regards to the *beta* criterion, is high over Europe with values mostly greater than 0.8. This indicates that the seasonal forecasting system, here composed of ECMWF System 4 forecasts and E-HYPE trained with WFDEI data, reproduces well the water volumes for the month ahead. Patches of lower performance can be observed around the Caspian Sea, in Norway and in the Alps in winter, and in Turkey, in the Iberian Peninsula and in the northernmost parts of Europe in summer. As expected, the quality of the forecasts decreases with increasing lead time, especially in Central Europe and in the Mediterranean region. Nevertheless, forecast performance remains greater than 0.8, indicating that water volumes can be generally well represented by System 4 even four months prior to the target month. Wider patches of lower performance can be observed than at one-month lead: in the Southern half of Europe and in Norway in winter, and in Scandinavia and Turkey in summer.

3.2. Comparative analysis - Ranking of descriptors

To spatially interpret hydrological skill and identify the key controls of poor/good model skill, we investigated potential relationships between predictive skill and physiographic-hydrological-climatic characteristics. First, the 15 descriptors were analyzed for inter-dependence, and omitted when these were highly inter-dependent to avoid potential artefacts in the CART regression analysis. Consequently a set of nine significant descriptors was statistically identified for application in the CART analysis, which further allowed us to estimate the descriptors' importance. Figure 2 shows the ranking of nine descriptors (ranked by importance, with 1 being the most important descriptor) for all months and lead months. Results show that the dominant descriptors resulting in poor/good model performance are the FlowID, which describes the hydrological behavior of the basin, elevation and remaining bias in temperature (BiasTemp). It is generally expected that remaining biases in temperature will have an impact on the form of precipitation (rainfall or snowfall) during the cold months, and the processes (i.e. changing from (to) snow accumulation to (from) melting). For example, this occurs in northern Europe for April where the mean average temperature for April is close to 0^{o}C and hence small deviations in the meteorological forecasts will affect the basin response. CART indicated elevation (Elev.) to also be an important factor. This could

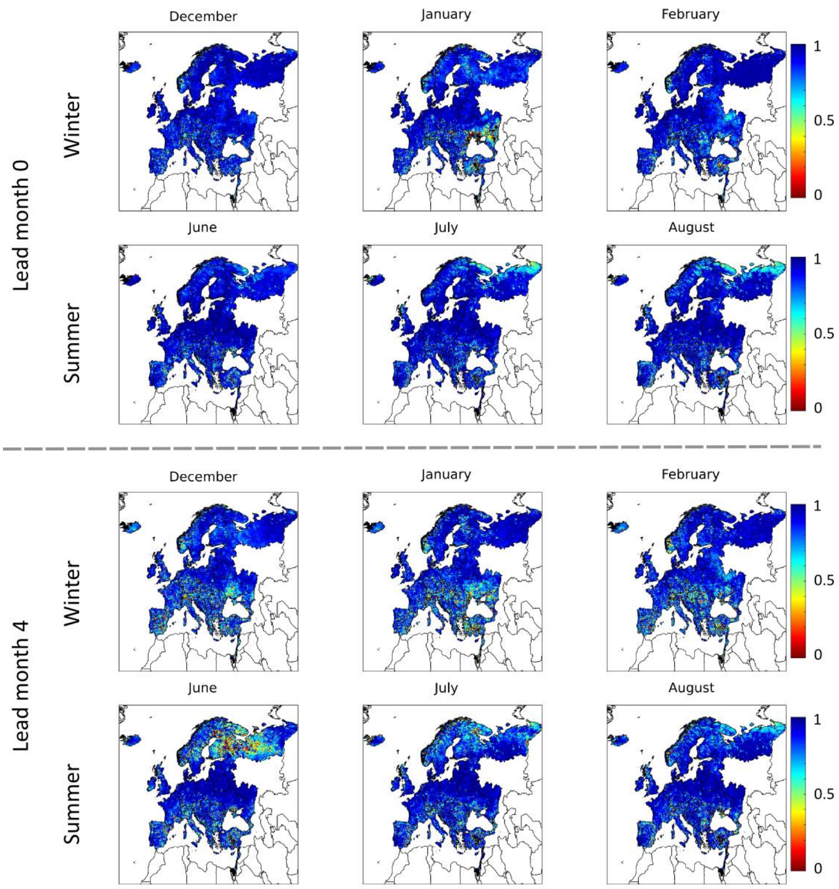

Figure 1. Spatial variability of the *beta* performance for two seasons and lead months 0 and 4.

be due to the high reliability of the climate forecasts in predicting the climatological variability in highly elevated (usually snow dominated) basins in comparison to low elevated rain-fed basins. The basin hydrological behavior (FlowID) seems to be the most important descriptor with basins of similar river flow properties achieving similar skill. It is known that river systems experience processes with high memory in comparison to the natural phenomena occurring in the atmosphere. Hence it is expected that hydrological variables can have higher predictability than meteorological variables. However, this cannot be linearly translated since the precipitation-discharge process is also not linear, and therefore different systems are expected to respond differently to the meteorological signal.

