Understanding and seasonal forecasting of hydrological drought in the anthropocene

Xing Yuan[1], Miao Zhang[1,2], ~~and~~ Linying Wang[1,2], and Tian Zhou[3]

带格式的: 字体: 小四, 上标

[1]CAS Key Laboratory of Regional Climate-Environment for Temperate East Asia (RCE-TEA), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[2]University of Chinese Academy of Sciences, Beijing, 100049, China
[3]Pacific Northwest National Laboratory, Richland, WA 99352, USA

带格式的: 上标

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~~ decrease the ~~nonlinear response of~~correlation between hydrological ~~drought to the~~and meteorological droughts, and ~~increase~~ make the hydrological drought respond to longer time scale of meteorological drought ~~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).

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 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 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), because many reservoirs were built for meeting the irrigation demand more reliably. 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 (Wada et al., 2013; Zhou et al., 2016) 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 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 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 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 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 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).

2. Data and Method

2.1 Study domain and hydroclimate observation data

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, with more than 80% of annual precipitation falls within May-September (Fig. 2). For each hydrological gauge, the sub-basin mean precipitation was calculated to investigate the relationship between meteorological and hydrological droughts.

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 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 (Fig. 3). Therefore, the Yellow River is an ideally large river basin to investigate the hydrological drought processes and predictability in the anthropocene.

In general, human interventions decrease streamflow over upper and middle reaches of Yellow River during rainy season while increase it during dry season (Figs. 3b-3h). This suggests that reservoirs in the upper and middle reaches of Yellow River store rain water in wet season and distribute it in the remaining time of the year according to the need, which is similar to regulations in other parts of the world (Wada et al., 2014). Actually, Figure 4a shows that the annual mean observed streamflow at upper reaches can be higher than the naturalized streamflow during dry years due to the reservoir water release (e.g., years 2000, 2002, 2006 and 2010 for Lanzhou gauge). Over the lower reaches, the observed streamflow is significantly lower than the naturalized streamflow during wet season due to heavy water consumption (Figs. 3i-3l), but the former is close to the latter during dry season because of no significant water consumption or reservoir management (Figs. 4c-4d). Figure 4 also shows that the magnitudes of reservoir storage changes are quite small as compared with streamflow. In fact, the mean absolute changes of reservoir storage during 1998-2010 are about 14%-38% and 12%-14% of observed and naturalized streamflow, respectively. This suggests that other human interventions, such as direct withdrawal of surface water for agricultural, industrial and civil consumptions, account for a large part of streamflow variations over Yellow River.~~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). Note that SSI was similar to the standardized runoff index (SRI) defined by Shukla and Wood (2008), except that streamflow was used here for a standardization. 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 ($\sum_{i=1}^{n}(-0.8-\mathrm{SSI}_i)$, 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). The observed meteorological forcing datasets including daily precipitation, daily maximum and minimum surface air temperature, and surface wind were interpolated from 324 China Meteorological Administration stations. The VIC model version 4.0.5 was used to predict runoff in a water balance mode over the entire Yellow River basin with 1321 grid cells at 0.25-degree resolution, and a routing model was used to translate the runoff into streamflow at each 0.25-degree grid cell, and to route the flow into rivers and finally into the ocean (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). Here, the 1-degree NMME global hindcasts of monthly precipitation and temperature were bilinearly interpolated into 0.25-degree, the interpolated monthly hindcasts for each NMME model were then bias-corrected independently against observations by using the quantile-mapping method (Wood et al., 2002) in a cross-validation mode (i.e., dropping observation and forecast in the target year when generating the climatology), and these bias-corrected monthly hindcasts were finally temporally disaggregated to daily by historical sampling and rescaling (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~~ 5 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 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 time scale of SPI that streamflow responds to~~optimal response time of streamflow to sub-basin averaged precipitation~~. For the naturalized streamflow, ~~the~~ it ~~response~~ responds ~~times are about~~to 6-12 months SPI over the upper and middle reaches of Yellow River, and to about 4 months SPI 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 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 significantly lower than for naturalized streamflow for gauges from Xunhua down to Huayuankou, with p values less than 0.01 (Figs. 5b-5j). There is also a significant difference in correlation for Gaocun gauge with $p<0.05$ (Fig. 5k), but the difference is not statistically significant for Lijin gauge with $p>0.1$ (Fig. 5l). Except for the Tangnaihai gauge in headwater region, ~~and~~ the SPI ~~response times~~time scales with the maximum correlation are longer for the observed streamflow than the naturalized streamflow, suggesting that human interventions basically make the hydrological drought respond to longer time scale of meteorological drought~~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 River. The optimal response time even never reaches at Lijin gauge (Fig. 2l), 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 correlations between SPI at different time scales and monthly SSI for different target months are plotted in Figure ~~3~~6. It is found that ~~the~~ streamflow ~~response~~ responds ~~times are~~to shorter time scale of SPI in wet and warm seasons and longer time scale 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~~6). The differences are larger during summer seasons than the winter seasons, ~~indicating~~ which is consistent with the seasonality of human water use as shown by different streamflow in Figure 3.

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~~7. It is found that human water management sometimes has positive influence on ~~the streamflow~~hydrological drought over the upper reaches. For example, the observed

SSI can be larger than the naturalized SSI at Lanzhou gauge (Fig. ~~4b~~7b). However, those increases mostly occur in winter seasons, while they do decrease in summer season when the water demand and water consumption are ~~large~~high. For the middle and lower reaches, observed SSIs are basically lower than the naturalized SSIs (Figs. ~~4d~~7d-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~~8. 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 reaches (Fig. ~~5a~~8a). 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~~8a), but they increase the drought frequency by 200% or even 400% over the lower reaches (e.g., Lijin gauge at the outlet).

For the drought duration in natural conditions, it is generally longer over upper reaches which is again due to a drier climate (Fig. ~~5b~~8b). With human interventions, there is a slight decrease in drought duration for the upper gauges down to Xiaheyan gauge. This suggests that human interventions reduce the persistency of drought over the upper reaches of the Yellow River basin. From Shizuishan gauge (the 6th gauge in Fig. ~~5b~~8b) 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 Xiaheyan gauge, but the severity increases by up to 6 times over the lower reaches (Fig. ~~5c~~8c). 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 number of drought events are plotted in Figure ~~6~~9. 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~~ 9 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 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 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~~ 10 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~~10) 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~~10) close to the drought threshold lines. This suggests the added values from climate forecast models in the hydrological drought forecasting.

The results shown in Figure ~~7~~ 10 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~~11. 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 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~~10), but the spreads vary among upstream and downstream gauges in the anthropogenic conditions (Fig. ~~8~~11). 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.

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:

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

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 k^{th} pair, with 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~~ 12 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) 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~~12), 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 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~~12), 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~~ 11 (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 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 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 decrease the correlation between hydrological and meteorological droughts, and make the hydrological drought respond to longer time scale of meteorological drought especially during rainy seasons~~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 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 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 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~~ become 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. 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 (or uncertainty if it is not fully understood) 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. Another popular method is to parameterize the human interventions directly in hydrological models (e.g., Wada et al., 2014; Zhou et al., 2016 among many others), where the irrigated area can be estimated for crop water demand, and the diversion and return flows can also be simulated in those models. This modeling framework could be further pushed forward by the availability of Surface Water and Ocean Topography (SWOT) satellite data in the near future, where surface water storage in reservoirs and rivers can be at 250m resolution (Biancamaria et al., 2016). ~~But~~ However, it is difficult to account for the inter-annual variability of human water use in a “real” forecasting mode. During severe hydrological drought events, the human water use tends to be more intensive to adapt to the drought conditions. How 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. We would like to thank Prof. Dennis Lettenmaier and two anonymous reviewers for their helpful comments. This work was supported by the National Natural Science Foundation of China (No. 91547103), China Special Fund for Meteorological Research in the Public Interest (Major projects) (GYHY201506001), National Key Research and Development Program (2016YFA0600403), CAS Key Research Program of Frontier Sciences (QYZDY-SSW-DQC012), 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.

带格式的: 字体: 10 磅, 图案: 清除 (白色)

带格式的: 字体: 10 磅

带格式的: 字体: (中文) Times New Roman

带格式的: 字体: 10 磅, 图案: 清除 (白色)

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 standardised indicators, Hydrol. Earth Syst. Sci., 20, 2483–2505, 2016.

Biancamaria, S., Lettenmaier, D. P., and Pavelsky, T. M.: The SWOT mission and its capabilities for land hydrology, Surv. Geophys., 37, 307-337, 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, 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.

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, 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.

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.

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.

Shukla, S., and Wood, A. W.: Use of a standardized runoff index for characterizing hydrologic drought, Geophys. Res. Lett., 35, L02405, doi:10.1029/2007GL032487, 2008.

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.

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.

Wada, Y., van Beek, L. P. H., Wanders, N., and Bierkens, M. F. P.: Human water consumption intensifies hydrological drought worldwide, Environ. Res. Lett., 8, 034036, 2013.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dynam., 5, 15-40, doi:10.5194/esd-5-15-2014, 2014.

