Reply to Editor's comments on "Irrigation, damming, and streamflow fluctuations of the Yellow River"

Z. Yin on behalf of all co-authors

1 Point-by-point replies to Editor's comments

EC: "Thank you for your detailed responses. I would say most of the reviewers' comments have been addressed except the model performance. I also raised that issue before external review."
AC: We would like to thank you very much for your kindly support during the reviewing process and constructive suggestions. Below is our point-by-point replies to your comments.

EC: "Isn't it possible to calibrate the model without reservoir and irrigation against the naturalized streamflow ..."
AC: Thank you very much for the suggestion. We did consider this suggestion carefully. However, we found that the naturalized streamflow contained a lot of uncertainties, which may be not suitable for model calibration in our case.

Firstly, the naturalized streamflow is not equivalent to the real streamflow without human perturbations (e.g., simulated streamflow without anthropogenic factors), if interactions between irrigation and local climate is considered. The naturalized streamflow overlooked the impact of irrigation on evapotranspiration via decreasing surface temperature (T_s), therefore the potential evapotranspiration (i.e., the atmospheric demand). The consequence is that the reconstructed streamflow may be lower than the real natural streamflow (see Sect. 2). Therefore we are worried that a model calibrated against naturalized streamflow may bring unexpected uncertainties to the land surface processes, because our model captured significant irrigation-induced cooling, which has been verified by numerous studies at global scale (Bonfils and Lobell, 2007; Luyssaert et al., 2014; Sacks et al., 2009; Thiery et al., 2017) and over China (Yang et al., 2020).

Secondly, an indirect evidence shows that the effective irrigated area may be underestimated in several provinces (see Sect. 3). The statistical irrigation water consumption is based on monitoring of large irrigation districts, which can be regarded as the lower limit of real irrigation consumption. Therefore, we infer that bias might exist in the water consumption for irrigation in census data which further affects the accuracy of naturalized streamflow.

Thirdly, we found that the water balance is not closed if all observations and census data are reliable (see Sect. 4). It confirms that if a model is calibrated by naturalized streamflow, it might lead to unrealistic trend of total water storage (TWS).

We stress that we don't disagree with models calibrated by naturalized streamflow for short-term prediction. We applaud the efforts of the YRCC to remove anthropogenic impacts on the streamflow of the Yellow River, which is very helpful for the planning of water conservancy projects (Li et al., 2001). We remain convinced, however, that if a model includes interactive mechanisms among water cycle, local climate variation, and human activities, calibration against naturalized streamflow may bring unexpected uncertainties.

Nevertheless, we attempted to re-construct the simulated naturalized streamflow under our IR simulation framework and compare it to the naturalized streamflow based on census data. The results (Table R1) showed that the simulated naturalized streamflow can reflect the features of naturalized streamflow (coefficient of determination $R^2 > 0.45$; index of agreement $d > 0.7$). More importantly, the mean annual values of simulated naturalized streamflow at HuaYuanKou and LiJin stations are 1654.7 and 1683.2 $\mathrm{m}^3.\mathrm{s}^{-1}$, only 1.4% and 2.7% higher than that based on census data, respectively.

In the revision, we have added the Table R1 in the online supplementary, and mentioned this comparison (Page 9, Line 224–225) as "A validation against naturalized streamflows is shown in Table S1.".

EC: "... and then calibrate that with reservoir and irrigation against the observed streamflow?"

AC: We agree that calibration is helpful to adjust unmeasured parameters of a model in an area of interest to improve its performance. However, a precondition is that the model has included all key biophysical processes of the system, at least in the specific region. If several key processes are missing in the model, we are afraid that a detailed calibration will cover up the shortcomings, which is not helpful for either mechanism detection or for long-term prediction. For example, the HBV model was well-calibrated by different approaches in 156 catchments in Austria but failed in predicting the streamflow change after then (Duethmann et al., 2020). One of the key reasons is that the vegetation dynamics is not described in the HBV model.

Regarding to the YRB, many models were used to simulate the streamflows (Liu et al., 2020; Xi et al., 2018; Yuan et al., 2018). However, as far as we know, most model studies only focused on the natural part, simulating the natural streamflow and comparing to the naturalized streamflow. Studies with model simulation validated against observed streamflow are quite rare. Jia et al. (2006); Tang et al. (2008) are the only two studies having notable simulated streamflows validated against the observations. However, Jia et al. (2006) used census irrigation and reservoir operation data as inputs. The physical processes of those anthropogenic factors are not integrated in the model. Tang et al. (2008) coupled the irrigation process in the DBH model and investigated the long-term trend of annual streamflow. But the irrigation demand is driven by satellite-retrieved LAI, suggesting that the crop dynamics is not fully represented by physical laws. Several global hydrological models are able to simulate both irrigation and reservoir operation processes, and were applied on future projection of water resources either regionally (Liu et al., 2019) or globally (Hanasaki et al., 2018; Wada et al., 2014, 2016). Although all studies admitted the complicated situation of the YRB, none has focused on the YRB and providing a complete error budget of the catchment water balance.

Therefore, by developing a irrigation module in ORCHIDEE and a reservoir operation model, we try to demonstrate whether the streamflow fluctuations of the YR can be reproduced by generic parameterizations, which is the primary objective of this study. We think our aim has been reached and calibration is not compulsory at current stage.

Firstly, we developed physical-process based irrigation module and reservoir operation model to simulate streamflows of the YR. Our model only requests basic atmospheric forcings (incoming short/long wave radiation, surface pressure, 2 m humidity and temperature, snowfall, rainfall, and u-v wind speeds) and limited boundary conditions (e.g., PFT fractions, soil texture, and CO_2 concentration). Crop dynamics, irrigation, and dam operations are mechanistically simulated without any observations as drivers. By running the model, we already can give a reasonable simulation of streamflows (mean $d > 0.61$; mean mKGE> 0.4) and reservoir operations ($r \approx 0.9$ of LongYangXia+LiuJiaXia) with generic parameters. This is already a large amount of work in the perspective of model development.

Secondly, our results show that the errors of simulated streamflow decrease dramatically after considering irrigation and reservoir operations, suggesting that both irrigation and reservoir operations are first-order perturbations of streamflow fluctuation. Our first objective has been approached. However, during our diagnosis, we found that three mechanisms associated with water management (see our reply to the next comment) have to be considered in the model before introducing detailed calibration of the entire system.