3.3. Comparative analysis – River basin response

To get a better understanding of the basin characteristics that are characterized by a good/poor skill, Figure 3 shows the 11 spatially variable clusters, their distribution of flow signatures, and the distribution of skill in each cluster group. Similarity in catchment behavior for each class was interpreted and dominant flow generating processes could be distinguished. Results give a clear separation between basins with poor and good skill. Basins in cluster 5 achieve the highest skill. These basins are characterized by high ranges of baseflow (BFI), low monthly variability (DPar), and high values of low flows (q95 and q70). These properties describe basins where short-memory precipitation is aggregated and converted into long-memory discharge. Clusters 6, 7 and 9 gather basins with similar behavior to that of cluster 5, yet these characteristics are less pronounced than in cluster 5. Basins in cluster 8 and 10 are short-memory rivers characterized by flashy response and high seasonal variability (DPar and CV). These basins are responding quite fast to the precipitation signal and with strong dynamics (RLD) whilst contribution from base flow is small (BFI). Basins that belong to clusters 1, 2 and 3 perform adequately and are generally characterized by the same flow signatures. These basins are mainly located in the Scandinavian region and in highly elevated regions of central Europe. They are characterized by medium to high slope in their flow distribution (mFDC), which is an indicator of a regime driven by snowmelt.

4. Conclusion

The evaluation spots the strengths and weaknesses of ensemble seasonal forecasts from ECMWF System 4 (15 members), including trends of performance in various months and lead times. We identified links between forecasting skill and different physiographic and hydro-climatic characteristics.

1. The forecasting system shows good performance in reproducing water volumes in most of Europe (depending on the season); however skill deteriorates as a function of lead time, particularly in central Europe and in the Mediterranean region.

2. CART shows that forecast quality is dependent on the basin's hydrologic regime. Elevation and remaining bias in temperature were also identified as key characteristics to explain skill (hydrologic response in mountainous basins depends on temperature).

3. Skill seems to be limited in relatively flashy basins experiencing strong flow dynamics over the year (less memory in the system).

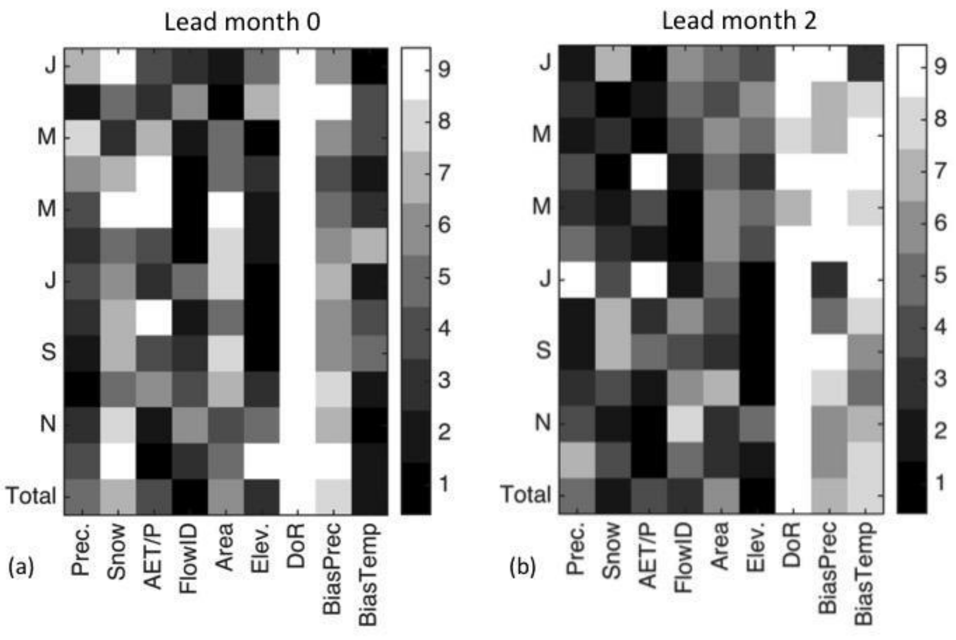

Figure 2. Importance ranking of key descriptors that influence the hydrological forecasting skill over Europe for all months and in lead month: (a) 0, and (b) 2.