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.

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., Mauer, E. P., Kumar, A., and Lettenmaier, D.P.: Long-range experimental hydrologic forecasting for the eastern United States,J.Geophys. Res., 107, 4429, doi:10.1029/2001JD000659, 2002.

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 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, 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.

带格式的: 正文，左，缩进：左侧： 0 厘米，悬挂缩进：2.25 字符，首行缩进： -2.25 字符，段落间距段前：0 磅，在相同样式的段落间添加空格，行距：单倍行距，非在对开页上不交换缩进

带格式的: 字体：（中文）Times New Roman，10 磅，字体颜色：文字 1

带格式的: 缩进：左侧： 0 厘米，悬挂缩进：2.25 字符

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.

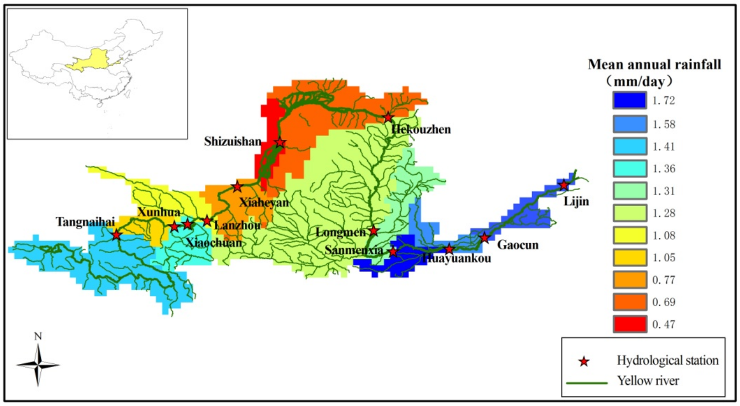
Yuan, X., Roundy, J. K., Wood, E. F., and Sheffield, J.: Seasonal forecasting of global hydrologic extremes: system development and evaluation over GEWEX basins, B. Am. Meteorol. Soc., 96, 1895–1912, doi:10.1175/BAMS-D-14-00003.1, 2015b.

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 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, 2015.

Zhou, T., Nijssen, B., Gao, H., and Lettenmaier, D. P.: The contribution of reservoirs to global land surface water storage variations. *J. Hydrometeor.*, 17, 309-325, 2016.



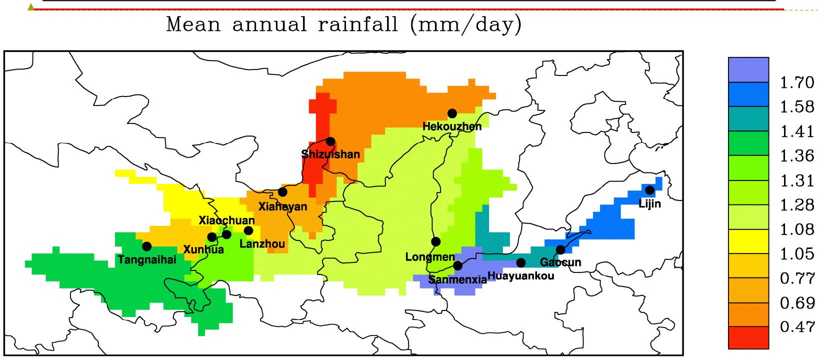


Figure 1. Locations of hydrological stations over the Yellow River basin. Shaded areas are regional mean annual rainfall (mm/day) averaged during 1961-2010.

带格式的: 字体:(中文)+中文正文, 10 磅, 加粗

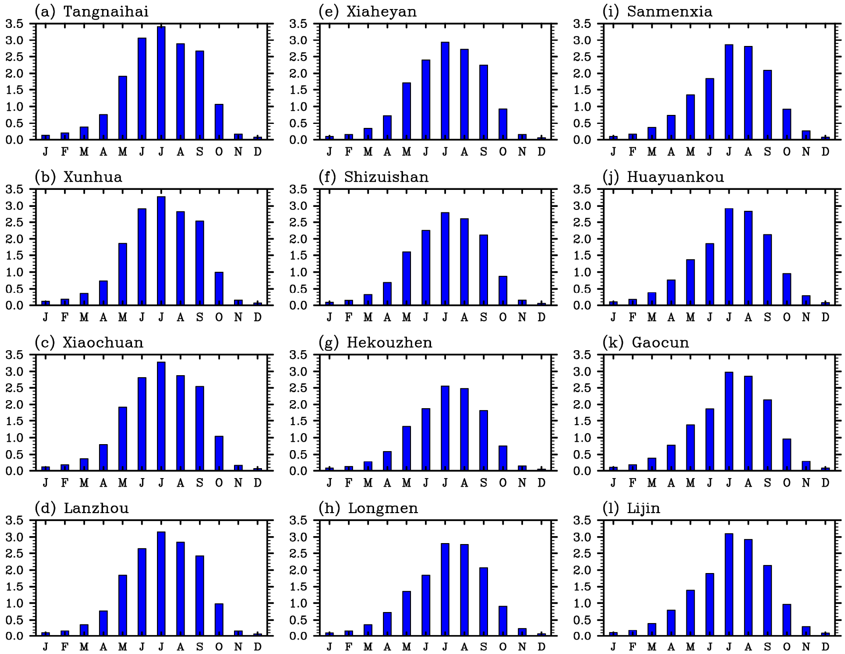


Figure 2. Monthly mean rainfall (mm/day) averaged over 1961-2010 for each sub-basin.

带格式的: 字体: (中文) +中文正文
带格式的: 居中
带格式的: 字体: (中文) Times New Roman, 加粗
带格式的: 字体: (中文) Times New Roman
带格式的: 字体: (中文) Times New Roman
带格式的: 字体: (中文) Times New Roman

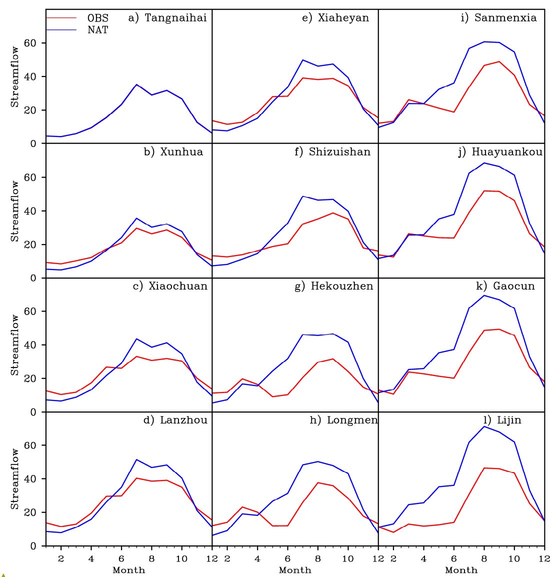
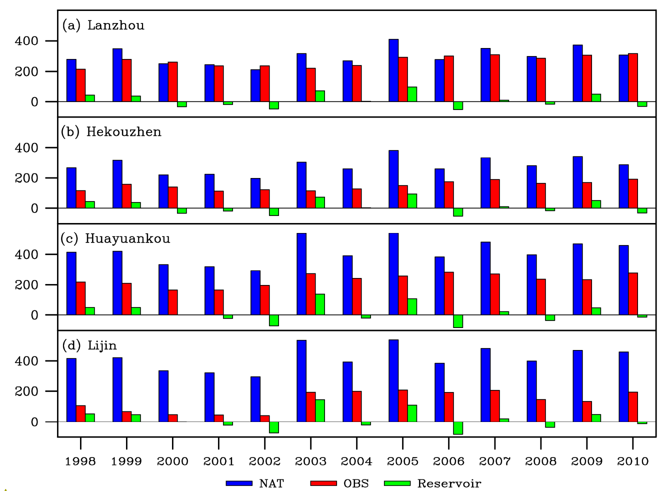


Figure 3. Monthly mean naturalized (blue) and observed (red) streamflow (10^8 m^3) averaged over 1961-2010 for 12 hydrological gauges located from upper to lower mainstream of the Yellow River.



带格式的：字体：(中文) +中文正文
带格式的：居中
带格式的：行距：1.5 倍行距
带格式的：字体：(中文) +中文正文
带格式的：居中，行距：1.5 倍行距

Figure 4. Annual mean naturalized (blue) and observed (red) streamflow (10^8 m^3), and reservoir storage change (10^8 m^3, negative green values represent reservoir water distribution) accumulated within four selected sub-basins (from headwater down to the gauge) during 1998-2010.