Thirdly, a comprehensive calibration against observed streamflow needs more data as support. For example, to fit the seasonality of observed streamflows, we can either modify soil and routing related parameters (e.g., the maximum infiltration rate, flow velocities of surface and subsurface runoff, etc), or reservoir operation related parameters. Although both calibration approaches may improve the simulated streamflow, it is hard to say which way reflects the reality. Obviously, these two ways will give different future projections under climate change! Thus it is important to demonstrate which factor results in the mismatch before calibration. However, both of them are difficult to diagnose due to lack of data. The former (soil and routing parameters) is hard to be measured and the latter (reservoir operations) has rare public data. Nevertheless, with the upcoming Surface Water and Ocean Topography (SWOT) mission, it will be possible to monitor the water level and surface extent of reservoirs, which will be helpful to improve and validate the reservoir operation simulations (Ottlé et al., 2020).

In summary, detailed calibration is important and necessary to provide a precise estimation of streamflows fluctuations and future projection. However, it is the step after model development and mechanism diagnosis which confirms that all known mechanisms has been considered in the model. Our study developed irrigation module and a reservoir operation model, and provided reasonable simulated land surface elements (soil moisture (Yin et al., 2018), irrigation amount, total water storage (Yin et al., 2020), leaf area index (Xi et al., 2018), crop phenology and yield (Müller et al., 2018; Wang, 2016; Wang et al., 2017), evapotranspiration, streamflow, multi-year mean reservoir operation) with generic parameterizations and limited inputs. Furthermore, we uncovered several processes that may strongly affect the streamflow fluctuations, which should be described in the model before comprehensive model calibration. In addition, more observations (like reservoir operation) should be collected to trace the source of uncertainty and preform accordingly calibration.

Therefore, we think that extensive calibration is necessary but is not compulsory at current stage. The priorities of mechanism integration and extensive calibration has been discussed in the manuscript and in our reply to the 3rd comment of the first referee. In the revision, we firstly demonstrated the importance and priority of key mechanism integration in the Introduction as: (Page 4, Line 56–67) "To model present water resources in the YRB and make future projections, not only natural mechanisms, but also anthropogenic ones must be represented in a model. If a key mechanism is missing in a model, a calibration of its parameters to match observations can compensate for structural biases and projections may be erroneous. For example, the HBV model (Hydrologiska Byråns Vattenbalansavdelning) was well-calibrated with different approaches in 156 catchments in Austria but failed in predicting streamflow changes due to climate warming (Duethmann et al., 2020). One of the key reasons being that the response of vegetation to climate change was missing in the model. In this study, we integrate two key anthropogenic processes (irrigation and dam operation) in the land surface model ORCHIDEE (ORganizing Carbon and Hydrology in Dynamic EcosystEms) which has a mechanistic description of plant-climate and soil water availability interactions and of river streamflows. Through a set of simulations with generic parameter values, we aim to preliminarily diagnose how irrigation and dam operation improve the simulations of observed YRB streamflows. After making sure we understand the impact of adding these two new and crucial processes, the model will be calibrated against a suite of observations so that it can be applied for future projections."

In the Discussion, we showed our opinion first before extending detailed discussion as: (Page 15, Line 415–421) " ... for long-term projections, a model should include all key processes of the system studied. If key processes are missing in the model, a calibration will cover up the shortcomings, which lead to lack of predictive capacities for long time scales as shown by Duethmann et al. (2020). Therefore, by developing crop physiology and phenology, irrigation, and (offline) dam operation model, we have tried to demonstrate that streamflow fluctuations of the YR can be reasonable reproduced by a generic land surface model. Although mismatches exist in the simulated streamflows, they are more likely caused by missing processes (joint impact of multiple medium reservoirs, special mission of dams, irrigation system characteristics) than by poor calibration of existing processes, ..."

In addition, we adimt that this study is still at the early stage (mechanism integration and diagnosis) to model the hydrological cycles in the YRB. Thus we underlined that our study provided preliminary diagnosis of the hydrologicla cycle in the YRB under human perturbations in the abstract (Page 1, Line 5), introduction (Page 5, Line 104), and conclulsion (Page 15, Line 433).

EC: "More discussions on this issue, including future plans on how to improve the model performance is necessary."
AC: In this study, we demonstrated that irrigation and dam operation are first-order perturbations to streamflow fluctuation in the YRB. Besides, we discovered several key mechanims missing in current LSMs and GHMs, which strongly affected the streamflow of the YR as well. They are: 1) specific operation rules of dams (e.g., water-sediment mission); 2) integrated catchment management in extreme years; and 3) Effects of artificial reservoirs to local climate. Those processes must be included in the model. Furthermore, the dam operation model will be integrated in ORCHIDEE. Finally, extensive calibration will be applied to reduce the simulation errors of the entire hydrological cycles in the YRB (not only streamflow, but also LAI, evapotranspiration, soil moisture, surface temperature, etc).

These overlooked mechanisms has been discussed in the Sect. 4. In the revision, we introduced our plan to solve these issues, as: (Page 14, Line 386–389) "... are difficult to be well modeled due to lack of data. However, with the upcoming Surface Water and Ocean Topography (SWOT) mission, it will be possible to monitor the water level and surface extent of more reservoirs, which will be helpful to improve and validate the dam operation simulations (Ottlé et al., 2020)." (Page 15, Line 426–429) "In summary, our results show that the errors of simulated streamflows decreased dramatically after considering crops, irrigation, and dam operations, suggesting that these are first order mechanisms controlling streamflow fluctuations. Future work can be focused on completing the model by linking dam operation to the variable crop water demand."

EC: "Please refer to those crop/reservoir model papers not only over Yellow River, but also other areas around the globe."
AC: More than 20 references about crop/reservoir models, or module developments in LSMs and GHMs has been cited. For instance, (Page 4, Line 73–77) "... as well as other site-based crop models (e.g., EPICs (Folberth et al., 2012; Izaurralde et al., 2006; Liu et al., 2007, 2016; Williams, 1995), CGMS-WOFOST (de Wit and van Diepen, 2008), APSIM (Elliott et al., 2014; Keating et al., 2003), and DSSAT (Jones et al., 2003)) and land surface models (e.g., CLM-CROP (Drewniak et al., 2013), LPJ-GUESS (Smith et al., 2001; Lindeskog et al., 2013), LPJmL (Waha et al., 2012; Bondeau et al., 2007), and PEGASUS (Deryng et al., 2011, 2014)) (Wang et al., 2017; Müller et al., 2017)." and (Pages 5, Line 88–90) "However, the representation of dam operations in many global hydrological studies (e.g., Droppers et al. (2020); Haddeland et al. (2006, 2014); Hanasaki et al. (2018); Zhao et al. (2016); Yassin et al. (2019); Wada et al. (2014, 2016)) are based on the ideas of Hanasaki et al. (2006).".

