Journal of Hydrology 528 (2015) 720–733

Contents lists available at ScienceDirect

Journal of Hydrology

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

Modification of input datasets for the Ensemble Streamflow Prediction based on large-scale climatic indices and weather generator

Václav Šípek [a,*], Jan Daňhelka [b,1]

[a] The Institute of Hydrodynamics of the Academy of Sciences of the Czech Republic, Pod Paťankou 30/5, Prague 166 12, Czech Republic
[b] Czech Hydrometeorological Institute, Na Šabatce 2050/17, 143 06 Praha, Czech Republic

ARTICLE INFO

Article history:
Received 9 January 2015
Received in revised form 1 July 2015
Accepted 4 July 2015
Available online 10 July 2015
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Yu Zhang, Associate Editor

Keywords:
Seasonal forecasting
ESP
Large-scale climate
Weather generator

SUMMARY

Ensemble Streamflow Prediction (ESP) provides an efficient tool for seasonal hydrological forecasts. In this study, we propose a new modification of input data series for the ESP system used for the runoff volume prediction with a lead of one month. These series are not represented by short historical weather datasets but by longer generated synthetic weather data series. Before their submission to the hydrological model, their number is restricted by relations among observed meteorological variables (average monthly precipitation and temperature) and large-scale climatic patterns and indices (e.g. North Atlantic Oscillation, sea level pressure values and two geopotential heights). This modification was tested over a four-year testing period using the river basin in central Europe.

The LARS-WG weather generator proved to be a suitable tool for the extension of the historical weather records. The modified ESP approach proved to be more efficient in the majority of months compared both to the original ESP method and reference forecast (based on probability distribution of historical discharges). The improvement over traditional ESP was most obvious in the narrower forecast interval of the expected runoff volume. The inefficient forecasts of the modified ESP scheme (compared to traditional ESP) were conditioned by an insufficient restriction of input synthetic weather datasets by the climate forecast.

1. Introduction

Seasonal and long-term hydrological forecasts are based on two major approaches including statistical and physically based methods. Traditional statistical methods use an empirical relation between predictor variables and discharge/flow volume or water level. These relations are based on the analysis of the historical datasets and are thus inevitably influenced by their availability, quality and length. They are represented by the variety of techniques such as artificial neural networks (ANN) (Tawfik, 2003; Huang et al., 2004), regression analysis (Wilby et al., 2004; Opitz-Stapleton et al., 2007), time-series analysis (Pekárová and Pekár, 2006; Gámiz-Fortiz et al., 2011), fuzzy modelling (Mahabir et al., 2003), correlation analysis (Wedgebrow et al., 2002; Trigo et al., 2004), generalised additive models (Underwood, 2009; Ogtrop et al., 2011) or non-parametric methods (Piechota and Dracup, 1999; Oubeidillah et al., 2011). The physically based approach is represented by driving hydrological models with physically based climate model forecasts. The use of the physically based approach is increasing as the abilities of long-term climate forecasts are improving. Wood et al. (2002), Luo and Wood (2008) have applied the physically based forecasting system on the eastern coast of the USA. Coelho et al. (2006) have designed the system based on the climatic model in Brazil. In Europe, the abilities of physically based approach were demonstrated, e.g., by Céron et al. (2010), Fundel et al. (2013). However, compared to the statistical approach the use of the physically based methods remains less common. The reason is regionally variable precipitation predictive skills of current climate prediction models and their limited efficiency concerning longer time horizons (Lavers et al., 2009). Moreover, compared to complex and data demanding physically based models, the statistical approach still represents a convenient option to produce a reliable hydrological forecast. Although, recently observed improvements of numerical weather forecasts (Pappenberger et al., 2011; Gaborit et al., 2014) indicate that they will gradually replace the statistical approach (at least for shorter lead times). The major challenges and drawbacks of the physically based methods are summarised by Cloke and Pappenberger (2009).

* Corresponding author. Tel.: +420 233109026.
E-mail addresses: sipek@ih.cas.cz (V. Šípek), danhelka@chmi.cz (J. Daňhelka).
[1] Tel.: +420 244032300.

