Understanding and seasonal forecasting of hydrological drought in the anthropocene

Xing Yuan[1], Miao Zhang[1,2], and Linying Wang[1,2]

[1]RCE-TEA, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[2]University of Chinese Academy of Sciences, Beijing, 100049, China

Correspondence to: Xing Yuan (yuanxing@tea.ac.cn)

Abstract. Hydrological drought is not only caused by natural hydro-climate variability, but can also be directly altered by human interventions including reservoir operation, irrigation and groundwater exploitation, etc. Understanding and forecasting of hydrological drought in the anthropocene are grand challenges due to complicated interactions among climate, hydrology and human. In this paper, five decades (1961-2010) of naturalized and observed streamflow datasets are used to investigate hydrological drought characteristics in a heavily managed river basin, the Yellow River basin in North China. It is found that human interventions increase the nonlinear response of hydrological drought to the meteorological drought, and increase the response time especially during rainy seasons. Due to large water consumptions over the middle and lower reaches, there are two to four-fold increases in the hydrological drought frequency and up to six-fold increases in the drought severity, the drought duration increases by 12-83%, and the drought onset becomes earlier. A set of 29-year (1982-2010) hindcasts from an established seasonal hydrological forecasting system are used to assess the forecast skill of hydrological drought. In the naturalized condition, the climate-model-based approach outperforms the climatology method in predicting the 2001 severe hydrological drought event. Based on the 29-year hindcasts, the former method has a Brier Skill Score of 11%-26% against the latter for the probabilistic hydrological drought forecasting. In the anthropocene, the skill for both approaches increases due to dominant influence of human interventions that have been implicitly incorporated by the hydrological post-processing, while the difference between two predictions decreases. This suggests that human interventions can outweigh the climate variability for the hydrological drought forecasting in the anthropocene, and the predictability for human interventions needs more attention.

1. Introduction

Drought is a natural phenomenon occurring recursively due to climate variability that are associated with oceanic and/or terrestrial anomalies (Hong and Kalnay, 2000; Hoerling and Kumar, 2003). As the rainfall deficit reaches a certain threshold, the meteorological drought occurs. If the meteorological drought persists for a period of time, it will decrease the soil moisture and river flow, resulting in agricultural and hydrological droughts. Although the rainfall deficit is a primary driver for the agricultural and hydrological droughts, terrestrial hydrological processes (e.g., snow melting, evapotranspiration), geological and topographic conditions also play a non-trivial role in the drought propagation (Van Loon et al., 2012; Rimkus

et al., 2013; Teuling et al., 2013; Stoelzle et al., 2014; Staudinger et al., 2015). Therefore, monitoring and forecasting of agricultural and hydrological droughts not only provide more relevant guideline for the management of agricultural and water resources, but also raise challenging science questions on the mechanism and predictability of drought processes (Pozzi et al, 2013; Yuan et al., 2013; Wood et al., 2015).

5   Given that rainfall deficit is a major cause for the hydrological drought, many studies focus on the understanding of the propagation from meteorological to hydrological drought. For instance, Vicente-Serrano and López-Moreno (2005) investigated the relationship between streamflow and antecedent rainfall by using a correlation analysis, and found that hydrological drought index series has the highest correlation with a two-month accumulated meteorological drought index series in a mountainous Mediterranean basin. Such correlation analysis method was then widely used to understand the time

10   scale of hydrological drought (Rimkuset al., 2013; Haslingeret al., 2014; Bloomfield et al., 2015; Folland et al., 2015; Niu et al., 2015; Kumar et al., 2016). A recent study from Barker et al. (2016) comprehensively investigated the relationship between meteorological and hydrological droughts for 121 near-natural catchments in the United Kingdom, where the relationship was found to be associated with natural climate and catchment properties.

Besides natural climate and hydrological processes that affect the development of hydrological drought, human activities

15   such as land use and land cover change, urbanization, irrigation, reservoir operation and groundwater exploitation can also influence hydrological drought significantly (AghaKouchak et al., 2015; Van Loon et al., 2016a). López-Moreno et al. (2009) found that the second largest reservoir in Europe increased the duration and severity of hydrological drought over downstream areas. Similar studies found that the reservoir regulation might reduce the drought severity over upstream areas but increase it over downstream areas over Australian and Chinese catchments (Wen et al., 2011; Zhang et al., 2015).

20   However, there is limited knowledge on the integrated impact of human activities on hydrological drought processes over a large river basin due to the lack of human water use data, which is one of the major issues that hinders the understanding of hydrological drought in the anthropocene (Van Loon et al., 2016b). An alternative approach is to use a land surface hydrological model or a less complicated water balance model to recover the naturalized streamflow by assimilating the reported water use data. The naturalized streamflow data can be used to calibrate hydrological model without the human

25   component to investigate the natural response of hydrological processes to the climate variations (Yuan et al., 2016), it also can also be used to investigate the integral anthropogenic impact on the hydrological drought by comparing with the observed streamflow time series.

The ultimate goal of the understanding of drought processes is to facilitate the development of drought early warning systems for drought adaptation and mitigation, and the development of ocean-atmosphere-land coupled general circulation

30   models (CGCMs) provides an unprecedented opportunity to transfer the advances in seasonal forecasting research (Kirtman et al., 2014; Yuan et al., 2015a) into an integrated drought service. Besides the meteorological drought forecasts (Dutra et al., 2012; Yuan and Wood, 2013; Ma et al., 2015), the agricultural drought forecasts with dynamical seasonal climate forecast models have also been widely applied and evaluated (Luo and Wood, 2007; Mo et al., 2012; Sheffield et al., 2014; Yuan et al., 2015b; Thober et al., 2015). However, dynamical forecasting of hydrological drought based on the CGCM-hydrology

coupled approach (Yuan et al., 2015a) has received less attentions (Trambaueret al., 2015; Sikder et al., 2016), although there are many statistical forecasting studies for low flows. The reasons are threefold: 1) a skilful seasonal forecasting of streamflow usually occurs over basins with large storages of snow, surface and/or subsurface water (Wood and Lettenmaier, 2008; Koster et al., 2010), and the strong control from initial hydrological conditions limits the added value from climate

5 predictions (Wood et al., 2016; Yuan, 2016); 2) unlike the meteorological drought forecasts, both agricultural and hydrological drought forecasts are influenced by the uncertainty in hydrological model, and the hydrological drought forecasting tends to be more challenging since the errors from upstream areas can be transferred to or even amplified in downstream areas; and 3) many river basins are altered by human activities, where the management impacts are often neglected in most dynamical forecasting system. In fact, seasonal forecasting of hydrological drought in the anthropocene

10 raises the questions of how to define the predictability of the anthropogenic processes within a coupled hydro-climate system, how to distinguish the uncertainty from each component, and how to assess the forecast skill of hydrological drought with natural and anthropogenic forcings.