2 Naturalized streamflow overlooked human–local climate interactions

The naturalized streamflow (Q_{nat}) from YRCC (Li et al., 2001) is calculated by:

$$Q_{\mathrm{nat}} = Q_{\mathrm{obs}} + C_{\mathrm{agr}} + C_{\mathrm{ind}} + C_{\mathrm{urb}} + \Delta W_{\mathrm{res}}. \tag{1}$$

Q_{obs} is observed streamflow. C_{agr}, C_{ind}, and C_{urb} are water consumption of agriculture, industry and urban, respectively. ΔW_{r} is water storage change of reservoirs. The water consumption is the difference between water withdrawal from river channels and return flows.

Irrigation water consumption (C_{irr}) is the main component of water consumptions of the census data for streamflow naturalization. It is estimated as the deficit between irrigation water withdrawal (W_{w}) and return flow (W_{r}) to the river channel (as in Fig. R2). The total evapotranspiration in the cropland (ET_{t}) consists of ET rooted from precipitation (ET_P) and ET rooted from irrigation (ET_I). If no irrigation occurs, the long-term water balance in the cropland (the soil moisture change is neglected) can be written as:

$$P = \mathrm{ET}_P^{\mathrm{nat}} + W_{\mathrm{r}}^{\mathrm{nat}}. \tag{2}$$

The superscript $_{\mathrm{nat}}$ indicates variables under natural process only. If irrigation is applied, the water balance will be:

$$P + W_{\mathrm{w}} = \mathrm{ET}_P^{\mathrm{irr}} + \mathrm{ET}_I + W_{\mathrm{r}}^{\mathrm{nat}} + W_{\mathrm{r}}^{\mathrm{irr}}. \tag{3}$$

The superscript $_{\mathrm{irr}}$ indicates variables under irrigation case. $W_{\mathrm{r}}^{\mathrm{nat}}$ and $W_{\mathrm{r}}^{\mathrm{irr}}$ are return flow due to precipitation and irrigation events, respectively. Then we can estimate C_{irr} by combining the two equations above:

$$\begin{aligned} C_{\mathrm{irr}} &= W_{\mathrm{w}} - W_{\mathrm{r}}^{\mathrm{irr}} \\ &= \mathrm{ET}_P^{\mathrm{irr}} + \mathrm{ET}_I + W_{\mathrm{r}}^{\mathrm{nat}} - P \\ &= \mathrm{ET}_I - \left(\mathrm{ET}_P^{\mathrm{nat}} - \mathrm{ET}_P^{\mathrm{irr}}\right). \end{aligned} \tag{4}$$

If interactive mechanisms between irrigation and local climate are ignored ($\mathrm{ET}_P^{\mathrm{irr}} = \mathrm{ET}_P^{\mathrm{nat}}$), we will have $C_{\mathrm{irr}} = \mathrm{ET}_I$, which is the traditional algorithm (Li et al., 2001).

However, numerous studied demonstrated that irrigation led to surface cooling due to enhanced latent heat flux (Bonfils and Lobell, 2007; Luyssaert et al., 2014; Sacks et al., 2009; Thiery et al., 2017). According to Penman-Monteith equation (Monteith, 1965), surface cooling will result in the decrease of saturation vapor pressure, which consequently reduces evaporative rate. Therefore, although total ET is promoted by irrigation, the ET rooted from P with irrigation is less than that without irrigation due to the surface cooling, suggesting that $\mathrm{ET}_P^{\mathrm{nat}} \neq \mathrm{ET}_P^{\mathrm{irr}}$. Moreover, the $\mathrm{ET}_P^{\mathrm{nat}} - \mathrm{ET}_P^{\mathrm{irr}}$ returns to the river as a part of the return flow $W_{\mathrm{r}}^{\mathrm{irr}}$. In other words, the $W_{\mathrm{r}}^{\mathrm{irr}}$ does not only include irrigation water, but also contains a part of precipitation water, which should be evaporated if irrigation was not applied. For instance, precipitation occurs after an irrigation event. A part of the precipitation should infiltrate into soil for evapotranspiration, but it goes to runoff as soil moisture is much higher after irrigation.

According to our simulations, the irrigation-induced daily maximum temperature decrease is -0.15 K over the YRB (25.91 K in NI and 25.77 K in IR, Fig. R1a), which can be as large as -2 K in some large irrigation districts. Consequently, the potential evapotranspiration decreases can be over 80 mm.yr^{-1} with an average 53.1 mm.yr^{-1} over the YRB (Fig. R1b). This cooling reduces the $\mathrm{ET}_P^{\mathrm{nat}}$ from 364 mm.yr^{-1} to $\mathrm{ET}_P^{\mathrm{act}}$ 348 mm.yr^{-1} with approximately -15.7 mm.yr^{-1} difference, which returns to the river in the form of either surface runoff or deep drainage. The $\mathrm{ET}_P^{\mathrm{nat}}-\mathrm{ET}_P^{\mathrm{irr}}$ is as much as 24.8% of the census-based naturalized streamflow (annual mean is 63.2 mm.yr^{-1} at LiJin station from 1982 to 2014), which, however, was not taken into account during the naturalization process.

In addition, another neglected term is evaporation loss of artificial reservoirs (Li et al., 2001). By using the total surface area of those reservoirs in the YRB (1336 km^2 obtained from the GRanD dataset; `http://globaldamwatch.org/grand/`) and multi-year observed pan evaporation rate (1717 mm.yr^{-1}, Liu et al. (2013)) we preliminarily estimate the evaporative loss of artificial reservoirs is about 2.9 mm.yr^{-1} over the whole catchment, equaling to 4.6% of mean annual naturalized streamflow at LiJin.

Based on the analysis above, we infer that the natural streamflow of the YR is underestimated about 30% due to overlooking the effect of irrigation-induced surface cooling.

3 Uncertainties in the census-based water consumption data

Irrigation consumption accounts for as large as 82% of water consumption in the YRB (YRCC, 1998–2014). Water can be withdrawn from surface water resources or ground water to irrigate croplands and return to river channel or deep drainage. And the difference between water withdrawn and return flows is defined as water consumption which is consumed by evapotranspiration.

