Hydrology and
Earth System
Sciences

Open Access

Discussions

Attribution of hydrologic forecast uncertainty within scalable forecast windows

L. Yang[1], F. Tian[1], Y. Sun[1], X. Yuan[2], and H. Hu[1]

[1]State Key Laboratory of Hydro-science and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China
[2]Department of Civil and Environmental Engineering, Princeton University, New Jersey 08540, USA

Correspondence to: F. Tian (tianfq@tsinghua.edu.cn)

Abstract

Hindcasts based on the Extended Streamflow Prediction (ESP) approach are carried out in a typical rainfall-dominated basin in China, aiming to examine the roles of initial condition (IC), future atmospheric forcing (FC) and hydrologic model uncertainty (MU) in the streamflow forecast skill. The combined effects of IC and FC are explored within the framework of a forecast window. By implementing virtual numerical simulations without the consideration of MU, it is found that the dominance of IC could last up to 90 days in dry season, while its impact gives way to FC for lead times exceeding 30 days in the wet season. The combined effects of IC and FC on the forecast skill are further investigated by proposing a dimensionless parameter (β) that represents the ratio of the total amount of initial water storage and the incoming rainfall. The forecast skill increases exponentially with β, and varies greatly in different forecast windows. Moreover, the influence of MU on forecast skill is examined by focusing on the uncertainty of model parameters. Two different hydrologic model calibration strategies are carried out. The results indicate that the uncertainty of model parameters exhibits a more significant influence on the forecast skill in the dry season than in the wet season. The ESP approach is more skillful in monthly streamflow forecast during the transition period from wet to dry than otherwise. For the transition period from dry to wet, the low skill of the forecasts could be attributed to the combined effects of IC and FC, but less to the biases in the hydrologic model parameters. For the forecasting in dry season, the usefulness of the ESP approach is heavily dependent on the strategy of the model calibration.

1 Introduction

Reliable hydrologic forecasts are crucial in many hydrologic sectors, e.g., flood control, irrigation, water supply, etc. Forecast skill is mainly affected by three factors, i.e., the uncertainty of hydrologic models used to derive the streamflow from atmospheric forc-

ing (precipitation, temperature, etc.), the uncertainty of initial conditions of the basin at the beginning of the forecast, and the uncertainty of atmospheric forcing during the forecast horizon. These factors are referred to as the "uncertainty triplet" (Zappa et al., 2011). Examination of the three types of uncertainty and their joint impacts on the forecast skill is therefore not only beneficial for the understanding of rainfall-runoff processes, but also for facilitating hydrologic applications with processed-based hydrologic models.

Previous studies revealed that the dominance of IC and FC in the hydrologic forecast skill varies with season and location (Mahanama et al., 2008, 2011; Shukla and Lettenmaier, 2011; Singla et al., 2012; Maurer and Lettenmaier, 2003). For hydrologic forecasts, IC is frequently referred to as the initial water storage (e.g., soil moisture, snow water equivalent) of the entire basin. Li et al. (2009) found that IC dominates the forecast skill with lead time of up to 1 month, where beyond that FC becomes the main contributor to the forecast skill. However, for some basins over the US, initial soil moisture contributes significantly to the forecast skills of all seasons except spring, and the contribution could last up to six months (Mahanama et al., 2011). The dominance of IC is related to the persistence of soil moisture and/or snow water equivalent (SWE). It is not surprising that SWE dominates the forecast skill over those basins of which streamflow are mainly generated by the snow melt (Koster et al., 2010). For a short lead time (within the concentration time of the basin) or under a dry period, soil moisture is the only source of streamflow observed at the outlet of the basin (Kirkby, 1978). The persistence of IC is also affected by the future atmospheric conditions. Wood and Lettenmaier (2008) found that IC yields forecast skill for up to five months during the transition between the wet and dry seasons, but for the reverse transition, FC is critical. IC is also proved to have a stronger impact during the inter-monsoon seasons in Sri Lanka (Mahanama et al., 2008). More recently, Shukla et al. (2013) examined the relative roles of IC and FC in seasonal hydrologic forecast at a global scale. Their results are consistent with previous studies. Mahanama et al. (2011) defined a param-

eter κ (variance ratio of initial water storage and total rainfall within forecast period) to combine the effect of IC and FC, allowing the estimates of the potential forecast skills.

In this study, we will continue with the discussion of the impact of IC and FC on hydrologic forecast by implementing hindcasts with the ESP (Extended Streamflow Prediction) approach (Day, 1985). In particular, due to the persistence of IC and the impact of FC, we believe that there should be certain proper time ranges which we termed as "forecast windows" (see the definition in Sect. 3), and beyond the ranges the forecast is not reliable any more. The impact of IC and FC will be discussed under the framework of the "forecast window". In addition, we will try to combine the effect of the two factors by defining a new dimensionless parameter, based on which the relationships of forecast skill and the combined effect of IC and FC could be derived.

The third member of the "uncertainty triplet" is model uncertainty (MU). The rainfall-runoff model based forecasting approach (e.g. ESP) requires the model to provide accurate initial conditions of the basin at the beginning of the forecast. The forecast skill will undoubtedly be impaired with model uncertainty (Walker et al., 2003; Demirel et al., 2013). MU could be induced by systematic biases in model input, the structure of the model, and model parameters etc. (Walker et al., 2003). While errors in the model input and structure may not be readily reduced, we will only focus on the uncertainty of model parameters in this study through a process of model calibration, assuming that the input as well as the model structure is robust. Shi et al. (2008) demonstrated the influence of hydrologic model calibration in the improvement of seasonal streamflow forecasting.

The analyses in Sect. 3 are based on the following hypotheses: (1) forecast skill varies with initial forecast date and forecast window; (2) the combined effect of IC and FC determines the reliable forecast window in hydrologic forecasting; (3) hydrologic model calibration should be a multi-objective optimization problem for reliable simulations of both high flows and low flows. We will test those hypotheses over the upper Hanjiang River basin (UHRB) in China by examining the relative roles of IC, FC and MU on hydrologic forecast. UHRB is a typical rainfall-dominant river basin (see descrip-

tions in Sect. 2), which is also the headwater of South-to-North Water Transfer project (SNWT) in China. This study will broaden the application fields of ESP approach by testing its validity in rainfall-dominant rather than snow-dominant basins. The results will shed light on future application of ESP approach in similar basins.

5   The paper is organized as follows. Study area and methods will be described in Sect. 2; Sect. 3 will present results, with conclusions and discussion following in Sect. 4.