This paper focuses on the understanding and seasonal forecasting of hydrological drought over a heavily managed river basin in North China, the Yellow River basin. Both the naturalized and observed streamflow along the mainstream of Yellow

15 River will be used to investigate the relationship between meteorological and hydrological droughts under natural and anthropogenic conditions, to quantify the influence of human activities on the characteristic of hydrological drought (e.g., drought frequency, duration and severity, and the seasonality of hydrological drought onset), and to assess hydrological drought forecasting in the anthropocene with an experimental seasonal hydrological forecasting system established over the Yellow River basin (Yuan et al., 2016).

20 2. **Data and Method**

2.1 Study domain and hydroclimate observation data

Figure 1 shows the locations of 12 mainstream hydrological gauges used in this study, with Tangnaihai gauge in the headwater region and Lijin gauge at the outlet of the entire Yellow River basin that has a drainage area of 7.52×10^5 km^2. Details of the drainage areas for the 12 gauges can be found in Yuan et al. (2016). Both the natural and observed streamflow

25 datasets at monthly time scale during 1961-2010 were obtained from the Bulletin of Yellow River Water Resources (http://www.yellowriver.gov.cn/), where the naturalized streamflow was calculated by adding the water consumed by agricultural, industrial and civil sectors, and the water regulated by reservoirs, back to the observed streamflow (Yuan et al., 2016). Except for the Tangnaihai gauge in the headwater region, streamflow at the hydrological gauges used in this study were influenced by human interventions. Therefore, the Yellow River is an ideally large river basin to investigate the

30 hydrological drought processes and predictability in the anthropocene.

Precipitation dataset at 0.25-degree resolution during 1961-2010 was interpolated from 324 meteorological stations within the Yellow River basin (Yuan et al., 2016). Figure 1 shows that regional mean annual precipitation decreases from southeast to the northwest. Most precipitation over Yellow River occurs in summer season due to the influence of East Asian monsoon,

resulting in a strong seasonality of precipitation. For each hydrological gauge, the sub-basin mean precipitation was calculated to investigate the relationship between meteorological and hydrological droughts.

2.2 Definitions of drought indices and hydrological drought event

The Standardized Precipitation Index (SPI; McKee et al., 1993) was used as the meteorological drought index. The main advantage of SPI is its multiscale nature, i.e., it can be used to represent meteorological drought at different time scales. In this study, sub-basin mean precipitation datasets averaged over antecedent 1 to 24 months were used to calculate SPI-1 to SPI-24. To account for the seasonality, the SPI for each target month and for each time scale was calculated separately by using 50-year data. For example, SPI-6 at October 1982 was calculated by firstly fitting an empirical distribution based on the precipitation averaged between May and October during 1961-2010, and the May-October mean precipitation in 1982 was then used to determine the SPI6 value for October 1982.

Similarly, monthly naturalized streamflow datasets at 12 hydrological gauges during 1961-2010 were also standardized by using the same procedure, resulting in hydrological drought index named as Standardized Streamflow Index (SSI). For the anthropogenic streamflow, the same parameters of probabilistic distributions fitted from the naturalized streamflow were used, and the SSI values were then calculated.

A threshold of -0.8 was used to represent a drought condition for both SPI and SSI. And a hydrological drought event was selected when the SSI was below -0.8 for at least 3 continuous months (Yuan and Wood, 2013), where the drought onset month was the first month that the SSI fell below -0.8. Once a hydrological drought event occurred, both the duration months and the severity ($\sum_{i=1}^{n}(-0.8 - SSI_i)$), where n is the number of month for the drought event) were calculated. And the number of drought events, mean drought duration and severity were obtained for both naturalized and observed SSI.

2.3 Seasonal hydrological ensemble hindcast datasets

A number of seasonal hydrological ensemble hindcast datasets created by Yuan (2016) were used in this study. To have this paper self-contained, the hindcast experiments were briefly described below. Firstly, a continuous offline hydrological simulation with the calibrated Variable Infiltration Capacity (VIC) land surface hydrological model (Liang et al. 1996) and river routing model driven by observed meteorological forcings from 1951 to 2010, was conducted to generate the initial hydrological conditions (ICs) for the hydrological hindcasts (Yuan et al., 2016). Secondly, a set of 6-month Ensemble Streamflow Prediction (ESP) experiments were carried out by using the VIC and routing models, where the hydrological models were initialized with the generated ICs and were forced by 28 ensemble forcing during 1982-2010 excluding the target year. It is named as ESP/VIC hereafter. Thirdly, a grand ensemble of 99 realizations from eight North American Multimodel Ensemble (NMME; Kirtman et al., 2014) models was used to force hydrological models to generate the NMME/VIC hindcast dataset (Yuan, 2016). The forecast streamflow can be directly compared with offline simulated streamflow. To compare with observed streamflow, a hydrological post-processing procedure (Yuan, 2016) is applied to adjust the forecast streamflow statistically by using the Bayesian theory. While only the ensemble mean (deterministic) forecast skill for streamflow was evaluated in Yuan (2016), both deterministic and probabilistic forecast skill of streamflow (especially for the low flows) are assessed in this paper.

3. Results

3.1 Relation between meteorological drought and hydrological drought

Figure 2 shows the Pearson correlation coefficients between SPI at different time scales and monthly SSI both for naturalized and observed streamflow. There is an increase in the correlation as the SPI time scale increases, suggesting that

5 streamflow is not only influenced by concurrent precipitation, but also by antecedent precipitation up to a few months. Similar to Vicente-Serrano and López-Moreno (2005), the time scale with the maximum correlation is considered as the optimal response time of streamflow to sub-basin averaged precipitation. For the naturalized streamflow, the response times are about 6-12 months over the upper and middle reaches of Yellow River, and about 4 months over the lower reaches. In general, the response time usually increases from upper to lower reaches as the basin size increases. But for the Yellow River

10 basin, the lower reaches are wetter than the upper reaches in terms of annual mean precipitation (Fig. 1), which results in a faster response of streamflow to precipitation over lower reaches than that over upper reaches.

The correlations for the observed streamflow (red curves in Fig. 2) are lower and the response times are longer, suggesting that human interventions basically make the relationship between precipitation and streamflow more nonlinearly. For example, there are small cycles for the values of correlations at Xunhua gauge (Fig. 2b) over the upper reach of Yellow

15 River. The optimal response time even never reaches at Lijin gauge (Fig. 2n), which is the outlet of Yellow River. In other words, the relation between hydrological drought and meteorological drought is less linear in the anthropocene. The response time of hydrological drought to meteorological drought is not solely determined by the climate conditions or the catchment characteristics, while it is heavily determined by human water resources management.