带格式的: 字体: 10 磅

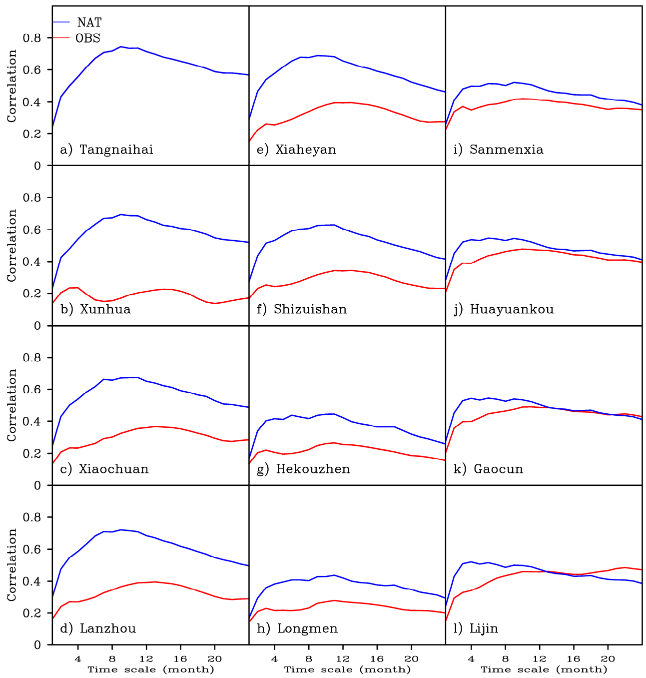


Figure ~~2~~5. Correlations between Standardized Precipitation Index (SPI) at different time scales and monthly naturalized (blue) or observed (red) Standardized Streamflow Index (SSI) for 12 hydrological gauges located from upper to lower mainstream of the Yellow River.

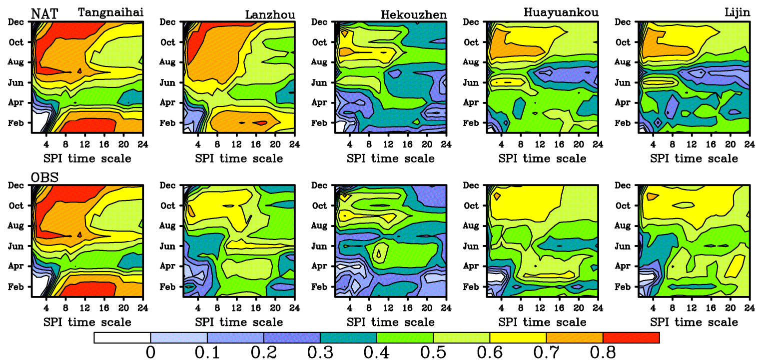
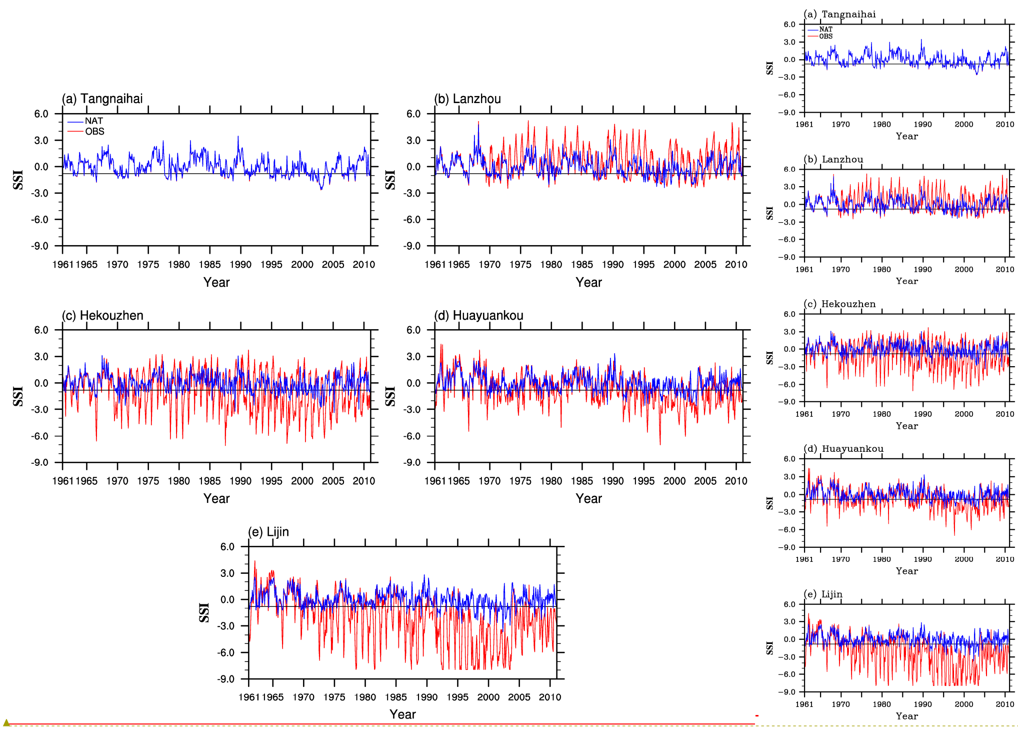


Figure ~~3~~6. The same as Fig. 2, but for each target month for five selected hydrological gauges.



带格式的：字体：(中文) +中文正文

Figure ~~4~~7. Time series of naturalized (blue) and observed (red) 1-month Standardized Streamflow Index (SSI) for five selected hydrological gauges. The horizontal black lines represent the threshold of -0.8 for drought conditions.

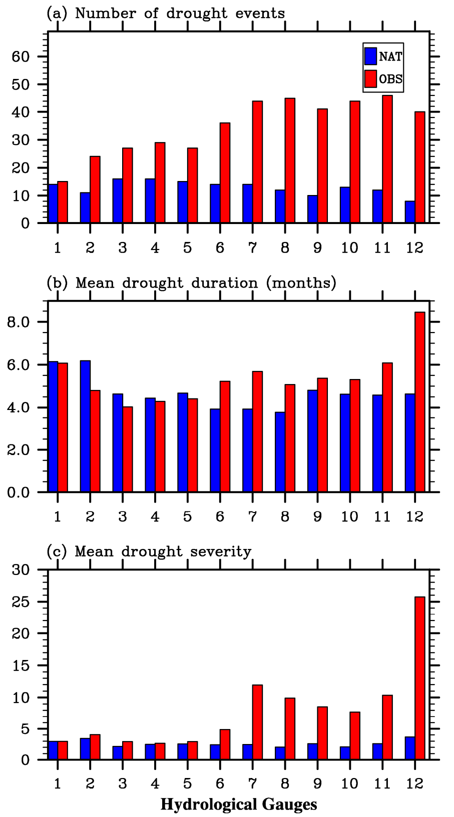


Figure ~~5~~8. Characteristics of hydrological drought events based on the streamflow time series with (red) or without (blue) human influences during 1961-2010 for 12 hydrological gauges. A hydrological drought event is selected when the Standardized Streamflow Index (SSI) is continuously below -0.8 for at least 3 months.

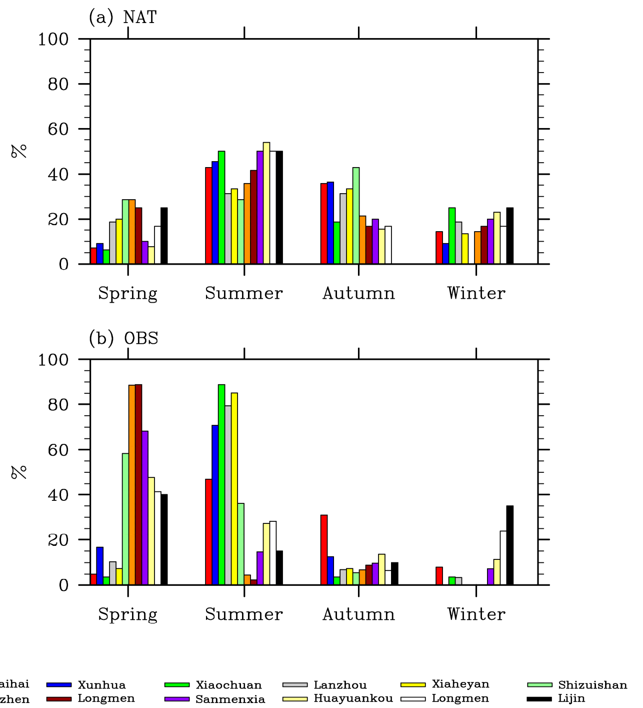


Figure ~~6~~9. Ratio of the number of hydrological drought onsets occurring in different seasons to the total number of hydrological drought events for the 12 hydrological gauges. Drought events are classified the same as those in Fig. 5.