2 Study area and methods

2.1 Study area

The study area is the upper Hanjiang River basin (abbreviated as UHRB below, see
10   Fig. 1 for its geographic location), which is a sub-basin of the Yangtze River in China. The drainage area of UHRB is 9.52×10^4 km^2. It is a typical monsoon-climate region, characterized by the summer dominant rainfall within the year and great distinctions between rainy and dry seasons (Guo et al., 2009). Figure 2 shows the intra-annual cycle of monthly rainfall and runoff averaged from 1970–2000 over UHRB. Clearly, the
15   runoff regime in this basin is dominated by the rainfall pattern. The integral runoff from May through October (the flood-season period) accounts for about 80 % of the annual total. Based on the rainfall and runoff regimes presented in Fig. 2, we divided a water year into four stages, i.e., pre-wet season (May to July), post-wet season (August to October), pre-dry season (November and December) and post-dry season (January to
20   April of the following year). The distinct variations of the basin state and rainfall pattern within the four stages have some implications in the forecast skill of the ESP approach, as will be discussed in the next section.

2.2 ESP approach

ESP is a widely used approach for hydrologic forecasting (Werner et al., 2004), and
25   usually serves as a reference for validating climate model-based seasonal hydrologic

prediction (Wood et al., 2005; Luo and Wood, 2008; Yuan et al., 2013). The basic idea of ESP is to run a candidate hydrological model with observed meteorological forcing through a spin-up period to the time of the forecast. Then, the model with the spun-up initial basin state is driven by an ensemble of forcing (precipitation, temperature, etc.) that is randomly sampled from the observed historical records (Day, 1985). An ensemble of the streamflow traces is then generated, containing the information of forcing uncertainty on the forecast. The arithmetic mean value of the ensemble forecasts is selected as the issued forecast, simply because the "mean" value is statistically regarded as an unbiased estimator of a sample (DeGroot and Schervish, 2012).

2.3 Model setup

The hydrological model used in this study is Tsinghua-Representative Elementary Watershed model (referred to as THREW model) (Tian, 2006). It is a semi-distributed hydrological model, based on the theory of representative elementary watershed raised by Reggiani et al. (1999). The model consists of a set of balance equations for mass, momentum, energy and entropy, including associated constitutive relationships for various exchange fluxes, at the scale of a well-defined spatial domain. Details of the model can be found in Tian (2006, 2009). The THREW model has been successfully applied in several previous studies (Mou et al., 2008; Yang et al., 2013; Li et al., 2012; Liu et al., 2012; Tian et al., 2006, 2012).

The study period is 1970–2000, which is divided into two parts with the purpose of model calibration and validation. The calibration period is 1970–1980 and the rest is used for model validation. The model calibration procedure is as follows: initial values and the reliable ranges of each parameter are determined according to the physical attributes of UHRB and previous THREW modeling experience (see Sun et al. (2013) for the information of the key parameters in the model). An automatic optimization algorithm, ε-NSGAII (Reed et al., 2003; Deb et al., 2002), is then used for further calibration. The value of each parameter is finally determined based on the automatic calibration results. This calibration procedure enables the parameters to bear clear

physical meanings and maintains the performance of the model as well. The objective function for the automatic calibration is the Nash–Sutcliff coefficient (NSE) (Nash and Sutcliffe, 1970), as widely used in previous hydrological modeling studies.

Another water balance related metric WB, together with NSE, is used to evaluate the performance of the model during the two periods. WB takes the form as:

$$WB = \frac{R_{obs} - R_{sim}}{R_{obs}} \tag{1}$$

where R_{obs} and R_{sim} are the total observed and simulated runoff (in mm), respectively.

The statistics are summarized in Table 1. The THREW model performs quite well in UHRB. The values of daily NSE in both the calibration and validation period are above 0.80. The value of monthly NSE is as high as 0.99 (see Yang et al. (2013) for more detailed evaluations of the THREW model performance in UHRB). The evaluation statistics indicate that the model accurately captures the dynamics of the streamflow during 1970–2000 over UHRB, given the observed atmospheric forcing.

However, this lumped calibration strategy might only present the overall performance of the model during the whole simulation period. It is not necessary to guarantee the performance over each sub-period (e.g., dry seasons). We will present further details of another model calibration strategy used in this study in the next section.

3 Results

In this section, we present the evaluation of the impacts of IC and FC on the forecast within different forecast windows (see Sect. 3.1 for the definition of forecast window) and at two contrasting initial states. The combined effects of IC and FC is then discussed by defining a new parameter β. The effect of MU (the uncertainty of model parameters) will be examined in Sect. 3.2.

3.1 The impact of IC and FC

3.1.1 Definition of forecast window (FW)

A forecast window (FW) proposed in this study can be regarded as an integration time window initiating from the forecast date. It differs from lead time, which is a frequently used term in hydrologic forecasting (see Fig. 3). Lead time is the gap between the time that the forecast is issued and the occurrence of the forecasted variable. For example, if we are interested in the total streamflow volume of this July, supposing the forecast time is issued some day in January, then the lead time should be about 6 months (the time gap between January and July). The model needs to run from January till the end of July. In this context, this study focused on the cases with no lead time (equal to zero). They could be regarded as "real-time" forecasts but with different FWs. The forecasted variable is the integral streamflow volume within each FW.

3.1.2 Evaluation of the impact of IC and FC

"Virtual" experiments are designed in this section for the reason of avoiding the incorporation of model uncertainty (MU) effect. In these "virtual" experiments, the forecasts are evaluated against retrospective streamflow simulations (driven by actual atmospheric forcing) instead of actual streamflow observations, assuming that the model is "perfect". The design of the experiment is summarized in Table 2. Eight different forecast windows are set: 7 days, 15 days, 30 days, 60 days, 90 days, 120 days, 150 days and 183 days. We also choose two contrasting forecast dates, 1 February and 1 July of each year, representing the dry and wet initial states of the basin, respectively. The forecasts are made at two initial dates for each year with 8 different forecast windows based on the ESP approach.

The results are evaluated by using the Pearson correlation coefficient ρ. ρ decreases with the extension of forecast window for both initial dates (Figs. 4 and 5). However, the patterns are a little different. For the dry scenario (the initial date is 1 February), ρ

is 0.99 for the 7 day window, and gradually drops below 0.50 for the window exceeding 90 days; while for the wet scenario (the initial date is 1 July), the maximum correlation coefficient is 0.70 for the 7 day forecast window, and quickly reduces to 0.46 for 30 day window. The forecasts almost equal the climatological mean when the forecast window exceeds 90 days under this scenario. The contrasting behaviors of the two scenarios indicate that the consistency of the forecast seems to bear a wider range of forecast windows for the dry scenario. Therefore, the performance of ESP is a function of initial date and the forecast window.