In order to analyze the relation during different seasons, five gauges from upper to lower reaches are selected and the

20 correlations between SPI at different time scales and monthly SSI for different target months are plotted in Figure 3. It is found that the response times are shorter in wet and warm seasons and longer in dry and cold seasons. The response times are about 8 months in the summer, while are about 16 months in the winter. Again, the correlations for observed streamflow for different target months are lower than that for naturalized streamflow, except for the headwater region (i.e., Tangnaihai gauge) without significant human interventions (Fig. 3). The differences are larger during summer seasons than the winter

25 seasons, indicating the seasonality of human water use.

3.2 Effect of human interventions on the hydrological drought characteristics

To demonstrate the effect of human water management on the streamflow variations directly, both the time series of naturalized and observed SSIs are plotted for the five selected gauges in Figure 4. It is found that human water management sometimes has positive influence on the streamflow over the upper reaches. For example, the observed SSI can be larger

30 than the naturalized SSI at Lanzhou gauge (Fig. 4b). However, those increases mostly occur in winter seasons, while they do decrease in summer season when the water demand and water consumption are large. For the middle and lower reaches, observed SSIs are basically lower than the naturalized SSIs (Figs. 4d-e), indicating that human interventions exacerbate the hydrological drought conditions in the lower reaches of the Yellow River basin.

By following the definition of hydrological drought event (it should last for at least 3 continuous months) in Section 2.2, the frequency, duration and severity of hydrological droughts under natural and anthropogenic conditions are calculated and shown in Figure 5. Under natural condition, seasonal hydrological drought occurs 8-16 times during the 50-year (1961-2010) period, where the frequency of hydrological drought is not necessarily higher over lower reaches than that over upper

5   reaches (Fig. 5a). This is partly because there is more precipitation over lower reaches, and the naturalized hydrological drought basically represents the response to the meteorological drought. In contrast, the observed hydrological drought frequency shows quite different characteristics, where the human interventions increase the drought frequency only by 7% to 80% from the upper gauges down to the Xiaheyan gauge (the 5th gauge in Fig. 5a), but they increase the drought frequency by 200% or even 400% over the lower reaches (e.g., Lijin gauge at the outlet).

10   For the drought duration in natural conditions, it is generally longer over upper reaches which is again due to a drier climate (Fig. 5b). With human interventions, there is a decrease in drought duration for the upper gauges down to Xiaheyan gauge. This suggests that seasonality of human interventions reduces the persistency of drought over the upper reaches of the Yellow River basin. From Shizuishan gauge (the 6th gauge in Fig. 5b) down to the outlet, the duration of hydrological drought increases by 12%-83% under human interventions. There is no significant change in drought severity down to

15   Xiaheyan gauge, but the severity increases by up to 6 times over the lower reaches (Fig. 5c). Therefore, human interventions not only increase the drought frequency over most areas of the Yellow River basin, but also increase the drought severity dramatically over the lower reaches.

To investigate the preference of the occurrence of seasonal hydrological drought, the onset seasons are identified based on individual hydrological drought events. And the ratios of the number of drought onsets during different seasons to the total

20   number of drought events are plotted in Figure 6. Without human interventions, most seasonal hydrological droughts start in summer, especially for the upper and lower gauges. While for the gauges over middle reaches, autumn is also a preferred season for drought onset. However, Fig. 6 shows that human intervention changes the seasonality of the hydrological drought onset significantly, where the spring is the preferred season for the drought onset for the lower reaches, and the ratio for summer season is also increased for the upper reaches. With human interventions, the hydrological drought onset

25   becomes earlier.

3.3 Seasonal forecasting of hydrological drought with human interventions

To demonstrate the capability of predicting hydrological drought in the anthropocene, a drought case of 2001 is selected to verify the ensemble forecasting of SSI. In terms of meteorological condition, 2001 is a moderate dry year, with the precipitation less than the climatology by 9.4%. However, 2001 is a severe hydrological drought year, with observed

30   streamflow less than the climatology by 26%-37% over the upper reaches and by 51%-86% over the lower reaches (http://www.yellowriver.gov.cn/). Figure 7 shows the ensemble forecasts started from February and June of 2001 for the selected five gauges from upper to lower reaches. As verified by the offline simulated SSI, the climatological forecast method (ESP/VIC) has some skill in the February forecast, but totally misses the drought in the June forecast with the ensemble mean SSIs (blue lines in the left panels of Fig. 7) close to zeros that are larger than the drought threshold lines

(black horizontal lines) of -0.8. By using the climate-model-based approach (NMME/VIC, see Yuan (2016) for details), there are not significant improvements for the forecasts started from February, but a number of ensembles can capture the hydrological drought conditions, with the ensemble mean SSIs (blue lines in the right panels of Fig. 7) close to the drought threshold lines. This suggests the added values from climate forecast models in the hydrological drought forecasting.

5   The results shown in Figure 7 neglect the errors or uncertainties in hydrological model because the offline simulated hydrological drought index is used as the reference to verify the forecasts. To compare with observed SSI, the forecast streamflow series are post-processed (see Section 2.3 for details), and the results for 2001 are shown in Figure 8. Both ESP/VIC and NMME/VIC can predict a drought condition with ensemble mean SSIs much lower than the -0.8 for the middle and lower gauges (Hekouzhen, Huayuankou and Lijin), but both underestimate the drought severity. NMME/VIC

10 shows non-trivial improvement against ESP/VIC in terms of the drought forecasting. An interesting difference is that five gauges have similar ensemble spreads in the natural conditions (Fig. 7), but the spreads vary among upstream and downstream gauges in the anthropogenic conditions (Fig. 8). This is because the hydrological post-processing is applied for each target months and for each gauge independently, the ensemble spreads do not necessarily increase over forecast leads after the post-processing, while they indeed depend on the magnitude or intensity of human interventions.

15 To assess the probabilistic forecast skill for hydrological droughts during the hindcast period of 1982-2010, the Brier Score (BS; Wilks, 2011) is used. It is defined as:

$$BS = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2,$$

where k denotes a number of n forecast-reference pairs of hydrological drought conditions; o_k is probability from the reference (offline simulated SSI for the natural condition, and observed SSI for the anthropogenic case) for the kth pair, with

20 o_k=1 if the drought occurs (SSI<-0.8) and o_k=0 if it does not (SSI>-0.8); y_k is the corresponding probability of drought occurrence from the forecast, for example, if 6 of 10 ensemble forecast members have SSI<-0.8, then y_k=0.6. BS is similar to the root mean squared error, so a smaller BS represents a better performance.

Figure 9 shows the BS for the ensemble hydrological drought forecasts at different hydrological gauges and at different lead times. The statistics are based a set of comprehensive hindcasts, where for each month during the 29-year (1982-2010)

25 period there is a 6-month hindcast. This consists of 348 hindcast cases, each one is 6-month long, and has 28 or 99 realizations for ESP/VIC and NMME/VIC respectively. For the natural condition (upper panels of Fig. 9), the BS values increase as the forecast leads increase, indicating that the performance generally decreases over leads. The performance of probabilistic hydrological drought forecasting is better over lower reaches than that over upper reaches both for ESP/VIC and NMME/VIC, suggesting the influence of catchment memory. As compared with ESP/VIC, NMME/VIC has a better

30 performance, with BS decreased by 11%-26% in the first month, by 3%-14% in the second and third months.