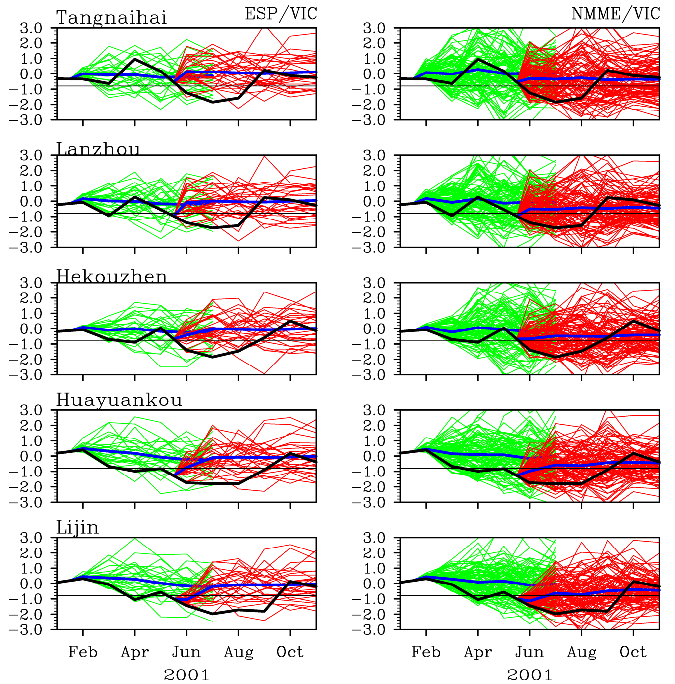


Figure ~~7~~10. Seasonal ensemble hindcast of the 2001 Yellow River hydrological drought from upstream to downstream gauges by using a climatology method (ESP/VIC) and the climate-model-based approach (NMME/VIC). Vertical axes are Standardized Streamflow Index (SSI), where SSI<-0.8 represents a hydrological 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.

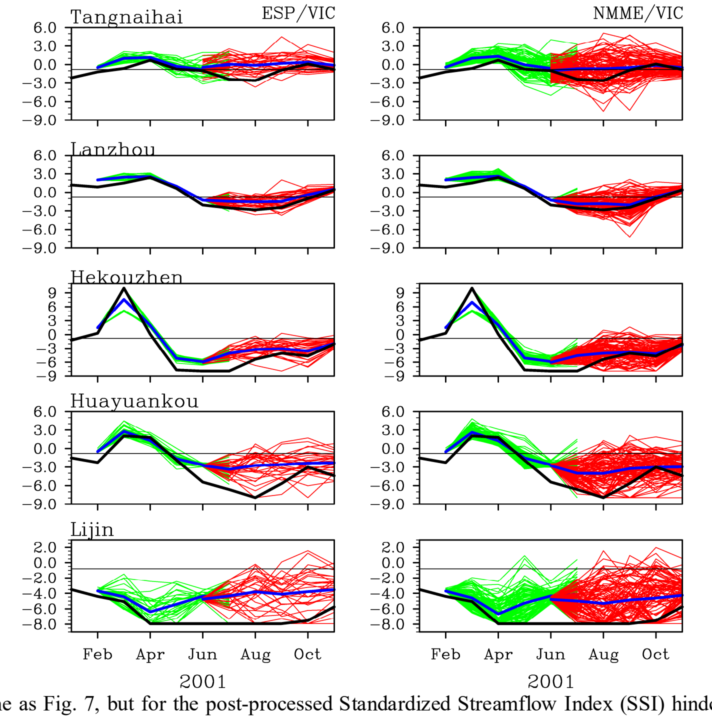


Figure 811. The same as Fig. 7, but for the post-processed Standardized Streamflow Index (SSI) hindcasts from ESP/VIC and NMME/VIC (see Section 2.3 for details) as verified by observed SSI.

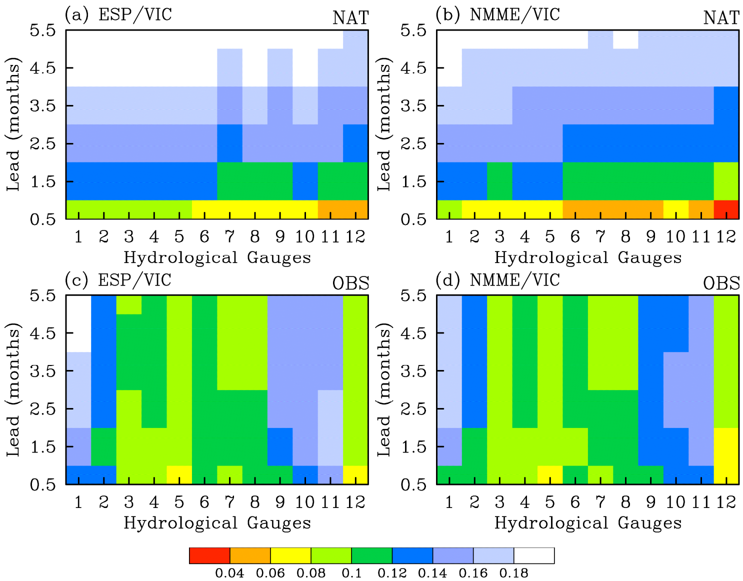


Figure ~~9~~12. Brier score (BS) for ensemble hydrological drought forecasts at different lead times from a climatology method (ESP/VIC) and the climate-model-based approach (NMME/VIC) for different hydrological gauges as verified against VIC offline simulated (a, b) and observed (c, d) streamflow over the Yellow River basin during the period of 1982-2010. For the verification against observed streamflow, the forecasts have been post-processed.

Xing Yuan
Professor/Dr
Institute of Atmospheric Physics
Chinese Academy of Sciences
Beijing 100029, China
Email: yuanxing@tea.ac.cn
Tel: +86-10-82995385
http://www.escience.cn/people/yuanxing
March 22, 2017

Prof. Dennis Lettenmaier
Special Issue Editor
Hydrology and Earth System Sciences

RE: manuscript #hess-2016-592

Dear Prof. Lettenmaier,

Thank you for your kind decision letter on our manuscript entitled “Understanding and seasonal forecasting of hydrological drought in the anthropocene” (hess-2016-592). We have carefully considered your and reviewer’s comments and incorporated them into the revised manuscript to the extent possible. Major changes include adding three new figures and the corresponding text to provide the climate and water management information in details, revising the interpretation of the relation between meteorological and hydrological drought, and clarifying the moviations and methods. We hope that you find the revised manuscript and the response to the reviews acceptable to *HESS*.
The detailed responses to the comments are attached.

We appreciate the effort you spent to process the manuscript and look forward to hearing from you soon.

Sincerely yours,

Xing Yuan

Responses to the comments from Editor

We are very grateful to the Editor for the positive and careful review. The thoughtful comments have helped improve the manuscript. The editor's comments are italicized and our responses immediately follow.

This paper analyzes precipitation and streamflow in the Yellow River basin using the standardized precipitation and streamflow indices (basically just Z-values, where the percentiles come from a fitted gamma distribution). They analyze both naturalized flows (it's not entirely clear where they came from, but probably an attempt by someone to back out the effects of irrigation and reservoir storage) and observed flows. As is well known, irrigation development in the basin has increased rapidly in recent decades, and during the irrigation season flows have been greatly diminished in some reaches. This shows up in their Figure 4, where for basins like Lijin in particular, the observed flows have SSI values that go as low as -9, which would be an absurdly small value for the naturalized flows (for which something in excess of 99% of the values should be between +/- 3).

Response: Thanks for the comments. As explained in the manuscript "… 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)." (P4L1-3 in this response letter)

Yuan et al. (2016) used the naturalized streamflow to calibrate the VIC model without water management model during 1961-1981, when the human interventions were supposed to be limited. The naturalized streamflow was then compared with VIC simulations during 1982-2010, and their NSEs vary between 0.71-0.91. And as shown in Figure S1 below, the naturalized streamflow agrees with VIC simulations quite well. In addition, we also compared the naturalized streamflow with another version of VIC model simulation (Zhou et al., 2016) during 1961-2010, and resulted in NSEs vary between 0.53-0.71 (Fig. S2). So, we believe the naturalized streamflow is a reliable source of data that supports the investigation of human impacts as compared with observed streamflow in this study.

For the low SSI values at Lijin gauge, we have checked the programs and scripts and find that those low SSI values (around -8) are associated with very low streamflow or even zero streamflow. Moreover, the observed SSI was calculated by using the distribution parameters fitted from naturalized streamflow instead of the observed streamflow, this also increases the chance of extremely low SSI values. As shown in Figure S3, the observed streamflow is significantly lower than the naturalized streamflow during 1981-2010. We have also tested the SSI by using observed streamflow to fit gamma distributions, and as commented by the editor, most values are between +/- 3.

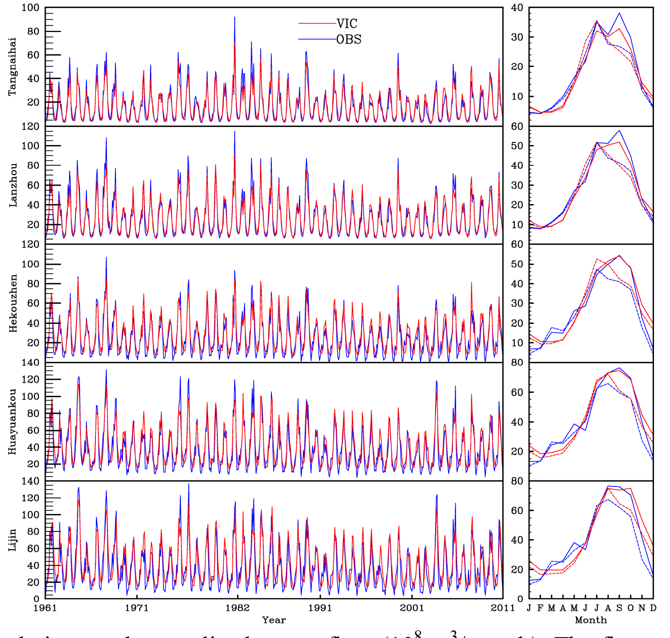


Figure S1. VIC simulations and naturalized streamflow (10^8 m^3/month). The figure was modified from Yuan et al. (2016).