.

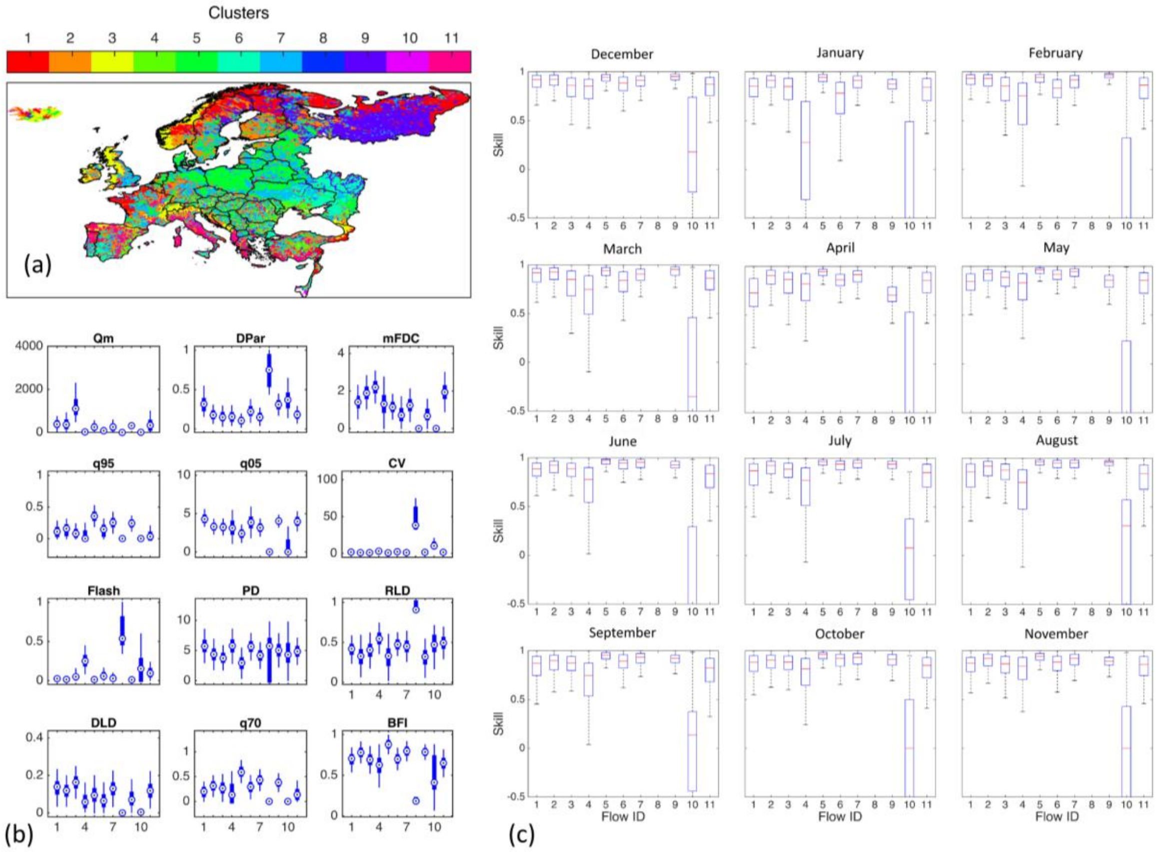

Figure 3. (a) Spatial distribution of hydrologically similar (clusters) basins over Europe, (b) distribution of flow signatures in each cluster group, and (c) distribution of *beta* skill in each cluster group.

References

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. doi:10.1016/j.jhydrol.2009.08.003

Hundecha, Y., Arheimer, B., Donnelly, C., Pechlivanidis, I.G., 2016. A regional parameter estimation scheme for a pan-European multi-basin model. J. Hydrol. Reg. Stud. 6, 90–111. doi:10.1016/j.ejrh.2016.04.002

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M., Viterbo, P., 2014. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. Water Resour. Res. 50, 7505–7514. doi:10.1002/2014WR015638

Wood, A.W., Lettenmaier, D.P., 2006. A test bed for new seasonal hydrologic forecasting approaches in the western United States. Bull. Am. Meteorol. Soc. 87, 1699–1712. doi:10.1175/BAMS-87-12-1699

Yang, W., Andréasson, J., Graham, P.L., Olsson, J., Rosberg, J., Wetterhall, F., 2010. Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. Hydrol. Res. 41, 211–229. doi:10.2166/nh.2010.004