As verified against observed SSI (lower panels of Fig. 9), the performances for ESP/VIC and NMME/VIC are surprisingly better than the natural cases, and the performances do not necessarily degrade over forecast lead times. This is because human interventions increase the occurrence and severity of hydrological drought, and outweigh the climate variations in

many cases. The hydrological post-processing imparts the first-order control in the forecasting, and many post-processed forecasts can represent drought conditions (SSI<-0.8), although they may underestimate the severity as shown in Fig. 8 (right panels). The differences in BS between ESP/VIC and NMME/VIC for the anthropogenic case are smaller than that for the natural case. In other words, seasonal predictability of hydrological drought in the anthropocene greatly depends on the
5   information of human water use, or the predictability of human interventions.

4. Concluding Remarks

This paper investigates the effects of human interventions on hydrological drought processes and the drought forecasting. Naturalized and observed monthly streamflow are standardized to calculate the hydrological drought index, the Standardized Streamflow Index (SSI). Comparison between naturalized and observed SSI at 12 hydrological gauges along the mainstream
10   of the Yellow River basin (the second largest river basin in China with a drainage area of 7.52×10^5 km^2) shows that human interventions basically increase the nonlinearity in the relationship between meteorological and hydrological droughts, and increase the response time of the latter to the former especially during summer seasons. Seasonal hydrological drought events are identified with monthly SSI<-0.8 for at least three continuous months. Due to heavy human water consumptions over the middle and lower reaches of the Yellow River, there are two to four-fold increases in the drought frequency and up
15   to six-fold increases in the drought severity, the drought duration increases by 12-83%, and the hydrological drought onset becomes earlier.

The naturalized streamflow datasets are used to calibrate the VIC land surface hydrological model and the routing model, and both the climatological forcings (ESP) and the climate model predicted forcings (NMME) are used to drive the hydrological models to provide seasonal streamflow forecasts. For a severe hydrological drought event occurred over the
20   Yellow River in 2001, ESP/VIC does not capture it while NMME/VIC has some skill when they are verified against naturalized SSI. The added values from climate-model-based seasonal hydrological drought forecasting are decreased in the anthropocene, where both methods can predict a drought condition after the hydrological post-processing but underestimate the severity. Unlike the naturalized hydrological drought forecasting, the ensemble spreads do not necessarily increase over forecast leads in the anthropocene because of the seasonality of human interventions that have been implicitly incorporated
25   in the hydrological post-processing. Based on the assessment of all hindcasts during 1982-2010, it is found that NMME/VIC decreases (improves) the Brier Score (BS) against ESP/VIC by 11%-26% in the first month and by 3%-14% in the second and third months for the probabilistic hydrological drought forecasting in the naturalized conditions. In the anthropocene, the performances for both forecast methods seem better in terms of BS, and the forecast skill does not necessarily decreases over forecast leads due to dominant influence of human water consumption on the hydrological drought processes.
30   While the effects of human interventions on hydrological drought processes have been studied in the past, to our knowledge, this study is among the first to investigate seasonal hydrological drought forecasting in the anthropocene. Intensive and direct influence of human water use challenges our understanding of hydrological drought predictability. Traditionally, hydro-climate predictability usually refers to the struggle between deterministic and chaotic physical processes. In the anthropocene, the human influence sometimes outweighs those natural hydro-climate variations and variability, and can be a

major source of predictability for hydrological drought. Current hydrological post-processing procedure accounts for the seasonality of human water use by adding/deducting water to/from the predicted streamflow, and adjusting the forecast results based on the Bayesian theory. But it is difficult to account for the inter-annual variability of human water use. During severe hydrological drought events, the human water use tends to be more intensive to adapt to the drought conditions. How

5   to quantify and model the variability of human water use is an interdisciplinary question for both physical and social sciences. In addition, the interdisciplinary collaboration is also indispensable to objectively quantify and to accurately predict drought impact, as the drought impact and drought are quite different.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (No. 91547103), China

10   Special Fund for Meteorological Research in the Public Interest (Major projects) (GYHY201506001), CAS Key Research Program of Frontier Sciences (QYZDY-SSW-DQC012), National Key Research and Development Program (2016YFA0600403), and the Thousand Talents Program for Distinguished Young Scholars. This paper is submitted to the special issue "Observations and modeling of land surface water and energy exchanges across scales" in honor of Professor Eric F. Wood. The first author would like to thank Eric for his inspiring guidance.

References

AghaKouchak, A., Feldman, D., Hoerling, M., Huxman, T., and Lund, J.: Water and climate: Recognize anthropogenic drought, Nature, 524, 409–411, doi:10.1038/524409a, 2015.

Barker, L. J., Hannaford, J., Chiverton, A., and Svensson, C.:From meteorological to hydrological drought using
5   standardised indicators, Hydrol. Earth Syst. Sci., 20, 2483–2505, 2016.

Bloomfield, J. P., Marchant, B. P., Bricker, S. H., and Morgan, R. B.: Regional analysis of groundwater droughts using hydrograph classification, Hydrol. Earth Syst. Sci., 19, 4327–4344, 2015.

Dutra, E., Magnusson, L., Wetterhall, F., Cloke, H. L., Balsamo, G., Boussetta, S., and Pappenberger, F.: The 2010–2011 drought in the Horn of African ECMWF reanalysis and seasonal forecast products, Int. J. Climatol., 33,1720–1729,
10   doi:10.1002/joc.3545, 2012.

Folland, C. K., et al.: Multi-annual droughts in the English Lowlands: a review of their characteristics and climate drivers in the winter half-year, Hydrol. Earth Syst. Sci., 19, 2353–2375, 2015.

Haslinger, K., Koffler, D., Schoner, W., and Laaha, G.: Exploring the link between meteorological drought and streamflow: Effects of climate catchment interaction, Water Resour. Res., 50, 2468–2487, doi:10.1002/2013WR015051, 2014.

15 Hoerling, M., and Kumar, A.: The perfect ocean for drought, Science, 299, 691–694, doi:10.1126/science.1079053, 2003.

Hong, S. Y., and Kalnay, E.: Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma–Texas drought, Nature, 408, 842–844, doi:10.1038/35048548, 2000.

Kirtman, B. P., Min, D., Infanti, J. M., et al.: The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction, B. Am. Meteorol. Soc., 95, 585–601,
20   doi:10.1175/BAMS-D-12-00050.1, 2014.

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, doi:10.1038/ngeo944, 2010.