To further illustrate the impact of IC/FC and their relative role in the forecast, we employed an analytical framework, as developed by (Wood and Lettenmaier, 2008) and used by Li et al. (2009). It is effective to discern relative impacts of IC and FC on the forecasts, as well as the dynamic competence of the two factors with the increase of forecast window. The basic idea of this framework is to re-sort the forecasts according to IC (soil moisture) and FC (precipitation forcing). 2-D images are produced based on the resorted matrix of the forecasts. The columns of each 2-D image in Figs. 6 and 7 can be regarded as a reverse-ESP (R-ESP) approach which was introduced by Wood and Lettenmaier (2008) (see also Fig. 2 of Li et al. (2009) for more details of R-ESP and ESP). The basic idea of R-ESP is to drive an ensemble of ICs, which are derived by resampled meteorological ensembles during the spinup period, using the accurate meteorological forcing. R-ESP reveals the influence of IC uncertainty on the forecast, while ESP focuses on the impact of FC. The patterns of the images reflect the relative impact of IC and FC: if these are horizontally structured, the forecasts are largely determined by IC, and FC has a relatively small impact; while FC dominates the forecast if the image is vertically structured. A "hatched" image indicates a combined influence of IC and FC.

Organized structures could be observed in both Figs. 6 and 7. For the forecasts initialized on the 1 February (Fig. 6), the patterns of the images are characterized by horizontal stripes for forecast windows less than 60 days (2 months). The patterns change into more vertical stripes when the window exceeds 90 days. For the forecast

of 90 day forecast window, a hatched pattern is displayed. IC is the dominant impact on the forecasts within 2 month forecast window when ESP has most skill; while this dominance changes to FC for larger forecast windows (exceeding 90 days), and ESP loses its skill as well. The 90 day window is the divide of the dominance of the two factors. Both the IC and FC could impact the forecasts within the 90 day forecast window. For the forecasts initialized on the 1 July (Fig. 7), the dominance of IC is constrained to 30 days. For the rest of forecast windows (exceeding 30 days), FC dominates the forecasts and the forecast skill decays as well.

The possible explanations for the results are: the 1 February represents relatively dry state of the basin within the year, since it is in the middle of the dry seasons and the antecedent precipitation is scarce. The lack of soil moisture to the saturation state can not be easily filled up with subsequent rainfall events. Thus, the IC could persist for a long period until the total rainfall within the forecast window is able to compensate the initial soil moisture anomaly. This could explain why the "hatched" image occurs on the 90 day case (Fig. 6) and is converted to "vertical" when flood season (May to August) begins. For the wet scenario (forecasts initialized on the 1 July), the basin has already been saturated or near-saturated. In this case, the accumulated soil moisture in the basin will contribute significantly to the future streamflow regime, just as most snow-dominated basins behave. However, unlike those snow-dominate basins, UHRB will experience rainy weather in the subsequent months (August, September and October). Several heavy precipitation events could easily recharge the basin and eliminate the persistence of the soil moisture completely, so the IC is not able to persist for a long period. It seems that the persistence of IC is not solely determined by the magnitude of initial anomaly, but also by the subsequent meteorological conditions. The forecast skill of the ESP approach is determined by the combined effects of IC and FC, which will be examined by defining a new parameter in the next section.

3.1.3 Combined effects of IC and FC

Based on the analyses in Sect. 3.1.2, we further come up with a new parameter, which tries to synthesize the effect of IC and FC on the forecasts and generalize the relationships to the forecasts with more diverse ICs. The new parameter beta (β) is a dimensionless ratio of total initial water storage and incoming rainfall within forecast windows, defined here as

$$\beta = \log(R_{sm}/R_p) \tag{2}$$

where R_{sm} is the total initial water storage (e.g., soil moisture, residual water in the river, expressed as "depth" and the unit is mm), representing the influence of IC; R_p is the total rainfall (in mm) within the forecast windows, representing the influence of FC. We take the logarithm of the ratio for mathematical reasons. The form of beta (β) is similar with the parameter kappa (κ) defined by Mahanama et al. (2011). However, they focus on the variance instead of the total amount, since the integration window is fixed to 3 months in their study.

We employed MARE (Mean Absolute Relative Error) as the evaluation metric of the forecasts. The form is:

$$\text{MARE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\text{Fst}_i - \text{Obs}_i}{\text{Obs}_i} \right| \times 100\% \tag{3}$$

where Fst_i and Obs_i are the forecasted and observed total streamflow volume within forecast windows, respectively; n is the number of forecast years, and $n = 30$ in this study.

A new set of forecast windows are used, they are: 3 days, 7 days, 10 days, 15 days, 20 days, 30 days, 60 days and 90 days, since ESP approach presents no skill for the forecasts beyond 90 days in this study. The first day of each month is chosen as the forecast date. The details of the experiment are summarized in Table 3. The experiment configurations in this section are also devoid of model uncertainty. An exponential

function ($y = ae^{bx}$) is employed to fit the relationship between the parameter (β) and evaluation metric (MARE) of the forecasts. The parameters and evaluation values (R^2) of the fitted lines are listed in Table 4.

The accuracy of the forecasts increases exponentially with β, corresponding to our above analysis that the forecasts skill will be mainly determined by IC if the initial water storage has a comparatively larger value than the total rainfall within the forecast windows. However, the variability of total precipitation within forecast window also plays a role. As shown in Fig. 8, the forecast accuracy becomes less sensitive to β as the forecast window increases. This is probably because the total rainfall within larger forecast windows enables the basin to be fully recharged several times, which eliminates the persistence of IC. In addition, the inter-annual variability of the total rainfall decays with the integrating time window (e.g., the variability of seasonal precipitation is usually smaller than that of weekly precipitation), which should come as no surprise. Thus, a larger value of total precipitation with smaller variability will induce smaller sensitivity of forecast accuracy to β.

It is also noteworthy that although the fitted lines asymptotically approximate to 0, there will be upper bounds for the forecast accuracy in reality. As can be observed in Fig. 8, the upper bounds of forecast skill for 3 days are much higher than 90 days, indicating that short-term forecasts are potentially more reliable than long-term forecast using the ESP approach. However, we also notice that short-term forecasts are not necessarily more accurate than long-term counterparts (the relatively larger vertical spans for the smaller forecast window lines in Fig. 8). Since the model is assumed to be error free, the possible explanation could be the variability of the total precipitation within forecast windows. For short-term forecast, the variability of FC should also be considered in addition to IC, especially for the cases when FC takes the dominance of the forecast skill.

Figure 8 also shows the potential ability of ESP approach within different forecast windows over UHRB. Although the newly defined parameter (β) only focuses on the ratio of the total water storage and rainfall, the derived relationships between the β

and forecast skill reveal the effect of variance as well. We note that this analytical framework of β is able to combine the effects of IC and FC, and provide a first-order understanding of when and to what extent the forecast skill may be achieved based on the ESP approach in UHRB.