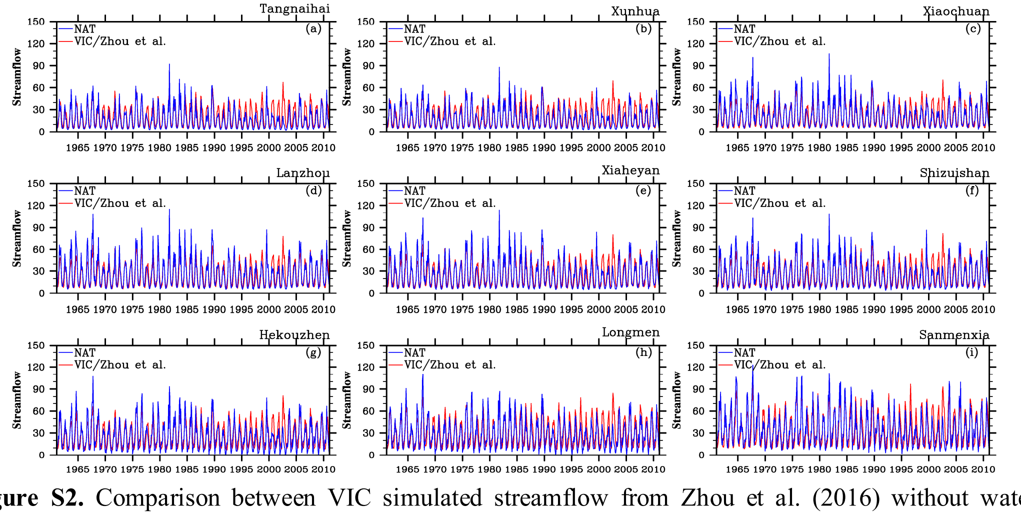


Figure S2. Comparison between VIC simulated streamflow from Zhou et al. (2016) without water management model and the naturalized streamflow (10^8 m^3/month).

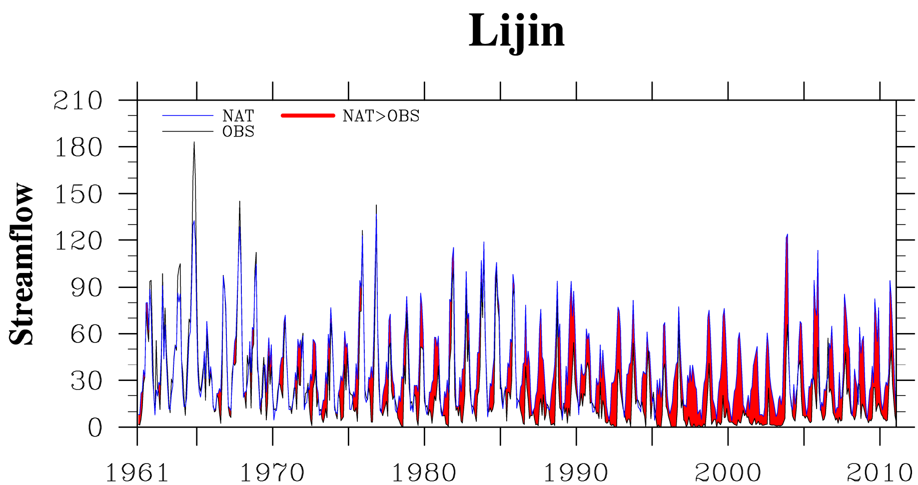


Figure S3. The naturalized and observed monthly streamflow (10^8 m^3/month) at Lijin gauge. The red shaded areas refer to the positive difference between the naturalized and observed monthly streamflow

The problem I have with this paper is the motivation and general approach. Basically, they are analyzing the managed flows as if the management effects were some kind of natural phenomenon. But in fact, the river flows post-management are the result of a set of decisions –either planned, or ad hoc. Those management decisions should be predictable – the irrigated area can be estimated, as can the crop water demands, hence the diversions and return flows. There's a whole branch of water resources systems analysis that does just that. So why treat those decisions as a black box, and do a statistical analysis at all? What they have found essentially is that if you look at the season in which the irrigation diversions are made, the river flows go down. Some of that water gets back into the river, later, and perhaps at a different place, and that effect may increase the flows relative to natural (say, somewhere downstream). But all of that should be predictable at some level.

Response: We totally agree with the editor for the usefulness of water resources systems analysis. This study presents a very preliminary but quantitative analysis for the human interventions impacts on the relationship between meteorological and hydrological drought, and the hydrological drought characteristics (frequency, duration, severity and drought onset seasonality), where the naturalized streamflow that was basically estimated by using such water resources systems analysis, was used to compare with observed streamflow. As we respond above, the naturalized streamflow has accounted for water consumed by agricultural, industrial and civil sectors, and the water regulated by reservoirs, although all the results are shown as an integral anthropogenic impacts in this study. We are now trying to collaborate with local authority to obtain the water use data in each sectors, and hope to incorporate those impacts into hydrological model. But we do believe that assessing the integrated human impacts on hydrological droughts as conducted in this study, can also provide useful information for local water resource management. We have added a discussion on the prospect of water resource model in drought analysis as follows:

"Another popular method is to parameterize the human interventions directly in hydrological models (e.g., Wada et al., 2014; Zhou et al., 2016 among many others), where the irrigated area can be estimated for crop water demand, and the diversion and return flows can also be simulated in those models. This modeling framework could be further pushed forward by the availability of Surface Water and Ocean Topography (SWOT) satellite data in the near future, where surface water storage in reservoirs and rivers can be at 250m resolution (Biancamaria et al., 2016)." (P10L10-15)

Another motivation of this study is to explore the seasonal hydrological drought predictability in the anthropocene, which is to our knowledge among the first in the climate-model-based seasonal hydrological forecasting community. Parts of our findings are that "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." (P1L22-26). So we totally agree with the editor that the management decisions should be predictable to some extent, and considering those predictability would improve the hydrological forecasting in the anthropocene. This study provides a first look on the integrated impacts of human interventions on hydrological drought predictability by using a time series model, and can serve as a benchmark for assessing the physical hydrological model-based approach to fully distinguish the impacts of reservoir regulation, irrigation and groundwater pumping individually in the future. However, the performance of those water management models would also be highly dependent on the data availability, which we believe they are essentially data-driven models in the foreseeable future unless we developed a physical model that fully addresses water and energy balances and the coupling among atmospheric water, surface soil, soil water and groundwater in the anthropocene.

Also, another concern I have – perhaps not so much with the Yellow River, but with the basins referenced in the Zhang et al, (2014) and Wen et al. (2011) studies that they cite as motivation. The Zhang et al. publication is somewhat obscure, and I could only get the abstract. I did read Wen et al., which is a study of a basin in Australia. In Table 1 of Wen, they give the various water management perturbations to the basin, which include construction of what appears to be a couple of km3 of reservoir storage. There must be an operating policy for those reservoirs, and it must be based on an objective function, presumably having something to do with meeting the irrigation demand. Whether or not the policy deals with instream flows at all isn't clear in the paper. My point is that if you look at the statistics of the reliability of the reservoir system in meeting the irrigation demands, presumably it's *higher than without the reservoirs – after all, that's the reason for building reservoirs. But if the operating policy doesn't consider the instream flows, of course eventually enough irrigation will be added to dry up the river. But we don't need a statistical study to tell us that. My concern is that in all of these papers that look at instream drought statistics (including the authors'), that's completely ignored. The situation is slightly different in the Yellow, as I think (I could be wrong) that there isn't currently a lot of storage, so the diversions for irrigation are mostly run of the river. But as I implied above, that could be modeled as well.*

Response: We agree with the editor that the statement in last version of the manuscript might cause confusion about the usefulness of building reservoirs. The reservoirs definitely increase the reliability of meeting the irrigation demand, although it may increase the severity of hydrological drought. Moreover, we also agree with the editor the human activities could be modeled. We have revised the statement as follows:

"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), because many reservoirs were built for meeting the irrigation demand more reliably… An alternative approach is to use a land surface hydrological model (Wada et al., 2013; Zhou et al., 2016) or a less complicated water balance model to recover the naturalized streamflow by assimilating the reported water use data." (P2L20-27)

As we respond above, one of the motivations of the study is to quantify the integral impacts of human interventions on hydrological over Yellow River basin, where the naturalized streamflow data generated by a water balance model play an important role. For the situation in the Yellow River, the editor's comment is absolutely correct, the reservoir plays limited role for the mainstream flows at annual scale (please see Figure 4 in the revised manuscript or our response to reviewer #2 below), but their impacts on local small catchments might be significant.