Kumar, R., et al.: Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator, Hydrol. Earth Syst. Sci., 20, 1117–1131, 2016.

25 Liang, X., Wood, E. F., and Lettenmaier, D. P.: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications, Global Planet. Change, **13**, 195–206, doi: 10.1016/0921-8181(95)00046-1, 1996.

López-Moreno, J. I., Vicente-Serrano, S. M., Beguería, S., García-Ruiz, J. M., Portela, M. M., and Almeida, A. B.: Dam effects on droughts magnitude and duration in a transboundary basin: The Lower River Tagus, Spain and Portugal, Water Resour. Res., 45, W02405, doi:10.1029/2008WR007198, 2009.

30 Luo, L. and Wood, E. F.: Monitoring and predicting the 2007 U.S. drought, Geophys. Res. Lett., 34, L22702,doi:10.1029/2007GL031673, 2007.

Ma, F., Yuan, X., and Ye, A.: Seasonal drought predictability and forecast skill over China, J. Geophys. Res.-Atmos., 120, 8264–8275, doi:10.1002/2015JD023185, 2015.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales. Preprints, Eighth Conf. on Applied Climatology, Anaheim, CA, Amer. Meteor. Soc., 179–184, 1993.

Mo, K., Shukla, S., Lettenmaier, D. P., and Chen, L.: Do Climate Forecast System (CFSv2) forecasts improve seasonal soil moisture prediction?, Geophys. Res. Lett., 39, L23703,doi:10.1029/2012GL053598, 2012.

Niu, J., Chen, J., and Sun, L.: Exploration of drought evolution using numerical simulations over the Xijiang (West River) basin in South China, J. Hydrol., 526, 68-77, 2015.

Pozzi, W., et al.: Towards global drought early warning capability: Expanding international cooperation for the development of a framework for global drought monitoring and forecasting, Bull. Amer. Meteor. Soc., 94, 776–785,doi:10.1175/BAMS-D-11-00176.1, 2013.

Rimkus, E., et al.: Dynamics of meteorological and hydrological droughts in the Neman river basin, Environ. Res. Lett., 8, 045014, 2013.

Sheffield, J., Wood, E. F., and Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., and Ogallo, L.: A Drought Monitoring and Forecasting System for Sub-Sahara African Water Resources and Food Security, B. Am. Meteorol. Soc., 95, 861–882, doi:10.1175/BAMS-D-12-00124.1, 2014.

Sikder, S., Chen, X., Hossain, F., Roberts, J. B., Robertson, F., Shum, C. K., and Turk, F. J.: Are general circulation models ready for operational streamflow forecasting for water management in the Ganges and Brahmaputra river basins?, J. Hydrometeorol., 17, 195–210, doi:10.1175/JHM-D-14-0099.1, 2016.

Staudinger, M., Weiler, M., and Seibert, J.: Quantifying sensitivity to droughts-an experimental modeling approach, Hydrol. Earth Syst. Sci., 19, 1371-1384, 2015.

Stoelzle, M., Stahl, K., Morhard, A., and Weiler, M.: Streamflow sensitivity to drought scenarios in catchments with different geology, Geophys. Res. Lett., 41, 6174–6183, doi:10.1002/2014GL061344, 2014.

Teuling, A. J., et al.: Evapotranspiration amplifies European summer drought, Geophys. Res. Lett. 40, 2071–2075, doi: 10.1002/grl.50495, 2013.

Thober, S., Kumar, R., Sheffield, J., Mai, J., Schafer, D., and Samaniego, L.: Seasonal soil moisture drought prediction over Europe using the North American Multi-Model Ensemble (NMME), J. Hydrometeorol., 16, 2329–2344,doi:10.1175/JHM-D-15-0053.1, 2015.

Trambauer, P., Werner, M., Winsemius, H. C., Maskey, S., Dutra,E., and Uhlenbrook, S.: Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa, Hydrol. Earth Syst. Sci., 19, 1695–1711, doi:10.5194/hess-19-1695-2015, 2015.

Van Loon, A. F., Van Huijgevoort, M. H. J., and Van Lanen, H. A. J.: Evaluation of drought propagation in an ensemble mean of large-scale hydrological models, Hydrol. Earth Syst. Sci., 16, 4057-4078, doi:10.5194/hess-16-4057-2012, 2012.

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl,K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen,L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M.,Verbeiren, B.,Wagener, T., Rangecroft,

S.,Wanders, N., and VanLanen, H. A. J.: Drought in the Anthropocene, Nature Geosci., 9, 89–91, doi:10.1038/ngeo2646, 2016a.

Van Loon, A. F., et al.: Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches, Hydrol. Earth Syst. Sci., 20, 3631–3650, 2016b.

5   Vicente-Serrano, S. M. and López-Moreno, J. I.: Hydrological response to different time scales of climatological drought: an evaluation of the Standardized Precipitation Index in a mountainous Mediterranean basin, Hydrol. Earth Syst. Sci., 9, 523–533, doi:10.5194/hess-9-523-2005, 2005.

Wen, L., Rogers, K., Ling, J., and Saintilan, N.: The impacts of river regulation and water diversion on the hydrological drought characteristics in the Lower Murrumbidgee River, Australia, J. Hydrol., 405, 3-4, 382–391,2011.

10 Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Int. Geophys. Ser., vol. 100, 3rd ed., 676 pp., Academic, San Diego, Calif., 2011.

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

Wood, A. W., Hopson, T., and Newman, A., Brekke, L., Arnold, J., and Clark, M.: Quantifying streamflow forecast skill
15   elasticity to initial condition and climate prediction skill, J. Hydrometeorol.,17, 651–668, doi:10.1175/JHM-D-14-0213.1, 2016.

Wood, E. F., Schubert, S. D., and Wood, A. W., et al.: Prospects for advancing drought understanding, monitoring, and prediction, J. Hydrometeorol., 16, 1636-1657, doi: 10.1175/JHM-D-14-0164.1, 2015.

Yuan, X., and Wood, E. F.: Multimodel seasonal forecasting of global drought onset, Geophys. Res. Lett., 40, 4900–4905,
20   doi:10.1002/grl.50949, 2013.

Yuan, X., Wood, E. F., Chaney, N. W., Sheffield, J., Kam, J., Liang, M., and Guan, K.: Probabilistic seasonal forecasting of African drought by dynamical models, J. Hydrometeor.,14, 1706–1720, doi:10.1175/JHM-D-13-054.1, 2013.

Yuan, X., Wood, E. F., and Ma, Z.: A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development, WIREs Water, 2, 523–536,doi:10.1002/wat2.1088, 2015a.

25 Yuan, X., Roundy, J. K., Wood, E. F., and Sheffield, J.: Seasonal forecasting of global hydrologic extremes: system development and evaluation over GEWEX basins, B. Am. Meteorol. Soc., 96, 1895–1912, doi:10.1175/BAMS-D-14-00003.1, 2015b.