3.2 The impact of MU

Model Uncertainty (MU) is another factor that influences the forecast. The essence of the ESP approach is to utilize the IC of the basin provided by the hydrological model through spin-up period. Thus, the reliability of the IC greatly impacts the final accuracy of the forecast. The model used in this study has been calibrated against NSE, with the evaluation metrics summarized in Table 1. Although the model presents a remarkably good performance during the whole simulation period, its performance in each separate month varies. As shown in Fig. 9b, the value of NSE is above 0.90 for July, August, September and October. These four months are also the wettest during the year in UHRB. For dry months (January to April), NSE is below 0. The value of NSE is 0.58 and 0.64 for March and April, respectively.

Previous studies revealed the problems of model calibration using single objective function. The mathematical form of NSE results in the over-weighting of high flows in the calculation (Schaefli and Gupta, 2007). Even though the simulation of low flow deviate is poor, NSE will not be affected, since the magnitude is small relative to the flood period. In this study, we also carried out a different model calibration strategy. The low flows (streamflow from January to May) are extracted from each year, constituting a new time series. The model is again calibrated with the same calibration procedure, but only for the new low-flow time series. NSE is still used as the calibration metric because of its simplicity. Future studies could investigate the effectiveness of other metrics as suggested by Pushpalatha et al. (2012).

New NSE for each month (January to May) is shown in Fig. 9b. Model performance in March, April and May is improved, while for January and February, the new calibration does not make any improvement. Since the objective of this study is to examine

the influence of MU (the uncertainty of model parameters) on the forecast accuracy of ESP approach, we will not endeavor to improve the model performance for each month. The comparisons of the forecasts made on March and April will be focused. We set the forecast window to 30 days, which is frequently used in previous studies (e.g. Smith et al., 1992; Hashino et al., 2006; Wang et al., 2011) and has been proved to be effective under different ICs in previous section.

The skill of the forecast in ESP approach is evaluated by SS_{MSE}, defined as

$$SS_{MSE} = 1 - \frac{\sum_{i=1}^{n}(Fst_i - Obs_i)^2}{\sum_{i=1}^{n}(Obs_i - \overline{Obs})^2} \qquad (4)$$

Where Obs_i and Fst_i are the observed and forecasted monthly streamflow volume (in mm), respectively; $\overline{Obs}$ is the mean value of the observed streamflow volume (in mm), representing the climatological forecast; n is the total number of forecast years, and $n = 30$ in this study. This has the same form as Nash–Sutcliff Efficiency, but is being used for assessing the model performances in individual months. A score of 1 corresponds to a perfect forecast.

During October to December, ESP approach presents high forecast skill (Fig. 9a). For March and April, the skill score is far below 0, indicating no forecast skill for this period. However, this does not indicate the uselessness of ESP approach. When the forecasts are made based on the re-calibrated models (especially for March and April), the skill scores are significantly improved. The improvement of the forecast skill in March and April reinforces our hypothesis that MU (errors in model parameters) influences the forecast accuracy in ESP approach, especially for those flow events that are difficult simulated by hydrologic models.

It is noteworthy that the forecast skill score is low in May, June and July, although the model performs quite well in this month. The low skill of the forecasts in May, June and July implies that the ESP approach is no better than climatology for the monthly

streamflow forecast of pre-wet seasons in UHRB, which probably could be attributed to the combined effects of FC and IC in this season (as discussed in previous subsections). For post-wet season (August, September and October) and pre-dry season (November and December), the ESP approach is useful. Our results are similar to the study of Wood and Lettenmaier (2008), which showed that the forecast skill is higher during the transition period from the wet to dry season than otherwise. However, our analyses also show that the forecast skill is improved in post-dry season (March and April, especially) after re-calibrating the hydrologic model used in the ESP. The forecast skill in January and March is still comparable with that in post-wet and pre-dry seasons, indicating the potential of ESP approach with well-calibrated hydrologic model. Future efforts should focus on the improvement of the model performance in the post-dry season period (January to April).

4 Conclusions and discussion

In this paper, we have investigated the influence of IC, FC and MU on hydrologic forecast based on the hindcasts using ESP approach over UHRB in China. The concept of forecast window (FW) is introduced in this study and the combined effects of IC and FC within the forecast window is discussed by implementing "virtual" experiments without considering the uncertainty of hydrologic model (structure, parameter, etc.). The improvement of the model performance also increases the skill of the ESP approach significantly for low flow forecast, indicating the impact of MU (the uncertainty of model parameters) on hydrologic forecast. The results are summarized below.

(1) For dry initial states, the forecast is consistently good within forecast windows up to 90 days. For wet initial states the persistence decays with the window exceeding 30 days. The contrasting behaviors of the two scenarios highlight the role of IC and FC on the forecast.

(2) IC controls the forecast skills within smaller forecast windows. The dominance of IC on the forecast is shifted to FC with the increase of the forecast windows. A dimen-

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

sionless parameter β is proposed to depict the combined effects of IC and FC. The analytical framework of β provides a first-order understanding of when and to extent the forecast skills may be achieved based on ESP approach.

(3) Traditional model calibration strategy emphasizes the model behavior during the wet season, while for dry season the model behaviors would not be guaranteed. This affects the accuracy of ESP approach for predicting low flows. By re-calibrating the model for low flows, the uncertainty of model parameters is reduced within the dry season, which improves the model performance as well as the forecast skill.

(4) The performance of ESP approach for monthly streamflow forecast is more reliable during the transition period from wet to dry than otherwise over UHRB. For the transition period from dry to wet, the combed effects of IC and FC rather than the model performance play a more important role in the forecast skill. Other innovative ways should be explored for monthly streamflow forecast during this season.

According to the best of the authors' knowledge, the concept of forecast window (FW) is first proposed in this study. We suggest that the concept of forecast window (FW) might have some implications in the hydrologic forecasting based on the ESP approach. The dominance of IC and FC within FWs reflects the potential ability of ESP. For different scenarios with various combinations of IC and FC, there should be certain "behavioural forecast windows" (BFWs), beyond which ESP-based hydrologic forecasts are not accurate and not reliable. In addition, BFW can also be related to the physical attributes of each basin (e.g., sizes, maximum concentration time, etc.). Li et al. (2009) disclosed that the basin size matters in the ESP skill. Thus, the relationship we derived in this study is only basin-specific. More basins with diverse physical attributes need to be examined, so as to generalize the relationship obtained in this study, which, however, is the scope of future studies. BFW is crucial in the practical implementation of ESP approach by providing a guideline for the design of forecasting systems and the confidence of the forecast results. However, the framework designed to examine FW and BFW in this study does not consider lead time in the forecasts. They could be regarded as zero lead time scenarios. Besides, the influence of MU is

not considered in the framework. Future studies will further examine the behaviors of FW and BFW with varying lead time and model uncertainty (structure and parameter) incorporated.