A final concern I have about the paper is that the ensemble prediction doesn't seem to fit. What was the purpose of including it? Is it to show that if some change in operation was made based on the forecasts, the hydrologic drought statistics would improve? I don't see any argument to that effect. So to me, that part of the paper seems not to fit. I do question the results they show in Figure 7, take for instance for Lijin, which is the sub-basin most affected by diversions for irrigation per their Figure 4. That basin has all kinds of SSI values in the -3 to -9 range, but none of their ensemble members are anywhere close to those values – their smallest forecasts are in the -3 range. The reason of course is that they're using VIC, which (I assume in the version they used) doesn't deal with water management. So they must be forecasting naturalized flows. But who would care about a forecast of naturalized flows? What a management agency needs is a forecast of how much water will be in the river. So we're back to the same thing – to make this paper meaningful, they need a water management model.

Response: I am also a hydro-climate modeler, so I totally agree with the editor that eventually we should have an appropriate physical hydrological model with water management for the hydrological drought forecasting. But to my experiences in modeling and hydrological forecasting, I do not think physical model are ready to resolve everything in the forecasting. The method proposed in this study combines a physical hydrological model (i.e., VIC without human component) with a statistical model (Bayesian model to account for the seasonality of human influences). The disadvantage of this method is that we do not consider the inter-annual variability of human component in the forecasting, which we think it is also very challenging in physical hydrological modeling. Actually we are now collaborating with scientists in USA and China to try to develop a regional climate model that can explicitly account for the surface water storage change and its control on irrigation, in a land-atmosphere coupled mode. The editor's comments and suggestions are very useful for developing such model.

Back to this study and to this comment, we should clarify that human intervention is implicitly considered in Figure 8 (now Figure 11 in the revised manuscript). While for the Figure 7 (now Figure 10 in the revised manuscript), it is to compare with Figure 8 to show the human interventions on hydrological drought forecasting. In addition, Figure 7 itself shows the added value from climate forecasting in a natural world. For the Lijin station, Figures 7 and 8 show a very dry case, where the SSI is extremely low in the observed condition (Fig. 8), but we can see the improvement in the forecasting at its upstream gauges (e.g., Huayuankou, Hekouzhen, etc). The ensemble prediction results show that human intervention can outweigh climate prediction for the hydrological drought forecasting in the anthropocene. Therefore, we would like to keep the ensemble prediction results.

In addition, we have also collaborated with our colleague at PNNL to obtain a set of VIC simulations with and without reservoir regulations and irrigations, and use the difference to represent the human influence, and to correct the hydrological forecasting without considering human interventions. However, it does not perform very well for the hydrological drought forecasting as shown below (Fig. S4). The reasons are multi-folded, but a clear message is that the dynamical-statistical forecasting approach proposed in this study would be a useful intermediate method, before we can collect enough water use data to develop a physically sound model for hydrological drought forecasting.

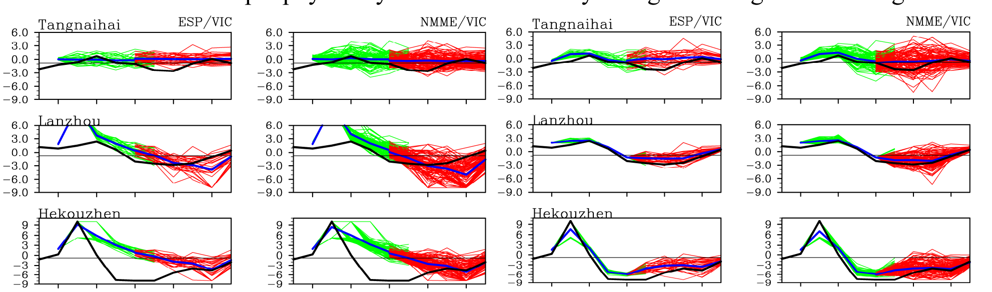


Figure S4. Left: Ensemble SSI prediction correctly by Zhou et al. (2016) VIC simulated human influences on streamflow, as validated by observed SSI. Right: Figure 8 in this manuscript (now Figure 11 in the revised manuscript), where the VIC forecasts are corrected by the Bayesian post-processing method to address the human influence.

I think the authors need to go back to the drawing board with the entire concept, and take a physically based, rather than statistical approach. As it is currently written, I don't find that the paper provides the reader with many insights into causality, which they could do.

Response: Causality is a very good suggestion, which would be our future target to either collecting more human water use data from different sectors with different purposes, or using a physical hydrological model with water management modules, although large uncertainty might exist without reliable water use data as input. While for this paper, we believe in the naturalized streamflow calculated from a water balance model by the local authorities, which is based on abundant water use datasets that is currently not publically available. So we think the quantification of the human influence on hydrological drought characteristic, and on the drought predictability, could provide some insights for the hydrological drought studies in the anthropocene.

References:

Biancamaria, S., Lettenmaier, D. P., and Pavelsky, T. M.: The SWOT mission and its capabilities for land hydrology, Surv. Geophys., 37, 307-337, 2016.

Wada, Y., van Beek, L. P. H., Wanders, N., and Bierkens, M. F. P.: Human water consumption intensifies hydrological drought worldwide, Environ. Res. Lett., 8, 034036, 2013.

Yuan, X., et al.: An experimental seasonal hydrological forecasting system over the Yellow River basin – Part 1: Understanding the role of initial hydrological conditions, *Hydrol. Earth Syst. Sci.*, 20, 2437–2451, 2016.

Zhou, T., et al.: The contribution of reservoirs to global land surface water storage variations. *J. Hydrometeor.*, 17, 309-325.

Responses to the comments from Reviewer #1

We are very grateful to the Reviewer for the positive and careful review. The thoughtful comments have helped improve the manuscript. The reviewer's comments are italicized and our responses immediately follow.

This paper examines hydrological drought in managed river basins in China. Based on the 29-year NMME forecasts, they compared the skill of hydrological drought forecasts between the naturalized and observed conditions. They found that human intervention out weighted the climate variability for hydrological drought forecasts. The paper is well written. I recommend to be published after minor revisions. My specific comments are listed below.

Response: We would like to thank the reviewer for the positive comments. Please see our responses below.

Statistical significance: This may be the weakness of this paper. You compared the correlation between naturalized and observed SPI or SSI. There is no statistical assessment. For example: Fig2. Some gauge like Xunhua, the differences are significant, but differences for gauge Lijin may be not. You need to add statistical significance test to results.

Response: Thanks for the comments. We have incorporated the statistical significance testing and revised the manuscript as follows, where Fig. 2 is Fig. 5 now due new addition of three figure before it.

"The correlations for observed streamflow are significantly lower than for naturalized streamflow for gauges from Xunhua down to Huayuankou, with p values less than 0.01 (Figs. 5b-5j). There is also a significant difference in correlation for Gaocun gauge with $p<0.05$ (Fig. 5k), but the difference is not statistically significant for Lijin gauge with $p>0.1$ (Fig. 5l)." (P6L11-14 in this response letter)

Is SSI similar to standardized runoff index (Shukla and Wood 2008) except you use streamflow?

Response: Exactly. We have clarified it in the revised manuscript as follows:

"Note that SSI was similar to the standardized runoff index (SRI) defined by Shukla and Wood (2008), except that streamflow was used here for a standardization." (P4L32-33)

Section 2.3 Please add more details For the VIC simulation, what are the sources for daily precipitation and temperature time series used to derive forcings? Did you run the VIC model for these 12 gauge sites or the whole domain? Did you use the VIC in the water balance mode (no observed radiation terms)? which version?

Response: Thanks for the comment. We have clarified the information for the VIC simulation as follows:

"The observed meteorological forcing datasets including daily precipitation, daily maximum and minimum surface air temperature, and surface wind were interpolated from 324 China Meteorological Administration stations. The VIC model version 4.0.5 was used to predict runoff in a water balance mode over the entire Yellow River basin with 1321 grid cells at 0.25-degree resolution, and a routing model was used to translate the runoff into streamflow at each 0.25-degree grid cell, and to route the flow into rivers and finally into the ocean (Yuan et al., 2016)." (P5L13-18)

Drought is usually defined as persistent low flow conditions. Does naturalized drought persist longer? Please comment on the persistence of low flow (SSI) conditions.