http://dx.doi.org/10.1016/j.jhydrol.2015.07.008
0022-1694/

Ensemble Streamflow Prediction (ESP) (Day, 1985) represents, in a certain way, a combination of both above-mentioned approaches. It combines the hydrological model with the climate outlook represented by historical meteorological series under the assumption that these meteorological inputs represent the distribution of possible conditions during the forecast period. The ESP system has been widely tested (Druce, 2001; Kim et al., 2001; Franz et al., 2003) and is currently implemented in the National Weather Service Advance Hydrologic Prediction Service (NWS AHPS) (McEnery et al., 2005). Several studies have demonstrated significant forecast improvements when climate information is used as part of the streamflow forecasting method. Hamlet and Lettenmaier (1999) restricted the number of input series to those with a similar phase of the El-Nino Southern Oscillation (ENSO) and the Pacific Decadal Oscillation. Perica (1998) proposed the integration of climate outlook as well as meteorological forecast into the ESP scheme by a pre-ESP additive/multiplicative adjustment to daily meteorological series. Stendinger and Kim (2010) adjusted probabilities of input series based on a frequency model of the unconditional and conditional distributions of key variables. Clark and Hay (2004), Gobena and Gan (2010) incorporated numerical weather forecast ranging from 14 days to 3 months into ESP in North America and achieved a significant improvement of the forecast efficiency. Moreover, Gobena and Gan (2010) stressed the limitations of ESP that concern the cases of relatively small number of historical weather series. The problem of statistical reliability of the probabilistic forecast results (caused by a limited number of ensemble members) was also recognised by Hamlet and Lettenmaier (1999). In contrast to a pre-ESP meteorological data treatment, Werner et al. (2004) weighted output hydrological ensembles according to the ENSO. Several post-ESP weighting schemes based on climate signal were published by Najafi et al. (2012). Kim et al. (2006), Bohn et al. (2010) investigated the use of multiple hydrological models, instead of a single one, concerning the ESP scheme. In order to represent the forecasted hydrological uncertainty more accurately, the output hydrological ensembles bias correction was proposed by Wood and Schaake (2008). Eum and Kim (2010) achieved better results than the conventional ESP approach in monthly forecasts by updating the observed values of meteorological variables twice a month. The role of initial hydrologic conditions (hereinafter IHC) was investigated using a reverse ESP approach by Wood and Lettenmaier (2008), Yossef et al. (2013).

This study introduces a new modification of the ESP method based on the replacement of the input historical meteorological series by synthetic weather datasets. These synthetic datasets should provide a sufficient number of weather series necessary to construct a probabilistic hydrological forecast when the observed meteorological series are not long enough, as pointed out by Hamlet and Lettenmaier (1999), Gobena and Gan (2010). This is convenient namely when there is an intention to restrict the number of input series based on climate/meteorological outlook or in the regions where the observed meteorological series of the required length are not available. In addition, in order to benefit from the sufficient number of synthetic weather series, these input datasets are restricted in number according to the values of large-scale climatic patterns and variables. The ESP approach was chosen because of the data availability and an intention to introduce a relatively easy operating forecasting system, which will be understandable for forecasters. Moreover, it does not require weather series downscaling (from numerical weather/climate forecasts operating on a significantly large scale), which will be essential namely for the small-scale catchments where the variability of meteorological data is an important factor.

2. Material and methods

2.1. Study area and data

The upper Cidlina River basin (CRB) is a tributary of the Elbe River in eastern Bohemia, Czech Republic (Fig. 1). The river basin area is 455 km^2 with elevations ranging from 225 m to 678 m. The average annual precipitation attains 659 mm and the average annual temperature is 9.0 °C. Arable land comprises 58% of the basin and 22% is forested. The agricultural character of the CRB has a significant influence not only on a hydrological regime but also on water quality measures (Pivokonský et al., 2001). The hydrological regime is characterised by the occurrence of low flow periods in autumn, therefore the estimation of the future water volume is important for water supply management. The basin is closed by the Nový Bydžov gauging station ($Q_a = 2.16\ m^3\ s^{-1}$).

The primary data used in this study were obtained from the Czech Hydrometeorological Institute database and consists of daily precipitation, temperature, solar radiation, humidity, and discharge. Precipitation is measured by nine rain gauges and the other variables are available from three climatic stations in the area of CRB. The duration of the datasets is 1990–2010. The period was split up into 17 years of calibration (1990–2006) and 4 years of forecasting periods (2007–2010). The length of the periods was a compromise determined by the requirements of the weather generator and the data availability. In February and March 1998 the discharge series were not available.

2.2. Forecasting system

The forecasting system is based on the ESP scheme developed at the NWS by Day (1985). The proposed modified ESP (hereinafter mESP) forecasting scheme is depicted in Fig. 2. The forecasting scheme consists of two fundamental parts: climate forecast and hydrological forecast. First, the climate forecast identifies the relations between large-scale climatic predictors (e.g. teleconnection patterns) and local meteorological variables (monthly mean areal precipitation/temperature). Second, synthetic 250 one-year ensembles of meteorological series are generated by LARS-WG (Semenov et al., 1998), based on observed historical series. The number of these synthetic datasets is reduced according to the climate forecast limitations (monthly MAP/MAT). Finally, only those meteorological series that are in agreement with the climate forecast are submitted to the hydrological model in order to issue the hydrological forecast. In this study, the mESP approach is used for the hydrological forecasts with a lead time of one month. The lead time was determined by the available climate forecast.

The proposed climate forecast with the one-month lead time comprises following steps:

- STEP 1 – Preparation of all necessary historical meteorological series (precipitation and temperatures) and large-scale climatic patterns (e.g. teleconnection patterns).
- STEP 2 – Calculation of mean monthly values. Mean areal precipitation (MAP) and mean areal temperature (MAT) are calculated using the Thiessen polygons.
- STEP 3 – Construction of the climate forecast based on data categorisation (Jin et al., 2005) into three subsets: above normal, near normal and below normal. The climate forecast benefits from the previously identified relations (Šípek, 2013) between the local meteorological series and the variety of large-scale climatic predictors.