The water consumption data is provided by water resources halls of provinces, which also reported effective irrigation area to the National Bureau of Statistics (NBS, from a private conversation with NBS on 16 Sept 2020). Figure R4 shows the time series of effective irrigation area of several provinces (based on the China Statistical Yearbook; http://www.stats.gov.cn/tjsj/ndsj/) that has weird trend in 2008. These weird trends cannot be properly interpreted by technical development due to their steep increase speeds (can be >40%) in the same year, and those provinces are not geographically linked. One possible reason is due to the second national land survey starting from 2007 and ending at 2009, during which overlooked irrigated areas are detected in those provinces. And 2009 is the year writing the China Statistical Yearbook for 2008. However, the water consumption data of these provinces does not have significant trends during this period, which may be due to lack of communication. A solid evidence is that the effective irrigation area in the Statistical Yearbook of Guangdong province (`http://stats.gd.gov.cn/gdtjnj/index.html`) maintains the same magnitude as before 2007. It is also reflected from the Yearbook of Qinghai province (`http://tjj.qinghai.gov.cn/tjData/qhtjnj/`) where the data of effective irrigation area stopped in 2008.

The irrigation water consumption data is based on those irrigation area recognized and monitored. Thus the irrigation water consumption from the census data is the lower limit of real irrigation water consumption. The discussion above implies that a part of irrigation area may be overlooked resulting in underestimation of irrigation water consumption. Although the magnitude of this error is unknown, it will enhance the underestimation of natural streamflow.

4 Water balance in the Yellow River Basin (YRB) might not be closed if constrained by both observed and census data

The water balance of the YRB can be written in either natural or actual ways, as:

$$\Delta S_{\mathrm{n}} = P - Q_{\mathrm{n}} - \mathrm{ET}_{\mathrm{n}}; \tag{5}$$

$$\Delta S_{\mathrm{a}} = P - Q_{\mathrm{a}} - \mathrm{ET}_{\mathrm{a}}. \tag{6}$$

Here P is precipitation; Q is streamflow; ET is evapotranspiration; and ΔS is the change of total water storage (TWS) in the considering period. Subscripts $_{\mathrm{n}}$ and $_{\mathrm{a}}$ denote naturalized and actual variables, respectively. The units of all variables are converted to mm.yr^{-1} for calculation.

According to the Yellow River Conservancy Commission (YRCC), the naturalized streamflow (Q_{n}) is estimated by the observed streamflow (Q_{a}), the surface water consumption (C_{sw}), and the change of water storages of all reservoirs upstream (ΔW_{res}) as:

$$Q_{\mathrm{n}} = Q_{\mathrm{a}} + C_{\mathrm{sw}} + \Delta W_{\mathrm{res}}. \tag{7}$$

The actual evapotranspiration (ET_{a}) is a sum of natural evapotranspiration (ET_{n}), surface water consumption (C_{sw}), and ground water consumption (C_{sw}), as:

$$\mathrm{ET}_{\mathrm{a}} = \mathrm{ET}_{\mathrm{n}} + C_{\mathrm{sw}} + C_{\mathrm{gw}}. \tag{8}$$

Both C_{sw} and C_{gw} are provided by the YRCC. They are defined as the difference between water withdrawal and the return flow back to either river channel or deep drainage.

Let's focus on the period 2003–2010 during which Zhang and Yuan (2020) provided a comprehensive diagnosis of water resources in the YRB by using a well-calibrated VIC model (Variable Infiltration Capacity) against the naturalized streamflow (Yuan et al., 2016) with NSE approximately 0.9. The eight-year averaged annual ET_{n} by VIC is 360 mm.yr^{-1}. According to Zhang and Yuan (2020), although GRACE and soil moisture data are available, there are still a large uncertainties in estimated ET_{a} by applying them in the water balance equation ($\mathrm{ET}_{\mathrm{a}} = 419 \pm 31$ and 401 ± 23 mm.yr^{-1} estimated by GRACE and soil moisture data, respectively). Therefore, by assuming that P, Q_{a}, Q_{n}, ET_{n}, C_{sw}, C_{gw}, and ΔW_{r} are reliable, which are either from in-situ measurements, or from census data, or from a simulation of a well-calibrated model, we try to demonstrate the distribution of water resources in the YRB.

Figure R3 shows the procedure of calculation. Dark and gray variables are what we assume reliable. Blue variables are unknowns and will be calculated by the water balance equation. The Q_n is calculated by YRCC (step ①) via Eq. 7. Based on rainfall-runoff ratio (from P and Q_n), the VIC model was calibrated and used to estimate the ET_n (step ②). Then from Eq. 5, we can estimate the ΔS_n (step ③), which is further used to compute ΔS_a by removing C_gw. Finally, we can get ET_a based on Eq. 8.

Table R2 shows the composition of natural and actual water balances of the YRB. The key point is that $\Delta S_\mathrm{a} = 14\,\mathrm{mm.yr}^{-1}$, which obviously conflict against all versions of GRACE products showing a slightly increasing trend (even if it is de-trended by coal mining data as shown in (Zhang and Yuan, 2020)). This result suggests that at least one of these components that we assumed reliable is not correctly estimated. Firstly, P, Q_a, and ΔW_r are based on direct in-situ observations, which are quite reliable. Secondly, the VIC model is broadly utilized and tested in numerous regional and global studies, which showed the robustness of the model. Thus, the most-likely uncertainty may be from the water consumption terms (C_sw and C_gw).

Thus we try to demonstrate whether the water consumption is overestimated or another way around. By giving a system bias factor λ, we can get the corrected water consumption as:

$$\begin{aligned} \hat{C}_\mathrm{sw} &= \lambda \cdot C_\mathrm{sw}; \\ \hat{C}_\mathrm{gw} &= \lambda \cdot C_\mathrm{gw}. \end{aligned} \tag{9}$$

Variables with hats denote corrected values by introducing the multiplier λ in the water consumption term. Thus the sum of corrected natural ET and ΔS will be:

$$\begin{aligned} \hat{\mathrm{ET}}_\mathrm{n} + \Delta \hat{S}_\mathrm{n} &= P - \hat{Q}_\mathrm{n} \\ &= P - Q_\mathrm{n} + (1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw}) \\ &= \mathrm{ET}_\mathrm{n} + \Delta S_\mathrm{n} + (1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw}) \end{aligned} \tag{10}$$

The $(1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw})$ is the difference between initial and corrected sum of natural ET and ΔS. If we assume that both $\hat{\mathrm{ET}}_\mathrm{n} - \mathrm{ET}_\mathrm{n}$ and $\Delta \hat{S}_\mathrm{n} - \Delta S_\mathrm{n}$ increase/decrease synchronously with positive/negative $(1-\lambda)(C_\mathrm{sw} - C_\mathrm{gw})$, the $\hat{\mathrm{ET}}_\mathrm{n}$ and $\Delta \hat{S}_\mathrm{n}$ can be written as:

$$\begin{aligned} \hat{\mathrm{ET}}_\mathrm{n} &= \mathrm{ET}_\mathrm{n} + \xi(1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw}), \\ \Delta \hat{S}_\mathrm{n} &= \Delta S_\mathrm{n} + (1-\xi)(1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw}). \end{aligned} \tag{11}$$