There are a number of ways (e.g., pre-processing, post-processing) to further improve the performance of ESP approach. For instance, Yang et al. (2013) improved the forecast skills of ESP approach by using a reduced set of the ensemble members. The members in the reduced ensemble were selected from the historical records when they have the same climate signals (e.g., SOI, PDO) as the forecast year. Similar studies could be found in Hamlet and Lettenmaier (1999), Lamb (2010) and Wang et al. (2011). Instead of using the randomly sampled members, the "pre-processing" process enhances the representativeness of future forcing ensembles. Similarly, "post-processing" aims to remove the bias of the streamflow ensembles, which could also improve the forecast skill. Possible "post-processing" methods were evaluated in previous studies (e.g. Kang et al., 2010; Wood and Schaake, 2008; Hashino et al., 2006). Shi et al. (2008) demonstrated that "post-processing" could reduce the forecast error as much as the correction of forecast errors through hydrologic model calibration. Yuan and Wood (2012) also found that post-processing streamflow forecasts directly from a global climate forecast model, where its land surface model is un-calibrated, has comparable performance to a well-calibrated hydrologic model driven by downscaled and bias-corrected meteorological forcing.

References

Day, G.: Extended streamflow forecasting using NWSRFS, J. Water Res. Pl.-ASCE, 111, 157–170, 1985.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, evolutionary computation, IEEE T. Evolut. Comput., 6, 182–197, doi:10.1109/4235.996017, 2002.

DeGroot, M. H. and Schervish, M. J.: Probability and Statistics, Addison-Wesley Pub. Co., Boston, MA, 2012.

Demirel, M. C., Booij, M. J., and Hoekstra, A. Y.: Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models, Water Resour. Res., 49, 4035–4053, doi:10.1002/wrcr.20294, 2013.

Guo, S., Guo, J., Zhang, J., and Chen, H.: VIC distributed hydrological model to predict climate change impact in the Hanjiang basin, Sci. China Ser. E, 52, 3234–3239, 2009.

Hamlet, A. and Lettenmaier, D.: Columbia river streamflow forecasting based on ENSO and PDO climate signals, J. Water Res. Pl.-ASCE, 125, 333–341, 1999.

Hashino, T., Bradley, A. A., and Schwartz, S. S.: Evaluation of bias-correction methods for ensemble streamflow volume forecasts, Hydrol. Earth Syst. Sci., 11, 939–950, doi:10.5194/hess-11-939-2007, 2007.

Kang, T. H., Kim, Y. O., and Hong, I. P.: Comparison of pre- and post-processors for ensemble streamflow prediction, Atmos. Sci. Lett., 11, 153–159, 2010.

Kirkby, M. J.: Hillslope Hydrology, John Wiely, Hoboken, New Jersey, 1978.

Koster, R. D., Mahanama, S. P. P., Livneh, B., Lettenmaier, D. P., and Reichle, R. H.: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow, Nat. Geosci., 3, 613–616, 2010.

Lamb, K.: Improving Ensemble Streamflow Prediction Using Interdecadal-Interannual Climate Variability, Doctoral thesis, University of Nevada, Las Vegas, 2010.

Li, H., Luo, L., Wood, E. F., and Schaake, J.: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting, J. Geophys. Res.-Atmos., 114, D04114, doi:10.1029/2008JD010969, 2009.

Li, H., Sivapalan, M., and Tian, F.: Comparative diagnostic analysis of runoff generation processes in Oklahoma DMIP2 basins: the Blue River and the Illinois River, J. Hydrol., 418–419, 90–109, 2012.

Attribution of hydrologic forecast uncertainty

L. Yang et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

|◀| ▶|

◀   ▶

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Liu, D., Tian, F., Hu, H., and Hu, H.: The role of run-on for overland flow and the characteristics of runoff generation in the Loess Plateau, China, Hydrolog. Sci. J., 57, 1107–1117, 2012.

Luo, L. and Wood, E. F.: Use of Bayesian merging techniques in a multimodel seasonal hydrologic ensemble prediction system for the eastern United States, J. Hydrometeorol., 9, 866–884, 2008.

Mahanama, S. P. P., Koster, R. D., Reichle, R. H., and Zubair, L.: The role of soil moisture initialization in subseasonal and seasonal streamflow prediction – a case study in Sri Lanka, Adv. Water Resour., 31, 1333–1343, 2008.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., and Reichle, R.: Soil moisture, snow, and seasonal streamflow forecasts in the United States, J. Hydrometeorol., 13, 189–203, 2011.

Maurer, E. P. and Lettenmaier, D. P.: Predictability of seasonal runoff in the Mississippi River basin, J. Geophys. Res.-Atmos., 108, 8607, doi:10.1029/2002JD002555, 2003.

Mou, L., Tian, F., Hu, H., and Sivapalan, M.: Extension of the Representative Elementary Watershed approach for cold regions: constitutive relationships and an application, Hydrol. Earth Syst. Sci., 12, 565–585, doi:10.5194/hess-12-565-2008, 2008.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models Part I – A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Pushpalatha, R., Perrin, C., Moine, N. L., and Andréassian, V.: A review of efficiency criteria suitable for evaluating low-flow simulations, J. Hydrol., 420–421, 171–182, 2012.

Reed, P., Minsker, B. S., and Goldberg, D. E.: Simplifying multiobjective optimization: an automated design methodology for the nondominated sorted genetic algorithm-II, Water Resour. Res., 39, 1196, doi:10.1029/2002WR001483, 2003.

Reggiani, P., Hassanizadeh, S. M., Sivapalan, M., and Gray, W. G.: A unifying framework for watershed thermodynamics: constitutive relationships, Adv. Water Resour., 23, 15–39, 1999.

Schaefli, B. and Gupta, H. V.: Do Nash values have value?, Hydrol. Process., 21, 2075–2080, 2007.

Shi, X., Wood, A. W., and Lettenmaier, D. P.: How essential is hydrologic model calibration to seasonal streamflow forecasting?, J. Hydrometeorol., 9, 1350–1363, 2008.

Shukla, S. and Lettenmaier, D. P.: Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill, Hydrol. Earth Syst. Sci., 15, 3529–3538, doi:10.5194/hess-15-3529-2011, 2011.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Shukla, S., Sheffield, J., Wood, E. F., and Lettenmaier, D. P.: On the sources of global land surface hydrologic predictability, Hydrol. Earth Syst. Sci., 17, 2781–2796, doi:10.5194/hess-17-2781-2013, 2013.