Yuan, X., Ma, F., Wang, L., Zheng, Z., Ma, Z., Ye, A., and Peng, S.: An experimental seasonal hydrological forecasting system over the Yellow River basin – Part 1: Understanding the role of initial hydrological conditions, Hydrol. Earth
30   Syst. Sci., 20, 2437–2451, doi:10.5194/hess-20-2437-2016, 2016.

Yuan, X.: An experimental seasonal hydrological forecasting system over the Yellow River basin – Part 2: The added value from climate forecast models, Hydrol. Earth Syst. Sci., 20, 2453–2466, doi:10.5194/hess-20-2453-2016, 2016.

Zhang, R., Chen, X., Zhang, Z., and Shi, P.: Evolution of hydrological drought under the regulation of two reservoirs in the headwater basin of the Huaihe River, China, Stoch. Env. Res. Risk A., 29(2), 487–499, 2014.

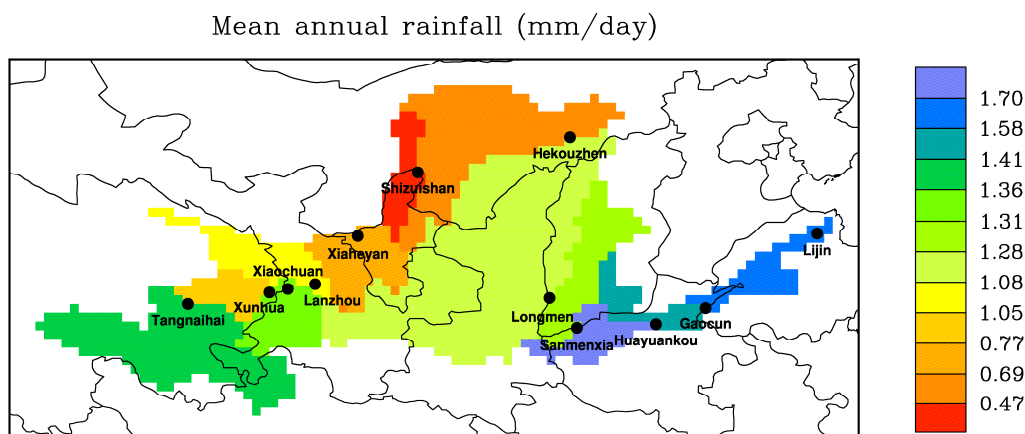



Figure 1. Locations of hydrological stations over the Yellow River basin. Shaded areas are regional mean annual rainfall (mm/day) averaged during 1961-2010.

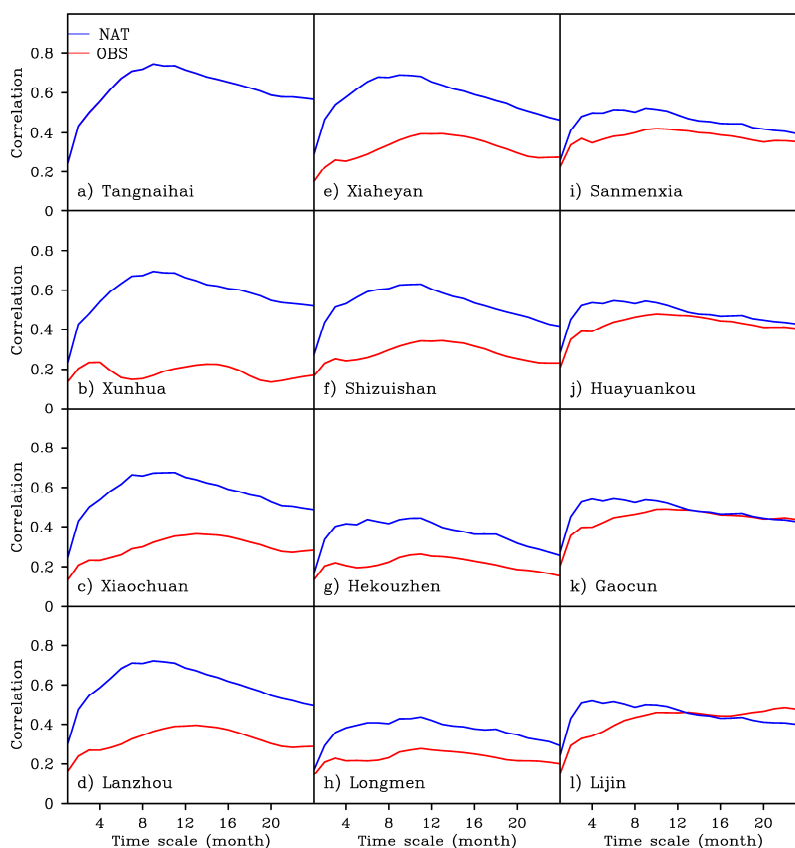



Figure 2. Correlations between Standardized Precipitation Index (SPI) at different time scales and monthly naturalized (blue) or observed (red) Standardized Streamflow Index (SSI) for 12 hydrological gauges located from upper to lower mainstream of the Yellow River.

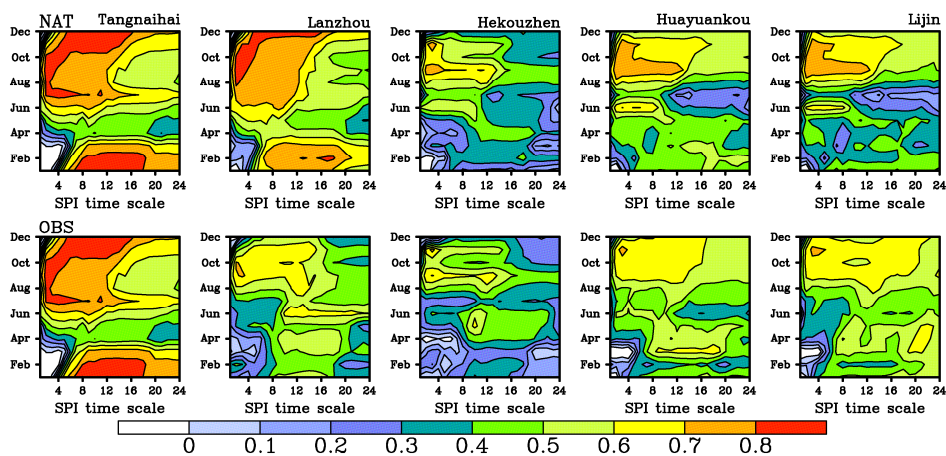
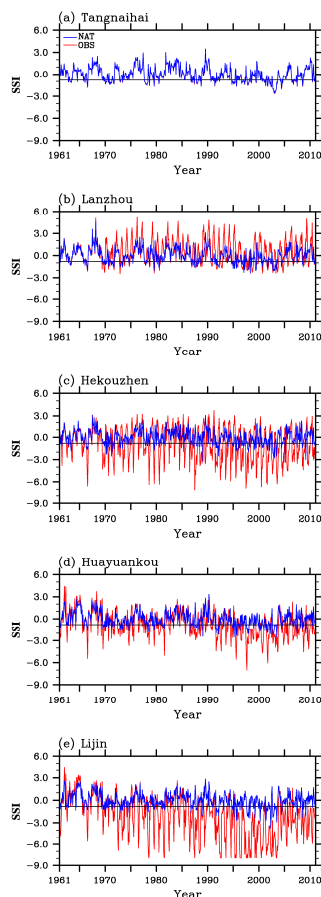



Figure 3. The same as Fig. 2, but for each target month for five selected hydrological gauges.