$\xi \in [0, 1]$ indicates the fraction of $\hat{\mathrm{ET}}_\mathrm{n}$ affected by $(1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw})$. By substituting $\Delta \hat{S}_\mathrm{n} = \Delta \hat{S}_\mathrm{a} + \hat{C}_\mathrm{gw} - \Delta W_\mathrm{res}$ into Eq. 11, we can get:

$$\Delta \hat{S}_\mathrm{a} = \Delta S_\mathrm{n} + (1-\xi)(1-\lambda)(C_\mathrm{sw} + C_\mathrm{gw}) - \lambda C_\mathrm{gw} + \Delta W_\mathrm{res}. \tag{12}$$

If $\pm 2\,\mathrm{mm.yr}^{-1}$ is a reasonable range of TWS trend in the YRB, we can estimate the range of λ by Eq. 12, as:

$$-2\mathrm{mm.yr}^{-1} \leq \Delta S_{\mathrm{n}} + (1-\xi)(1-\lambda)(C_{\mathrm{sw}} + C_{\mathrm{gw}}) - \lambda C_{\mathrm{gw}} \leq 2\mathrm{mm.yr}^{-1}. \tag{13}$$

And finally we can get:

$$\frac{69+46\xi}{57-46\xi} \leq \lambda \leq \frac{73+46\xi}{57-46\xi}. \tag{14}$$

From the inequality, we can conclude that: 1) $\lambda > 1$; 2) λ has the minimum value 1.21 when $\xi = 0$. The result suggests that the real water consumption in the YRB is at least 21% higher than that from the census data, confirming our speculation in Sect. 2 and 3.

Table R1: Comparison between census-based and simulated naturalized monthly streamflow of the Yellow River (1982–2000). R^2 is coefficient of determination. mKGE is modified Kling-Gupta Efficiency. d is degree of agreement.

Stations	R^2	mKGE	d
TangNaiHai	0.64	0.67	0.83
LanZhou	0.63	0.33	0.84
HuaYuanKou	0.52	0.18	0.77
LiJin	0.46	0.27	0.75

Table R2: Eight-year (2003–2010) averaged values of different water balance components based on observations, census data, and simulations. Numbers with normal font are based on in-situ measurements. Bold font indicates from census data. Typewriter font indicates output from a well-calibrated model Zhang and Yuan (2020). Italic font indicates variables estimated by water balance equations in Sect. 4.

Variables	P	Q	ET	ΔS	C_{sw}	C_{gw}	ΔW_{res}
Units	mm.yr^{-1}	mm.yr^{-1}	mm.yr^{-1}	mm.yr^{-1}	mm.yr^{-1}	mm.yr^{-1}	mm.yr^{-1}
Actual	444	24	*406*	*14*	**35**	**11**	**3**
Natural	444	**62**	`360`	*22*	–	–	–

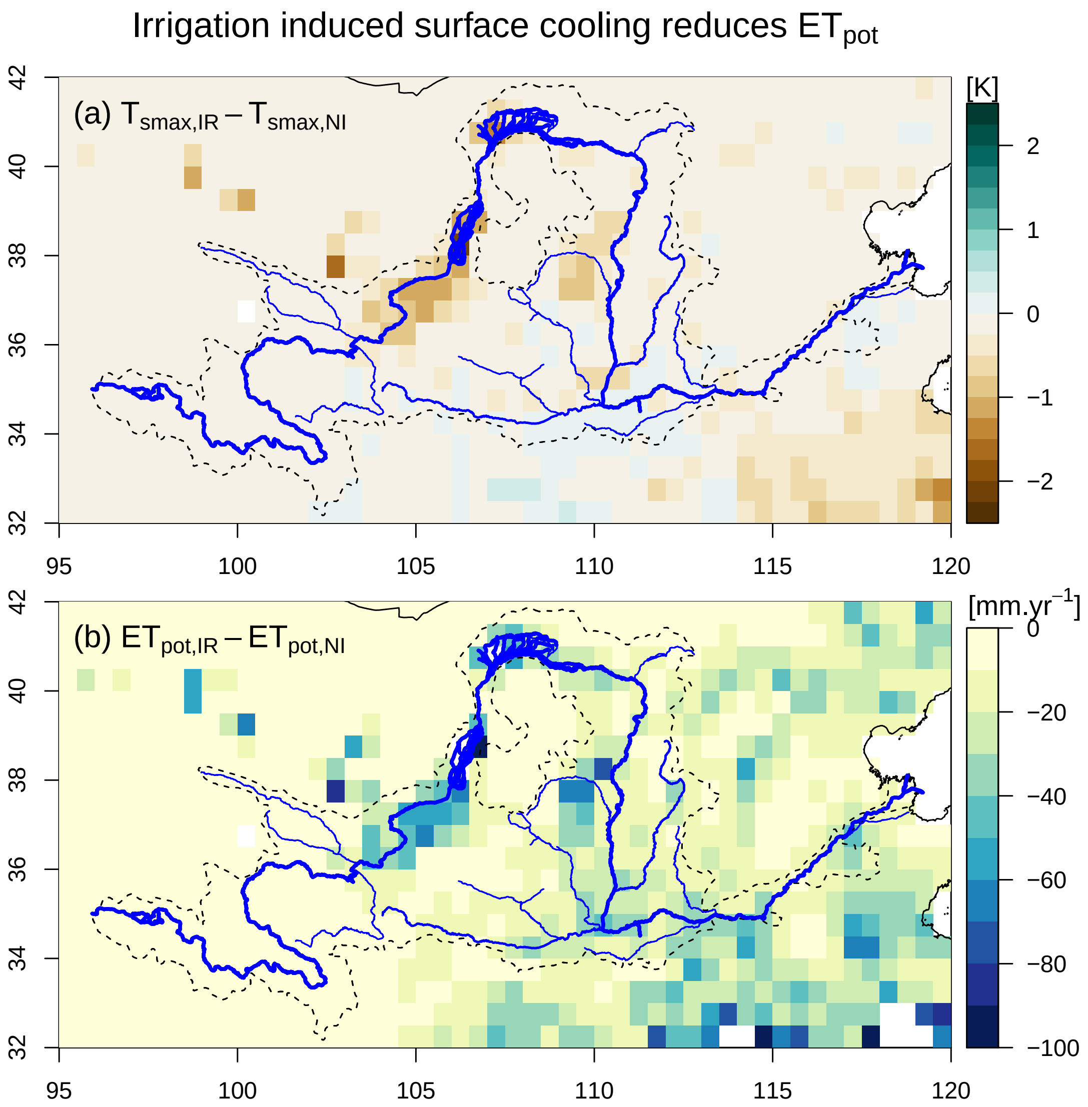

Figure R1: (a) Difference of multi-year mean of daily maximum surface temperature over the simulation period induced by irrigation. NI and IR denote the two simulations with and with out irrigation. Dashed lines denote the Yellow River Basin. Blue lines denote the networks of the Yellow River. (b) Same as (a) but for potential evapotranspiration.