Response: In this study, the monthly streamflow records were converted into percentiles to represent the low flow conditions at seasonal time scale. A hydrological drought event was defined as follows:

"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)…" (P5L3-4)

As shown in Figure 5b (Figure 8b in the revised manuscript), the naturalized drought persists a little longer than the observed streamflow at upper gauges, suggesting the positive influence of human intervention for reducing the drought duration. However, the former is basically shorter than the latter at middle and lower gauges, mainly due to intensive human water consumption. We clarified the duration changes in the manuscript as follows:

"For the drought duration in natural conditions, it is generally longer over upper reaches which is again due to a drier climate (Fig. 8b). With human interventions, there is a slight decrease in drought duration for the upper gauges down to Xiaheyan gauge. This suggests that human interventions reduce the persistency of drought over the upper reaches of the Yellow River basin. From Shizuishan gauge (the 6^{th} gauge in Fig. 8b) down to the outlet, the duration of hydrological drought increases by 12%-83% under human interventions." (P7L14-18)

You used NMME forecasts. Did you perform hydroclimate forecasts using VIC for each model separately and then took the ensemble means? How exactly did you process the NMME data? Readers need more details on that.

Response: Yes, the VIC model was driven by each ensemble members of nine NMME models with a grand ensemble of 99 realizations, and ensemble means were then calculated to evaluate the deterministic forecast skill. In addition, all 99 realizations were also used directly to evaluate the probabilistic forecast skill.

The downscaling process is the same as Yuan (2016). To have this paper self-contained, we have added some descriptions in the revised manuscript as follows:

"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). Here, the 1-degree NMME global hindcasts of monthly precipitation and temperature were bilinearly interpolated into 0.25-degree, the interpolated monthly hindcasts for each NMME model were then bias-corrected independently against observations by using the quantile-mapping method (Wood et al., 2002) in a cross-validation mode (i.e., dropping observation and forecast in the target year when generating the climatology), and these bias-corrected monthly hindcasts were finally temporally disaggregated to daily by historical sampling and rescaling (Yuan, 2016)." (P5L21-27)

You stated seasonal cycle plays a role in drought. (page 5 response time is different for summer and winter). How large is the precipitation seasonal cycle?

Response: Thanks for the comment, we have added a figure to show the precipitation seasonal cycle (Figure 2 in the revised manuscript), and have clarified the seasonal cycle as follows:

"Most precipitation over Yellow River occurs in summer season due to the influence of East Asian monsoon, resulting in a strong seasonality of precipitation, with more than 80% of annual precipitation falls within May-September (Fig. 2)." (P3L27-29)

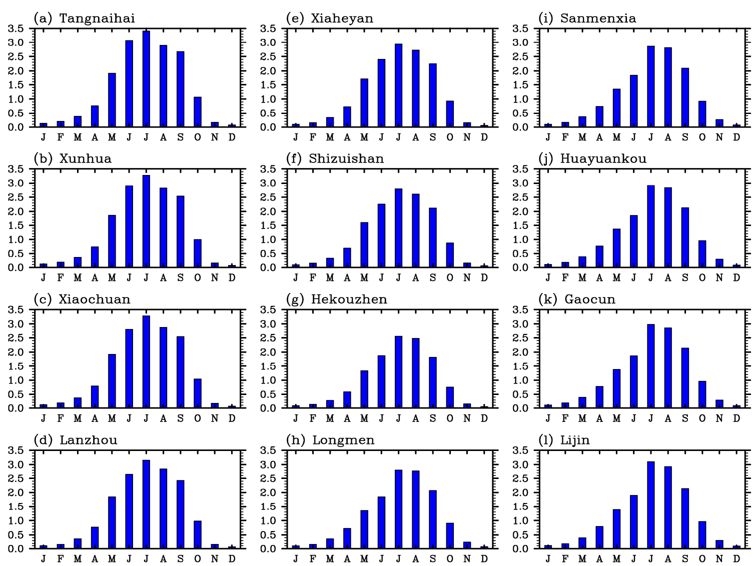


Figure 2. Monthly mean rainfall (mm/day) averaged over 1961-2010 for each sub-basin.

References:

Shukla, S., and Wood, A. W.: Use of a standardized runoff index for characterizing hydrologic drought, Geophys. Res. Lett., 35, L02405, doi:10.1029/2007GL032487, 2008.
Wood, A. W., Mauer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern United States, J. Geophys. Res., 107, 4429, doi:10.1029/2001JD000659, 2002.
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.

Responses to the comments from Reviewer #2

We are very grateful to the Reviewer for the positive and careful review. The thoughtful comments have helped improve the manuscript. The reviewer's comments are italicized and our responses immediately follow.

The paper presents a fairly interesting study on an important topic with substantial results and insights. The research therein is a good fit for HESS. The main focus is on the impact of human water use/regulation activities on drought. The authors also carried out a number of seasonal meteorological/hydrological forecast experiments and I find them very carefully designed and carried out. The results/discussions are clearly presented too. My major concerns are on the analysis methodology and the adequacy of supporting information. The study area is one large river basin in China while a quite minimum level of specific information on the local water management is provided. Usually, more information on the surface water use practices will be very useful in helping readers understand the findings and their implications across similar areas in other parts of the world. I recommend its publication in HESS with improvements on the analysis method and additional discussion on local water management and how that leads to what is seen in the results.

Response: We would like to thank the reviewer for the positive comments. We have added two new figures and the corresponding text to provide the water management information in details, and have revised the interpretation of the relation between meteorological and hydrological drought. Please see our responses below.

Specific Remarks

The paper (circa P. 5, L. 3-18) interprets the peak correlation time scale as the "optimal response time of streamflow to sub-basin averaged precipitation", while offering no supporting evidence (e.g. citation of previous research, data results). The 6-12 months (and later 8-16 months) "response time" seems incredibly long and beyond what a hydrologist can reasonably expect. Given the size of the Yellow River basin, it shouldn't take more than a month or two for water to travel from rain-falling hillslopes down to river gauging stations. And the local soil water stores or snowpack won't be able to defer the release of precipitated water for that long either. SPI/SSI does time averaging to the underlying parameters and this essentially smooths out noises at shorter time scales. A true "response time" is usually calculated from time lagged correlation analysis, e.g., between SPI-1 and SSI-1. Either the "response time" needs to be calculated differently or the same calculations need to be interpreted differently. Note that the change in the relationship between meteorology and hydrology is one of the major points in the paper as summarized in the abstract.

Response: We greatly appreciate the positive comment. We calculated the response time as suggested by the reviewer and found that the most significant lag correlations occur at lag-1 month both for naturalized and observed monthly streamflow along the mainstream of Yellow River, although the lag-correlations are again lower for the observed streamflow.

Actually in the last version of the manuscript, we followed the work done by Vicente-Serrano and López-Moreno (2005), and thought the time scale with the maximum correlation is considered as the time scale of SPI that streamflow responds to, i.e., the SPI time scale that has the most similar variations to the SSI. However, we have realized that using "response time" in the manuscript would be very confusing. So, we have removed all "response time" throughout the paper, and have re-written the corresponding text as follows:

Abstract—"It is found that human interventions decrease the correlation between hydrological and meteorological droughts, and make the hydrological drought respond to longer time scale of meteorological drought especially during rainy seasons." (P1L14-16 in this response letter)

Section 3.1—"Similar to Vicente-Serrano and López-Moreno (2005), the time scale with the maximum correlation is considered as the time scale of SPI that streamflow responds to. For the naturalized streamflow, it responds to 6-12 months SPI over the upper and middle reaches of Yellow River, and to about 4 months SPI over the lower reaches … Except for the Tangnaihai gauge in headwater region, the SPI time scales with the maximum correlation are longer for the observed streamflow than the naturalized streamflow, suggesting that human interventions basically make the hydrological drought respond to longer time scale of meteorological drought … It is found that streamflow responds to shorter time scale of SPI in wet and warm seasons and longer time scale in dry and cold seasons." (P6L4-7, L14-16, L24-26)

Concluding Remarks—"Comparison between naturalized and observed SSI at 12 hydrological gauges along the mainstream 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 decrease the correlation between hydrological and meteorological droughts, and make the hydrological drought respond to longer time scale of meteorological drought especially during rainy seasons." (P9L14-17)

Further, the notion of "nonlinear response" of hydrological drought to meteorological drought is a bit vague in the discussions. The rainfall-runoff process is by itself "nonlinear" and lagged in time, at least at short time scales. If the word "response" refers to the rainfall-runoff process (at any time scale), the research here should try to find out what exactly human interventions did to that process. Reduction in streamflow volume (i.e. significant amount of consumptive use)? Longer lag times (delayed release for flood control)? If the lag times become longer, should this be considered in the forecast post-processing procedure? (For example, a time series based procedure that looks at a prior history instead of just the current month.)