5 **Figure 4.** Time series of naturalized (blue) and observed (red) Standardized Streamflow Index (SSI) for five selected hydrological gauges. The horizontal black lines represent the threshold of -0.8 for drought conditions.

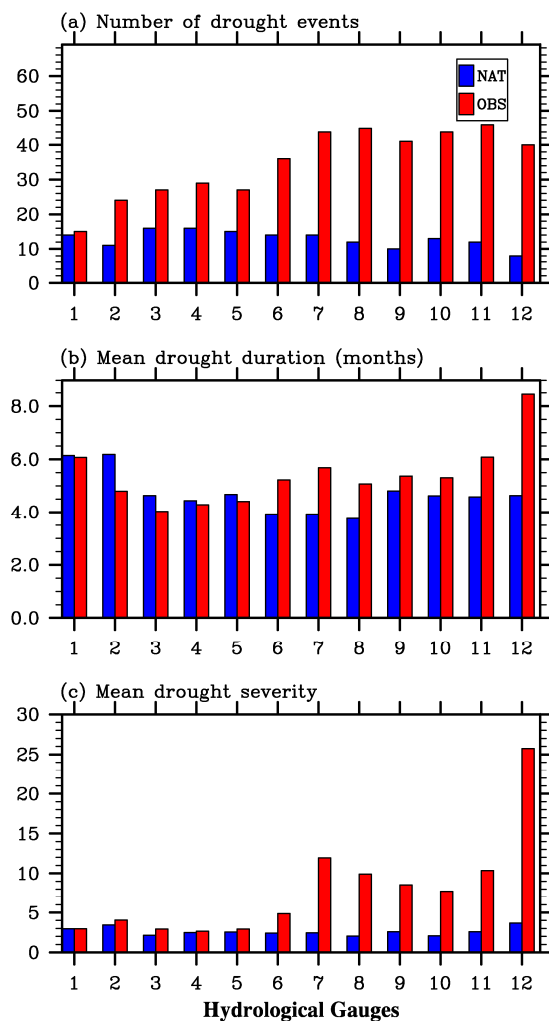



Figure 5. Characteristics of hydrological drought events based on the streamflow time series with (red) or without (blue) human influences during 1961-2010 for 12 hydrological gauges. A hydrological drought event is selected when the Standardized Streamflow Index (SSI) is continuously below -0.8 for at least 3 months.

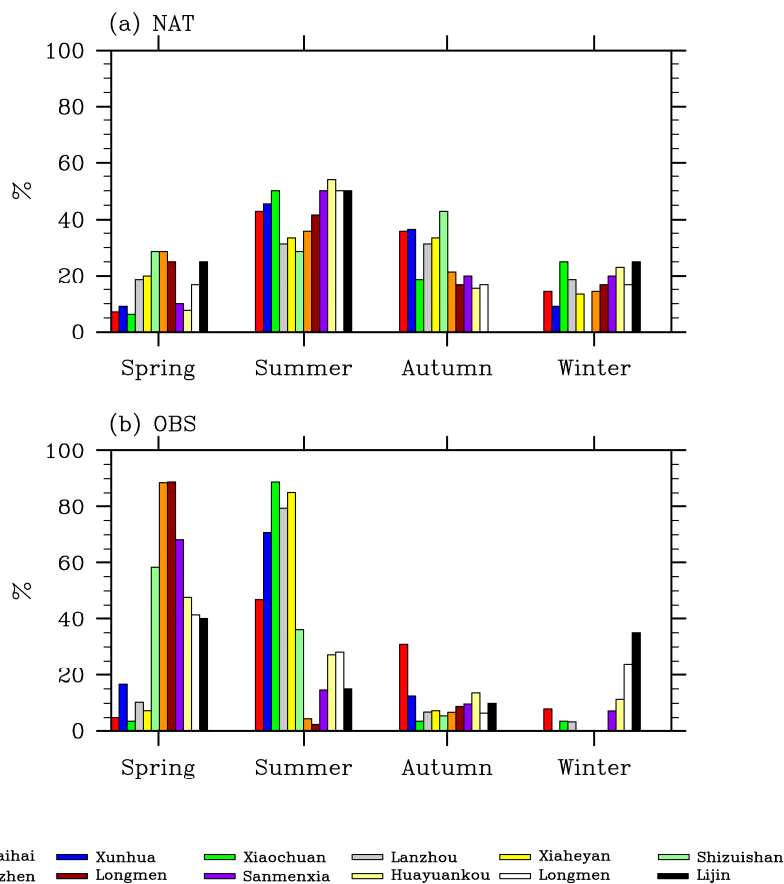



Figure 6. Ratio of the number of hydrological drought onsets occurring in different seasons to the total number of hydrological drought events for the 12 hydrological gauges. Drought events are classified the same as those in Fig. 5.