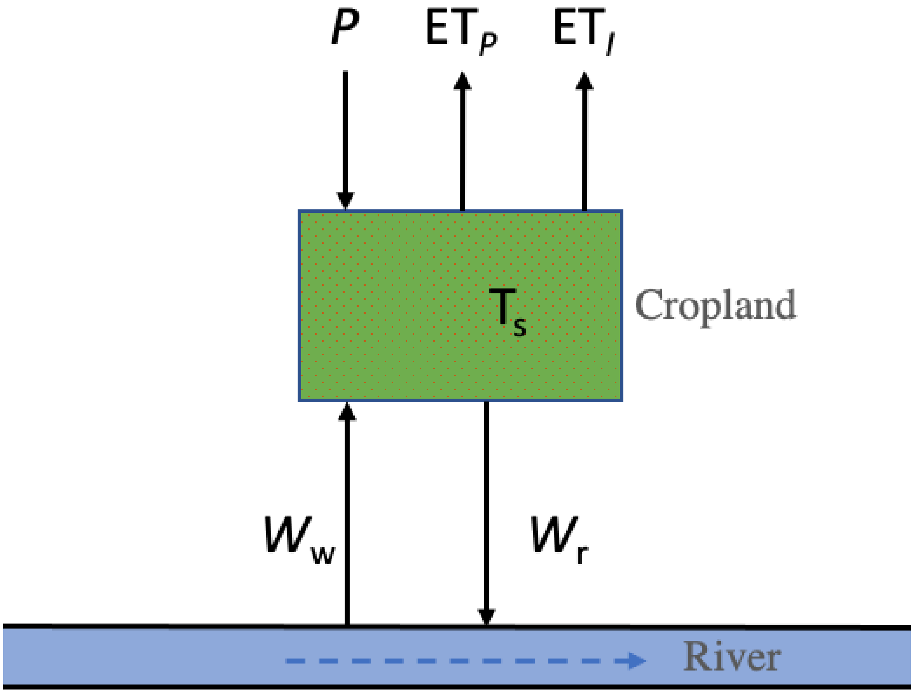

Figure R2: Conceptual figure of a irrigation process. Water is firstly withdrawn (W_w) from a river channel and irrigated in the cropland (water loss during transfer is neglected). After consumed by cropland evapotranspiration, a part of irrigated water will return back to the river channel (W_r). Both W_w and W_r can be measured durin the transport of water to crops, and the deficit is water consumption of irrigation. In the cropland, water is recharged by both precipitation (P) and irrigation (W_w). The irrigation may reduce the surface temperature (T_s) by increasing the heat capacity of soil and promoting evapotranspiration. The evapotranspiration (ET) can be divided into ET using water from P (ET_P) and using water from irrigation (ET_I).

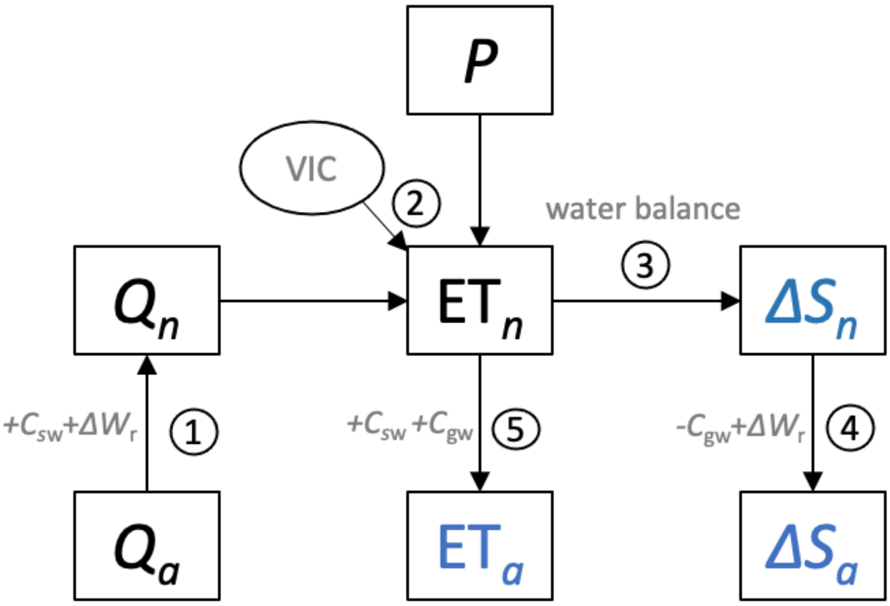

Figure R3: The flowchart of calculating water balances in the Yellow River Basin (YRB). Rectangles underline the components in the water balances. Black denotes observed variables. Blue denotes estimated variables. And gray denotes variables provided by census data. Numbers in circles show the calculation procedure.