Response: We have also removed "nonlinear response" in the revised manuscript given that the focus of this work is not to investigate the lag-correlation for forecasting. We have plotted the annual cycle of naturalized and observed streamflow in Figure 3 to show the effect of human interventions, and revised the section “2.1 study domain and hydroclimate observation data” as follows:

“In general, human interventions decrease streamflow over upper and middle reaches of Yellow River during rainy season while increase it during dry season (Figs. 3b-3h). This suggests that reservoirs in the upper and middle reaches of Yellow River store rain water in wet season and distribute it in the remaining time of the year according to the need, which is similar to regulations in other parts of the world (Wada et al., 2014). Actually, Figure 4a shows that the annual mean observed streamflow at upper reaches can be higher than the naturalized streamflow during dry years due to the reservoir water release (e.g., years 2000, 2002, 2006 and 2010 for Lanzhou gauge). Over the lower reaches, the observed streamflow is significantly lower than the naturalized streamflow during wet season due to heavy water consumption (Figs. 3i-3l), but the former is close to the latter during dry season because of no significant water consumption or reservoir management. (Figs. 4c-4d)” (P4L6-13)

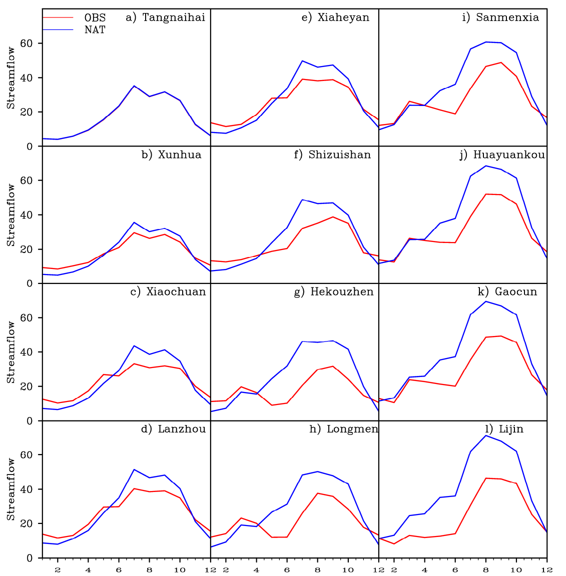


Figure 3. Monthly mean naturalized (blue) and observed (red) streamflow (10^8 m^3) averaged over 1961-2010 for 12 hydrological gauges located from upper to lower mainstream of the Yellow River.

For the same reasons, more information on the water regulation practices in the study area is needed for a (much) better understanding of the impacts and differences found in the results. For example, reservoirs may store rain water from wet season and distribute it in the remaining time of the year according to the need. How much of the streamflow water is being regulated in the Yellow River basin *(e.g. reservoir capacity relative to the annual total inflow) and for what purposes? How much of the streamflow is being modified (in both absolute and relative senses)?*

Response: We thank for the comment. We have now collected annual statistics for the reservoir storage change during 1998-2010, but failed to obtain the monthly data. Based on the data available, we have added Figure 4 to show the interannual variations of naturalized and observed streamflow and the reservoir storage change, and we have revised the manuscript as follows:

"Figure 4 also shows that the magnitudes of reservoir storage changes are quite small as compared with streamflow. In fact, the mean absolute changes of reservoir storage during 1998-2010 are about 14%-38% and 12%-14% of observed and naturalized streamflow, respectively. This suggests that other human interventions, such as direct withdrawal of surface water for agricultural, industrial and civil consumptions, account for a large part of streamflow variations over Yellow River." (P4L14-18)

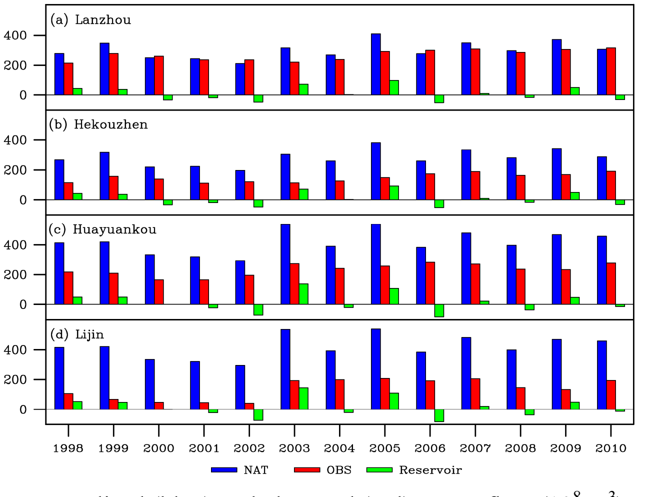


Figure 4. Annual mean naturalized (blue) and observed (red) streamflow (10^8 m^3), and reservoir storage change (10^8 m^3, negative green values represent reservoir water distribution) accumulated within four selected sub-basins (from headwater down to the gauge) during 1998-2010.

Fig. 4 helps to understand the scenario but direct comparisons between observations and naturalized values (in seasonal cycles and annual totals) can help explain what happened in Fig. 4 in a much better way. I guess the observed SSI in Fig. 4 is calculated against observed flow climatology and naturalized SSI against naturalized flow, right? (Please clarify.) If so, the comparisons between the two do not *reveal the difference between the observed and naturalized climatologies, e.g. reduced total flow volumes or lagged peak times.*

Response: Observed SSI in Fig.4 is not calculated against observed flow climatology. Actually both naturalized and observed SSI are calculated against the naturalized flow climatology, so they can be compared to detect the effect of human interventions on hydrological drought. Seasonal cycle of original values are now shown in Figure 3 (please see our response above) to support the SSI analysis.

Specific information on the local water management and water use practices is always helpful in understanding the findings and their implications across similar areas in other parts of the world (Wada et al., 2014). The study could be significantly stronger if more specific water management information is provided and related to the research findings.

Response: Thanks for the comment. Two figures regarding the seasonal cycle of monthly naturalized and observed streamflow, and the annual mean streamflow and reservoir storage change have been added into the revised manuscript. Please see our responses above.

P. 5, L. 13: nonlinearly -> nonlinear

Response: Revised as suggested.

Fig. 1: The map needs to show at least the Yellow River and its main tributaries (thicker lines for the main stream) under this study. Replace the political boundaries with sub-basin boundaries (keep the coast lines).

Response: We have revised Figure 1 as suggested.

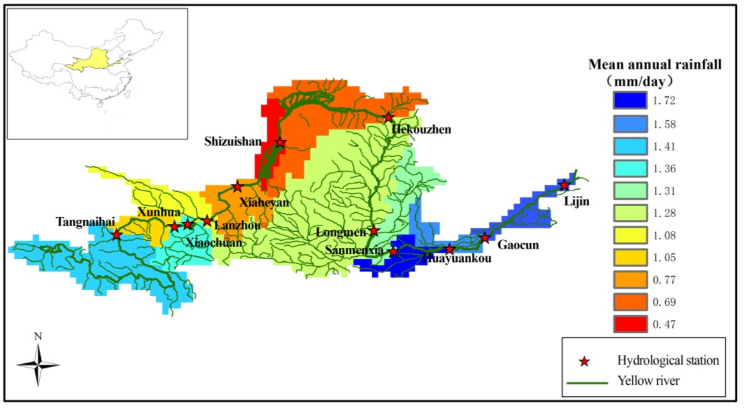


Figure 1. Locations of hydrological stations over the Yellow River basin. Shaded areas are regional mean annual rainfall (mm/day) averaged during 1961-2010.

Fig. 4: SSI at what time scale? 1-month? Subplots are too small and better if they are rearranged into multiple columns.

Response: Fig. 4 (Fig. 7 in the revised manuscript) has been replotted to show the panels in two columns. The SSI is at 1-month time scale, and it has been clarified in the revised figure caption:

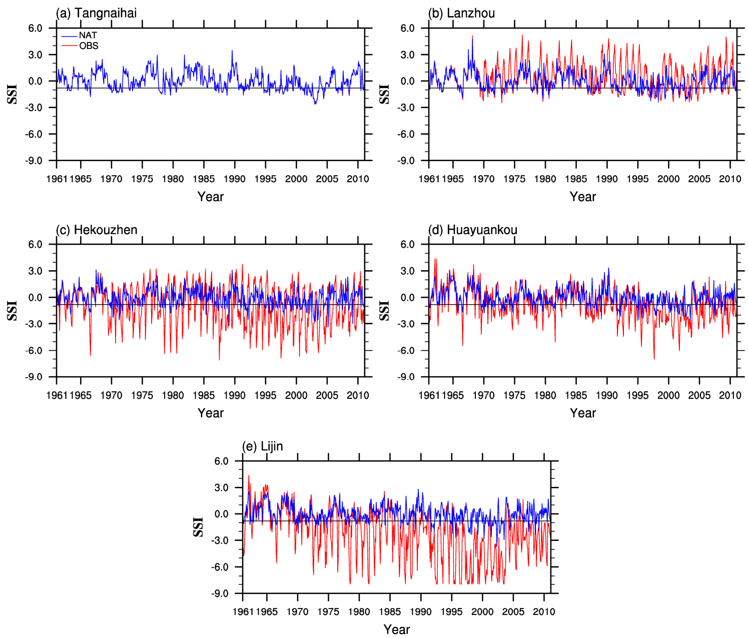


Figure 7. Time series of naturalized (blue) and observed (red) 1-month Standardized Streamflow Index (SSI) for five selected hydrological gauges. The horizontal black lines represent the threshold of -0.8 for drought conditions.

References:

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.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dynam., 5, 15-40, doi:10.5194/esd-5-15-2014, 2014.