Singla, S., Céron, J.-P., Martin, E., Regimbeau, F., Déqué, M., Habets, F., and Vidal, J.-P.: Predictability of soil moisture and river flows over France for the spring season, Hydrol. Earth Syst. Sci., 16, 201–216, doi:10.5194/hess-16-201-2012, 2012.

Smith, J., Day, G., and Kane, M.: Nonparametric framework for long-range streamflow forecasting, J. Water Resour. Pl.-ASCE, 118, 82–92, 1992.

Sun, Y., Tian, F., Yang, L., and Hu, H.: Attributing streamflow decrease to variations in climate conditions and catchment properties across the Upper Han River basin, China, J. Hydrol., in review, 2013.

Tian, F.: Study on Thermodynamic Watershed Hydrological Modeling (THModel), Doctoral thesis, Tsinghua University, Beijing, China, 168 pp., 2006.

Tian, F., Hu, H., Lei, Z., and Sivapalan, M.: Extension of the Representative Elementary Watershed approach for cold regions via explicit treatment of energy related processes, Hydrol. Earth Syst. Sci., 10, 619–644, doi:10.5194/hess-10-619-2006, 2006.

Tian, F., Mou, L., and Hu, H.: Characteristic of soil water retention curve on the macroscale, Sci. China Ser. E, 52, 2990–2996, 2009.

Tian, F., Li, H., and Sivapalan, M.: Model diagnostic analysis of seasonal switching of runoff generation mechanisms in the Blue River basin, Oklahoma, J. Hydrol., 418–419, 136–149, 2012.

Walker, W. E., Harremoes, P., Rotmans, J., Sluijs, J. P. v. d., Asselt, M. B. A. v., Janssen, P., and Krauss, M. P. K. V.: Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support, Integrated Assessment, 4, 5–17, 2005.

Wang, E., Zhang, Y., Luo, J., Chiew, F. H. S., and Wang, Q. J.: Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data, Water Resour. Res., 47, W05516, doi:10.1029/2010WR009922, 2011.

Werner, K., Brandon, D., Clark, M., and Gangopadhyay, S.: Climate index weighting schemes for NWS ESP-based seasonal volume forecasts, J. Hydrometeol. 5, 1076–1090, 2004.

Wood, A. W. and Lettenmaier, D. P.: An ensemble approach for attribution of hydrologic prediction uncertainty, Geophy. Res. Let., 35, L14401, doi:10.1029/2008GL034648, 2008.

Wood, A. W. and Schaake, J. C.: Correcting errors in streamflow forecast ensemble mean and spread, J. Hydrometeorol., 9, 132–148, 2008.

Attribution of hydrologic forecast uncertainty

L. Yang et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◀◀   ▶▶

◀   ▶

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Wood, A. W., Kumar, A., and Lettenmaier, D. P.: A retrospective assessment of national centers for environmental prediction climate model–based ensemble hydrologic forecasting in the western United States, J. Geophys. Res.-Atmos., 110, D04115, doi:10.1029/2004JD004508, 2005.

Yang, L., Tian, F., and Hu, H.: Modified ESP with the information of atmospheric circulation and teleconnection incorporated and its application, J. Tsinghua University (Science and Technology), 53, 606–612, 2013.

Yuan, X. and Wood, E. F.: Downscaling precipitation or bias-correcting streamflow? Some implications for coupled general circulation model (CGCM)-based ensemble seasonal hydrologic forecast, Water Resour. Res., 48, W12519, doi:10.1029/2012WR012256, 2012.

Yuan, X., Wood, E. F., Roundy, J. K., and Pan, M.: CFSv2-based seasonal hydro climatic forecasts over conterminous United States, J. Climate, 26, 4828–4847, doi:10.1175/JCLI-D-12-00683.1, 2013.

Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F.: Superposition of three sources of uncertainties in operational flood forecasting chains, Atmos. Res., 100, 246–262, 2011.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

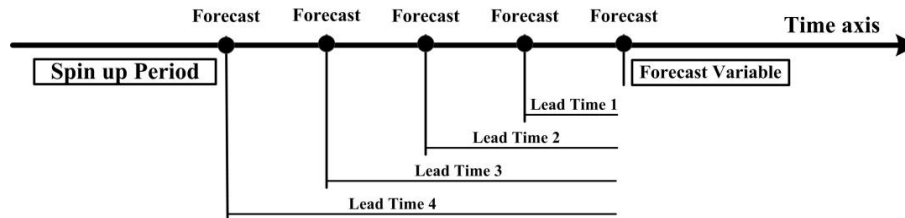Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Calibration and validation statistics for the simulated streamflow of UHRB at Danjiangkou Reservoir station.

	Nash–Sutcliff Efficiency (NSE)			Water Balance (WB)
	Daily	Weekly	Monthly	
Calibration Period (1970–1980)	0.85	0.98	0.99	−0.06
Validation Period (1981–2000)	0.81	0.96	0.99	−0.05

Table 2. Details of the experiment designs (for different forecast window).

Initial Forecast Date	Forecast Windows	Ensemble Members	Hindcast Period
1 Feb	7 days, 15 days, 30 days, 60 days, 90 days, 120 days, 150 days, 183 days	31	1970–2000
1 Jul			

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 3. Details of the experiment designs (for deriving the relationships of β and the forecast skills).

Initial Forecast Date	Forecast Windows	Ensemble Members	Hindcast Period
1 Jan, 1 Feb, . . . 1 Nov, 1 Dec	3 days, 7 days, 15 days, 20 days 30 days, 45 days, 60 days, 90 days	31	1970–2000

Table 4. Parameters of the exponential function and the R^2 value in the relationships of β and the forecast skills.