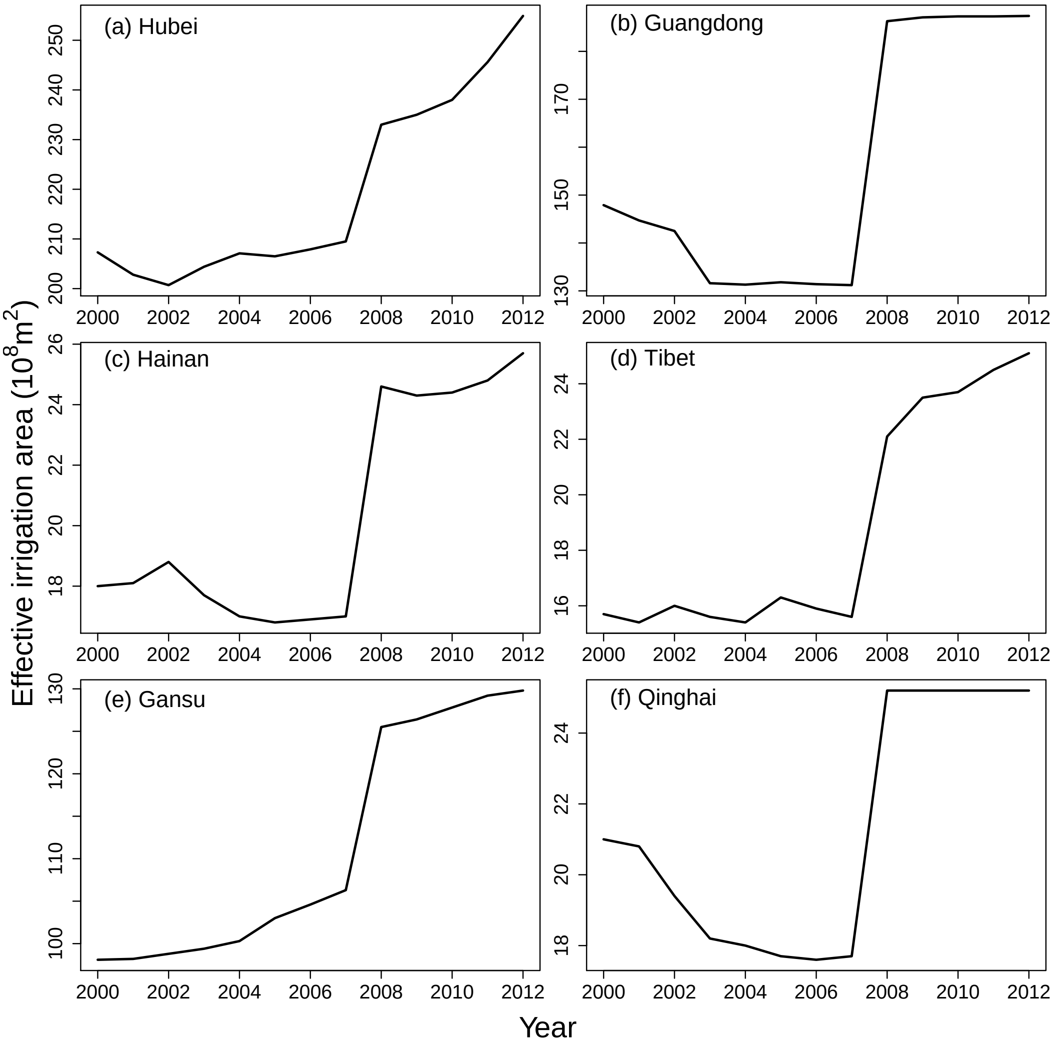

Figure R4: Time series of provincial effective irrigation areas ($10^8\,\mathrm{m}^2$) from 2000 to 2012. The data is from the China Statistical Yearbook (National Bureau of Statistics; `http://www.stats.gov.cn/tjsj./ndsj/`).

Bibliography

Bonfils, C. and Lobell, D.: Empirical evidence for a recent slowdown in irrigation-induced cooling, Proceedings of the National Academy of Sciences, 104, 13 582–13 587, https://doi.org/10.1073/pnas.0700144104, URL `http://www.pnas.org/cgi/doi/10.1073/pnas.0700144104`, 2007.

Duethmann, D., Blöschl, G., and Parajka, J.: Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change?, Hydrology and Earth System Sciences, 24, 3493–3511, https://doi.org/10.5194/hess-24-3493-2020, URL `https://hess.copernicus.org/articles/24/3493/2020/`, 2020.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrology and Earth System Sciences, 22, 789–817, https://doi.org/10.5194/hess-22-789-2018, URL `https://www.hydrol-earth-syst-sci.net/22/789/2018/`, 2018.

Jia, Y. W., Wang, H., Zhou, Z. H., Qiu, Y. Q., Luo, X. Y., Wang, J. H., Yan, D. H., and Qin, D. Y.: Development of the WEP-L distributed hydrological model and dynamic assessment of water resources in the Yellow River basin, Journal of Hydrology, 331, 606–629, https://doi.org/10.1016/j.jhydrol.2006.06.006, URL `http://linkinghub.elsevier.com/retrieve/pii/S0022169406003210`, 2006.

Li, D., Jiang, X. H., Wang, Y. M., and Li, H. L.: Analysis of calculation of natural runoff in the Yellow River Basin, Yellow River, 23, https://doi.org/10.3969/j.issn.1000-1379.2001.02.014, 2001.

Liu, C., Wang, S. R., Liang, Y. Y., and Leung, L. R.: Analysis of Pan Evaporation Change and the Influence Factors in the Yellow River Basin in 1961-2010, Climate Change Research, 9, 327—-334, https://doi.org/10.3969/j.issn.1673-1719.2013.05.003, 2013.

Liu, H., Jia, Y. W., Niu, C. W., Su, H. D., Wang, J. H., Du, J. K., Khaki, M., Hu, P., and Liu, J. J.: Development and validation of a physically-based, national-scale hydrological model in China, Journal of Hydrology, 590, 125 431, https://doi.org/10.1016/j.jhydrol.2020.125431, URL `https://doi.org/10.1016/j.jhydrol.2020.125431https://linkinghub.elsevier.com/retrieve/pii/S002216942030891X`, 2020.

Liu, X. C., Liu, W. F., Yang, H., Tang, Q. H., Flörke, M., Masaki, Y., Müller Schmied, H., Ostberg, S., Pokhrel, Y., Satoh, Y., and Wada, Y.: Multimodel assessments of human and climate impacts on mean annual streamflow in China, Hydrology and Earth System Sciences, 23, 1245–1261, https://doi.org/10.5194/hess-23-1245-2019, URL `https://www.hydrol-earth-syst-sci.net/23/1245/2019/`, 2019.

Luyssaert, S., Jammet, M., Stoy, P. C., Estel, S., Pongratz, J., Ceschia, E., Churkina, G., Don, A., Erb, K., Ferlicoq, M., Gielen, B., Grünwald, T., Houghton, R. A., Klumpp, K., Knohl, A., Kolb, T., Kuemmerle, T., Laurila, T., Lohila, A., Loustau, D., McGrath, M. J., Meyfroidt, P., Moors, E. J., Naudts, K., Novick, K., Otto, J., Pilegaard, K., Pio, C. A., Rambal, S., Rebmann, C., Ryder, J., Suyker, A. E., Varlagin, A., Wattenbach, M., and Dolman, A. J.: Land management and land-cover change have impacts of similar magnitude on surface temperature, Nature Climate Change, 4, 389–393, https://doi.org/10.1038/nclimate2196, URL `http://www.nature.com/articles/nclimate2196`, 2014.

Monteith, J. L.: Evaporation and environment, vol. 19, https://doi.org/10.1613/jair.301, 1965.