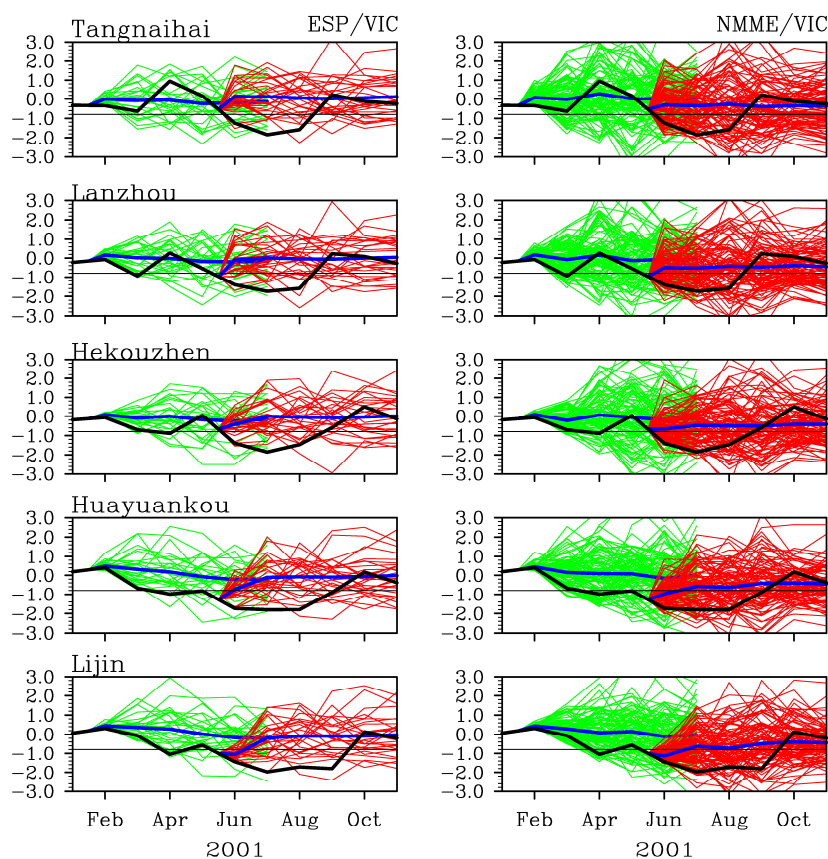



Figure 7. Seasonal ensemble hindcast of the 2001 Yellow River hydrological drought from upstream to downstream gauges by using a climatology method (ESP/VIC) and the climate-model-based approach (NMME/VIC). Vertical axes are Standardized Streamflow Index (SSI), where SSI<-0.8 represents a hydrological drought condition. Solid black lines represent the offline simulated SSI by the hydrological model VIC, green and red lines are for individual ensemble members from the hindcasts started from the beginnings of February and June respectively, and blue lines are the ensemble means of the hincasts.

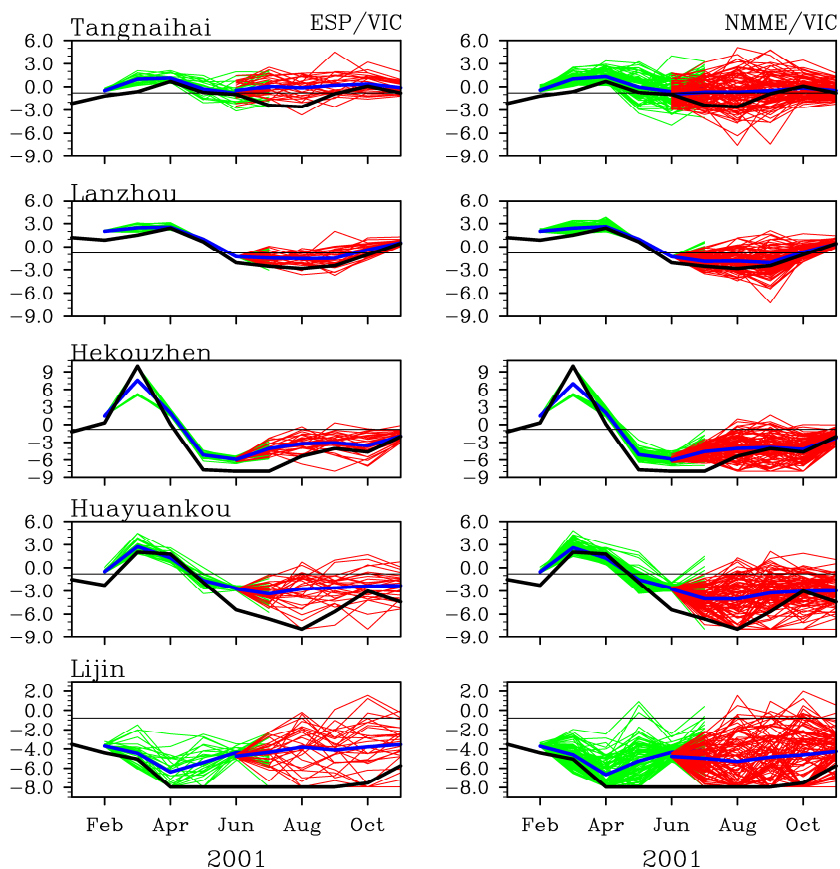



Figure 8. The same as Fig. 7, but for the post-processed Standardized Streamflow Index (SSI) hindcasts from ESP/VIC and NMME/VIC (see Section 2.3 for details) as verified by observed SSI.

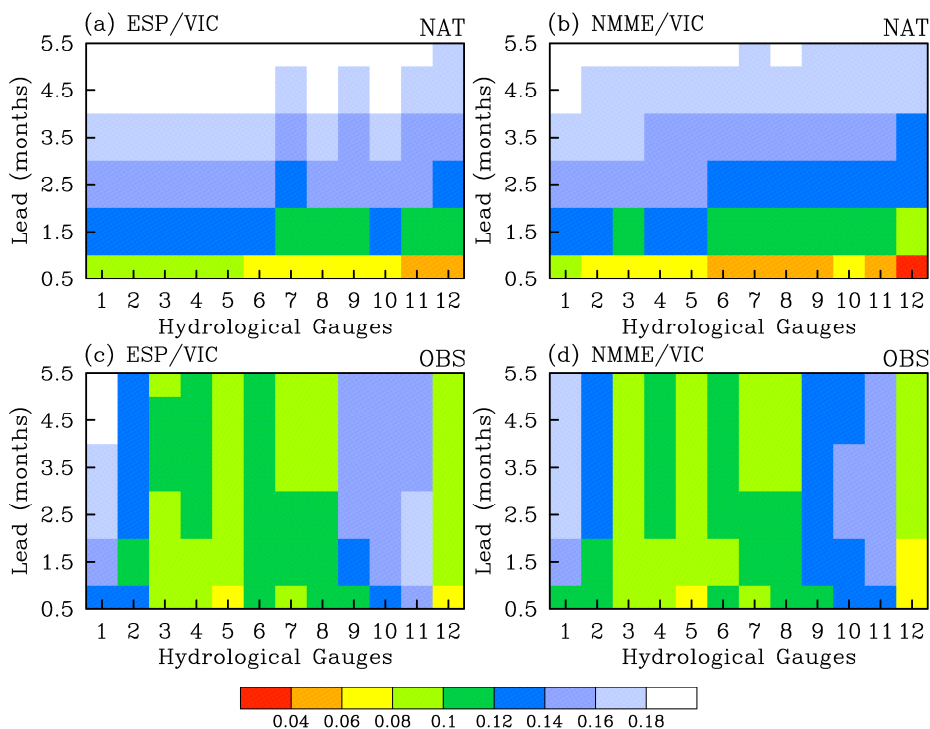



Figure 9. Brier score (BS) for ensemble hydrological drought forecasts at different lead times from a climatology method (ESP/VIC) and the climate-model-based approach (NMME/VIC) for different hydrological gauges as verified against VIC offline simulated (a, b) and observed (c, d) streamflow over the Yellow River basin during the period of 1982-2010. For the verification against observed streamflow, the forecasts have been post-processed.