Forecast Windows	a	b	R^2
3 days	3153	−1.1	0.90
7 days	1055	−0.89	0.92
10 days	561	−0.77	0.96
15 days	270	−0.62	0.92
20 days	155	−0.49	0.81
30 days	134	−0.54	0.82
60 days	67	−0.43	0.71
90 days	50	−0.32	0.56
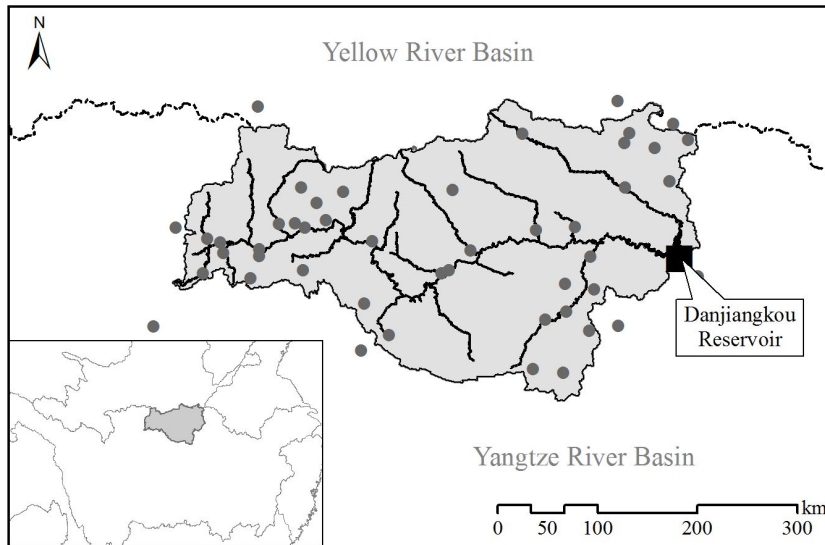
Fig. 1. Location of upper Hanjiang River basin (UHRB, denoted as shaded region in the figure) in China. Dashed line is the boundary of the Yangtze River basin. The grey dots are rain gauges. The black square represents the location of Danjiangkou Reservoir station.
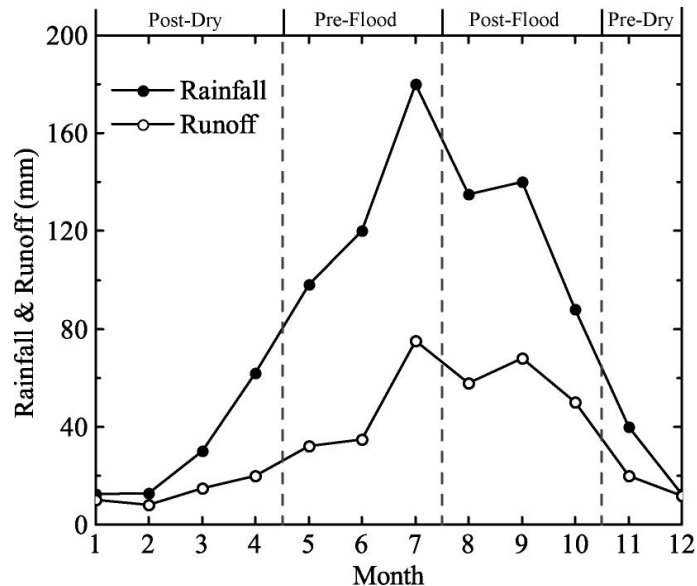
Fig. 2. Rainfall and runoff regimes over upper Hanjiang River basin (UHRB). Runoff is the total amount of monthly inflow to Danjiangkou reservoir divided by basin area. Rainfall is the mean value of all the rain gauges weighted by area. The unit is in mm.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
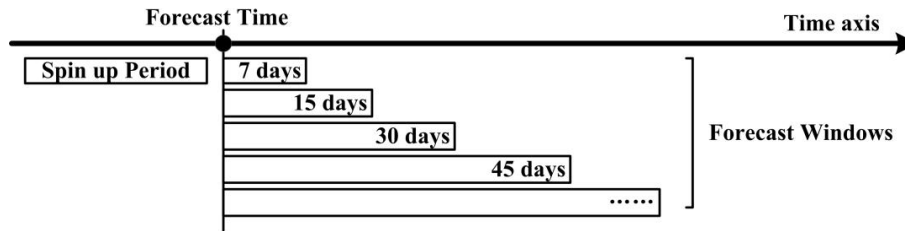
(a) Forecast window