Müller, C., Elliott, J., Pugh, T. A. M., Ruane, A. C., Ciais, P., Balkovic, J., Deryng, D., Folberth, C., Izaurralde, R. C., Jones, C. D., Khabarov, N., Lawrence, P., Liu, W. F., Reddy, A. D., Schmid, E., and Wang, X. H.: Global patterns of crop yield stability under additional nutrient and water inputs, PLOS ONE, 13, e0198 748, https://doi.org/10.1371/journal.pone.0198748, URL `https://dx.plos.org/10.1371/journal.pone.0198748`, 2018.

Ottlé, C., Bernus, A., Verbeke, T., Pétrus, K., Yin, Z., Biancamaria, S., Jost, A., Desroches, D., Pottier, C., Perrin, C., de Lavenne, A., Flipo, N., and Rivière, A.: Characterization of SWOT Water Level Errors on Seine Reservoirs and La Bassée Gravel Pits: Impacts on Water Surface Energy Budget Modeling, Remote Sensing, 12, 2911, https://doi.org/10.3390/rs12182911, URL `https://www.mdpi.com/2072-4292/12/18/2911`, 2020.

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S., and Helkowski, J. H.: Effects of global irrigation on the near-surface climate, Climate Dynamics, 33, 159–175, https://doi.org/10.1007/s00382-008-0445-z, URL `http://link.springer.com/10.1007/s00382-008-0445-z`, 2009.

Tang, Q. H., Oki, T., Kanae, S., and Hu, H. P.: Hydrological Cycles Change in the Yellow River Basin during the Last Half of the Twentieth Century, Journal of Climate, 21, 1790–1806, https://doi.org/10.1175/2007JCLI1854.1, URL `http://journals.ametsoc.org/doi/abs/10.1175/2007JCLI1854.1`, 2008.

Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., Hauser, M., and Seneviratne, S. I.: Present-day irrigation mitigates heat extremes, Journal of Geophysical Research: Atmospheres, 122, 1403–1422, https://doi.org/10.1002/2016JD025740, URL `http://doi.wiley.com/10.1002/2016JD025740`, 2017.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth System Dynamics, 5, 15–40, https://doi.org/10.5194/esd-5-15-2014, URL `http://www.earth-syst-dynam.net/5/15/2014/`, 2014.

Wada, Y., de Graaf, I. E. M., and van Beek, L. P. H.: High-resolution modeling of human and climate impacts on global water resources, Journal of Advances in Modeling Earth Systems, 8, 735–763, https://doi.org/10.1002/2015MS000618, URL `http://doi.wiley.com/10.1002/2015MS000618`, 2016.

Wang, X. H.: Impacts of environmental change on rice ecosystems in China: development, optimization and application of ORCHIDEE-CROP model, Ph.D. thesis, Peking University, 2016.

Wang, X. H., Ciais, P., Li, L., Ruget, F., Vuichard, N., Viovy, N., Zhou, F., Chang, J. F., Wu, X. C., Zhao, H. F., and Piao, S. L.: Management outweighs climate change on affecting length of rice growing period for early rice and single rice in China during 1991–2012, Agricultural and Forest Meteorology, 233, 1–11, https://doi.org/10.1016/j.agrformet.2016.10.016, URL `http://linkinghub.elsevier.com/retrieve/pii/S0168192316304087`, 2017.

Xi, Y., Peng, S. S., Ciais, P., Guimberteau, M., Li, Y., Piao, S. L., Wang, X. H., Polcher, J., Yu, J. S., Zhang, X. Z., Zhou, F., Bo, Y., Ottle, C., and Yin, Z.: Contributions of Climate Change, CO 2 , Land-Use Change, and Human Activities to Changes in River Flow across 10 Chinese Basins, Journal of Hydrometeorology, 19, 1899–1914, https://doi.org/10.1175/JHM-D-18-0005.1, URL `http://journals.ametsoc.org/doi/10.1175/JHM-D-18-0005.1`, 2018.

Yang, Q. Q., Huang, X., and Tang, Q. H.: Irrigation cooling effect on land surface temperature across China based on satellite observations, Science of the Total Environment, 705, 135 984, https://doi.org/10.1016/j.scitotenv.2019.135984, URL `https://doi.org/10.1016/j.matdes.2020.108947https://linkinghub.elsevier.com/retrieve/pii/S0048969719359790`, 2020.

Yin, Z., Ottlé, C., Ciais, P., Guimberteau, M., Wang, X. H., Zhu, D., Maignan, F., Peng, S. S., Piao, S. L., Polcher, J., Zhou, F., and Kim, H.: Evaluation of ORCHIDEE-MICT-simulated soil moisture over China and impacts of different atmospheric forcing data, Hydrology and Earth System Sciences, 22, 5463–5484, https://doi.org/10.5194/hess-22-5463-2018, URL `https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-699/https://www.hydrol-earth-syst-sci.net/22/5463/2018/`, 2018.

Yin, Z., Wang, X. H., Ottlé, C., Zhou, F., Guimberteau, M., Polcher, J., Peng, S. S., Piao, S. L., Li, L., Bo, Y., Chen, X. L., Zhou, X. D., Kim, H., and Ciais, P.: Improvement of the Irrigation Scheme in the ORCHIDEE Land Surface Model and Impacts of Irrigation on Regional Water Budgets Over China, Journal of Advances in Modeling Earth Systems, 12, 1–20, https://doi.org/10.1029/2019MS001770, URL `https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001770`, 2020.

YRCC: Yellow River Water Resources Bulletin 1998–2014, Tech. rep., Yellow River Conservancy Commission, Zhengzhou, URL `http://www.yrcc.gov.cn/other/hhgb/`, 1998–2014.

Yuan, X., Ma, F., Wang, L. Y., Zheng, Z. Y., Ma, Z. G., Ye, A. Z., and Peng, S. M.: An experimental seasonal hydrological forecasting system over the Yellow River basin – Part 1: Understanding the role of initial hydrological conditions, Hydrology and Earth System Sciences, 20, 2437–2451, https://doi.org/10.5194/hess-20-2437-2016, 2016.

Yuan, X., Jiao, Y., Yang, D., and Lei, H.: Reconciling the Attribution of Changes in Streamflow Extremes From a Hydroclimate Perspective, Water Resources Research, 54, 3886–3895, https://doi.org/10.1029/2018WR022714, 2018.

Zhang, M. and Yuan, X.: Crucial role of natural processes in detecting human influence on evapotranspiration by multisource data analysis, Journal of Hydrology, 580, 124 350, https://doi.org/10.1016/j.jhydrol.2019.124350, URL `https://doi.org/10.1016/j.jhydrol.2019.124350https://linkinghub.elsevier.com/retrieve/pii/S0022169419310856`, 2020.