(b) Lead time

Fig. 3. Schematic illustration of **(a)** forecast window and **(b)** lead time.
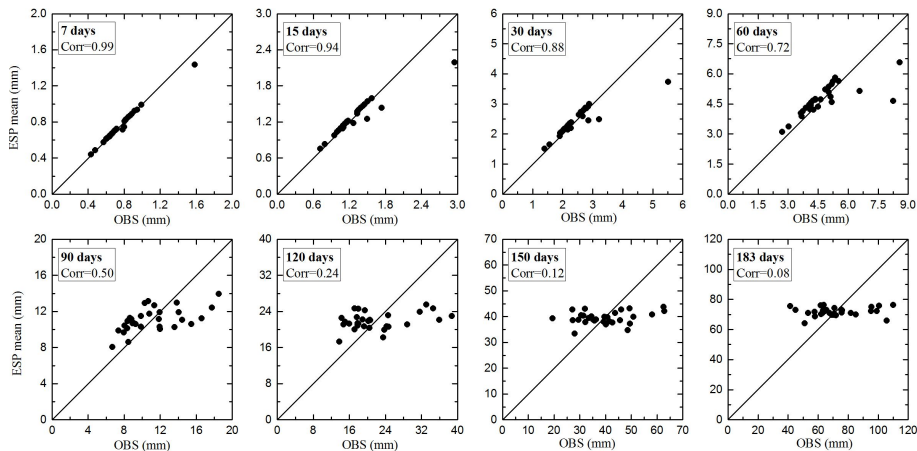
Fig. 4. Evaluation of the ESP approach for different forecast windows. The initial forecast date is 1 February.
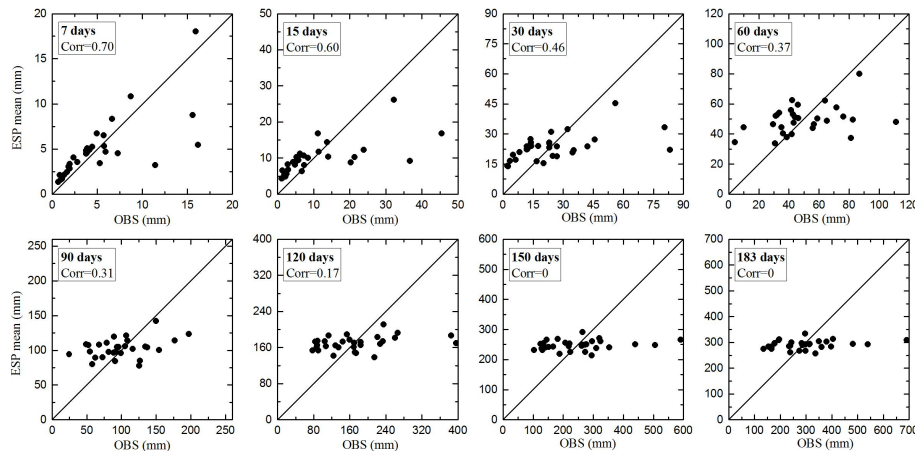
Fig. 5. The same as Fig. 4 but the initial forecast date is 1 July.
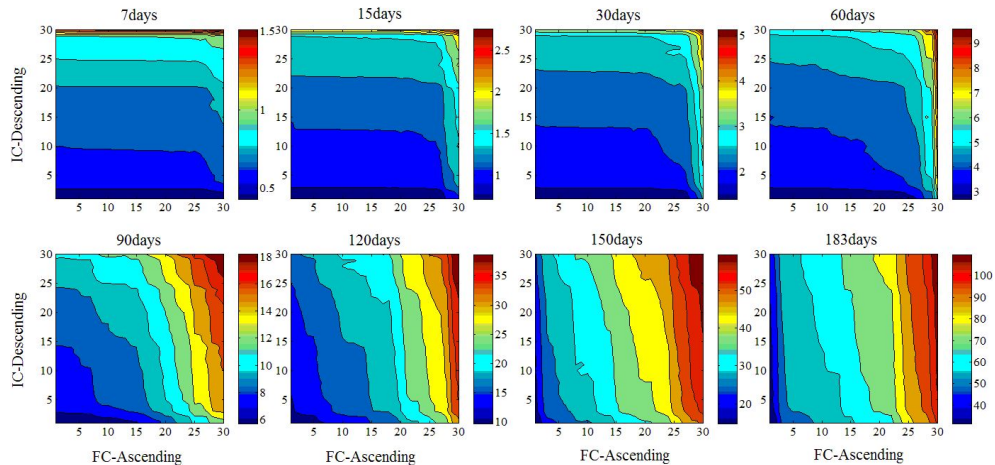
Fig. 6. Runoff forecasts (unit: mm) by ESP (row) and R-ESP (column) for UHRB. Forecasts are initialized on 1 February. Precipitation forcing (units: mm) is sorted from dry to wet in ascending order, and Initial conditions (units: mm) is sorted from wet to dry in descending order.
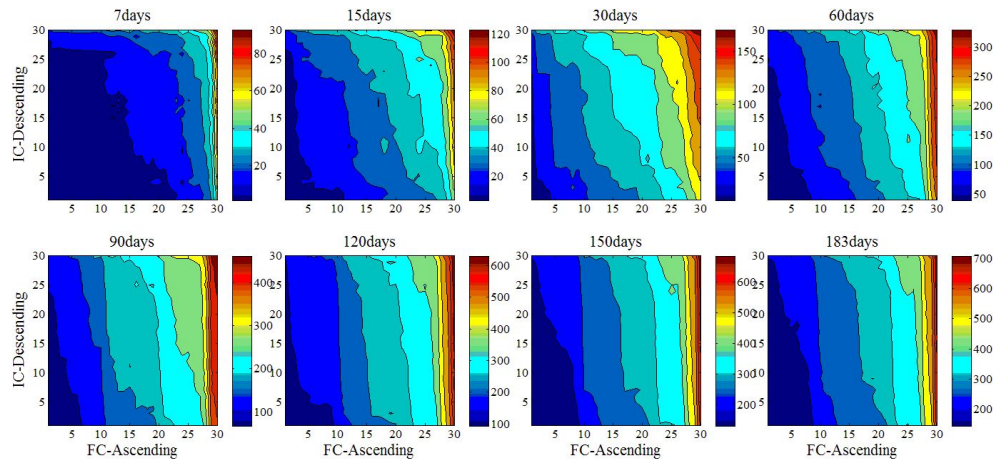
Fig. 7. The same as Fig. 6 but the forecasts are initialized on 1 July.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◀ | ▶|

◀ | ▶

Back | Close

Full Screen / Esc

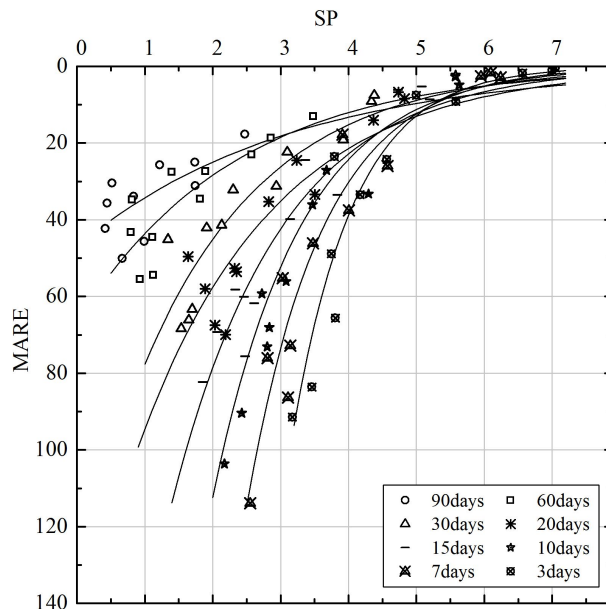Printer-friendly Version

Interactive Discussion

Fig. 8. Scatterplots of β and MARE. The curves are fitted exponential lines for each group of forecasts with the same forecast window. The parameters of the lines are summarized in Table 4.
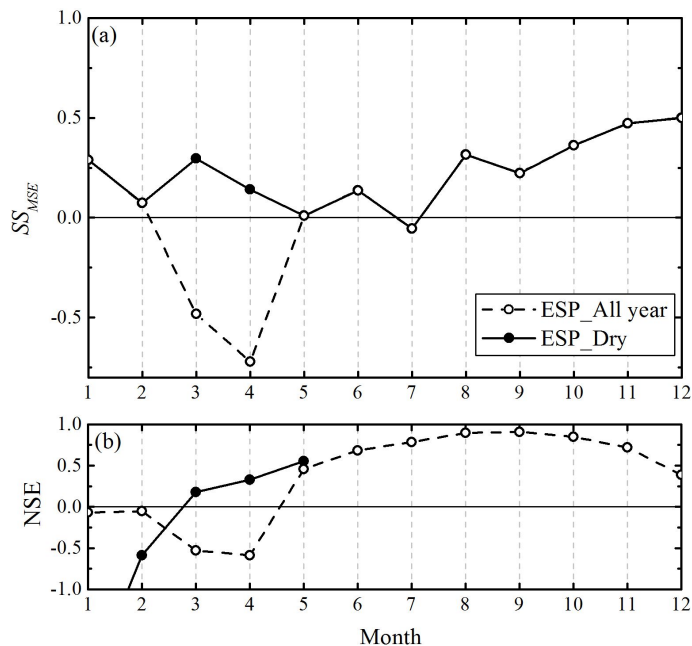
Fig. 9. (a) Forecast skills of ESP based on different sets of model parameters; **(b)** model performances for each month. "ESP_all year" (circles) represent using the default parameters of the model (the whole year). "ESP_Dry" (dots) uses the new-calibrated parameters of the model (only January to May).