

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2020WR027413

### Special Section:

Advancing Process Representation in Hydrologic Models: Integrating New Concepts, Knowledge, and Data

#### Key Points:

- Net irrigation can be quantified by evapotranspiration residuals derived from hydrologic models and satellite remote sensing
- Estimation of winter wheat irrigation has lower uncertainty than estimation of irrigation of summer crops
- Irrigation water use efficiency has improved in the North China Plain in the period 2002–2016

#### Supporting Information:

- Supporting Information S1

#### Correspondence to:

J. Koch,  
juko@geus.dk

#### Citation:

Koch, J., Zhang, W., Martensen, G., He, X., & Stisen, S. (2020). Estimating net irrigation across the North China Plain through dual modeling of evapotranspiration. *Water Resources Research*, 56, e2020WR027413. <https://doi.org/10.1029/2020WR027413>

Received 27 FEB 2020

Accepted 2 NOV 2020

Accepted article online 9 NOV 2020

## Estimating Net Irrigation Across the North China Plain Through Dual Modeling of Evapotranspiration

Julian Koch<sup>1</sup> , Wenmin Zhang<sup>2</sup>, Grith Martensen<sup>1</sup> , Xin He<sup>1,3</sup> , and Simon Stisen<sup>1</sup> 

<sup>1</sup>Department of Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark, <sup>2</sup>School of Geography, Nanjing Normal University, Nanjing, China, <sup>3</sup>Department of Water Resources, China Institute of Water Resources and Hydropower Research, Beijing, China

**Abstract** Irrigation is the greatest human interference with the terrestrial water cycle. Detailed knowledge on irrigation is required to better manage water resources and to increase water use efficiency (WUE). This study applies a framework to quantify net irrigation at monthly timescale at a spatial resolution of  $1 \text{ km}^2$  providing high spatial and temporal detail for regional water resources management. The study is conducted in the Haihe River Basin (HRB) in China encompassing the North China Plain (NCP), a global hot spot of groundwater depletion. Net irrigation is estimated based on the systematic evapotranspiration (ET) residuals between a remote sensing-based model and a hydrologic model that does not include an irrigation scheme. The results suggest an average annual net irrigation of  $126 \text{ mm yr}^{-1}$  ( $15.2 \text{ km}^3 \text{ yr}^{-1}$ ) for NCP and  $108 \text{ mm yr}^{-1}$  ( $18.6 \text{ km}^3 \text{ yr}^{-1}$ ) for HRB. It is found that net irrigation can be estimated with higher fidelity for winter crops than for summer crops. The simulated water balance for NCP is evaluated with Gravity Recovery and Climate Experiment (GRACE) data, and the net irrigation estimates can close the water balance gap. Annual winter wheat classifications reveal an increasing crop area with a trend of  $2,200 \text{ km}^2 \text{ yr}^{-1}$ . This trend is not accompanied by a likewise increasing trend in irrigation water use, which suggests an increased WUE in the NCP, which is further supported by net primary productivity data. The proposed framework has potential to be transferred to other regions and support decision makers to support sustainable water management.

**Plain Language Summary** The irrigation of agricultural fields is taking place at unsustainable rates in many regions of the world. Despite the fact that irrigation is the largest anthropogenic impact on the natural water cycle, there exists limited knowledge of the applied irrigation amounts. This study applies an approach to estimate net irrigation, that is, the evaporative loss of irrigation water. The approach is applied on the Haihe River Basin in North-Eastern China. For the estimation of net irrigation, two sources of evapotranspiration (ET) are considered. First, baseline ET is obtained from a rainfed hydrologic model without irrigation. Second, ET is obtained from a satellite remote sensing model, which represents rainfed and irrigated ET. We study the ET differences of the two sources to derive net irrigation amounts at monthly timescale at  $1 \text{ km}^2$  spatial resolution. Our analysis suggest an average annual net irrigation of  $108 \text{ mm yr}^{-1}$  ( $18.6 \text{ km}^3 \text{ yr}^{-1}$ ). The results are evaluated against annual winter wheat classification maps as well as satellite-based net primary productivity data and total water storage data (Gravity Recovery and Climate Experiment, GRACE). Our results indicate an increasing water use efficiency as a result of promoting water savings in the agricultural sector.

## 1. Introduction

It is estimated that 70% of the global freshwater withdrawals are attributed to irrigation, which makes agriculture the principal freshwater consumer (Foley et al., 2011; Siebert et al., 2010). Irrigated lands produces 40% of the global food on just 20% of the total agricultural lands (Vörösmarty & Sahagian, 2000). The steady population growth in combination with climate change will further increase the demand for irrigation agriculture (Rockström et al., 2012). Already today, over 40% of the applied irrigation originates from groundwater abstractions resulting in prolonged periods of persistent groundwater depletion (Famiglietti et al., 2011; Siebert et al., 2010). The irrigation-induced overexploitation of groundwater resources is likely to exacerbate in the coming decades, which will increase the need for quantification and mapping of irrigation in order to facilitate critical information for policy makers and water resources managers (Schwartz et al., 2020).

Even though irrigation is the most important direct human interference with the terrestrial water cycle and irrigation has a distinct role as climate forcing (Cook et al., 2015; Kang & Eltahir, 2019), there exists limited knowledge on the extent of irrigated areas and in particular on the amount of water applied for irrigation. Traditionally, irrigated areas and requirements have been documented and mapped based on census-based national agricultural maps and surveys in combination with crop water models. For instance, Siebert et al. (2010, 2015) have worked on inventories of irrigation extents at global scale. With the rise of modern satellite remote sensing systems, mapping the extent of irrigation has been an active field of research since the early 2000s. For example, Ozdogan and Gutman (2008) have mapped irrigated areas across the continental United States using remotely sensed data on vegetation phenology and climate. Similar work has been carried out for China by Zhu et al. (2014), for northern India by Thenkabail et al. (2005), and at global scale by Thenkabail et al. (2009). Recently, the focus has moved to high-resolution mapping of irrigated areas using data from the Landsat or Sentinel satellite missions (Bazzi et al., 2019; Deines et al., 2019; Xiang et al., 2019). Despite the advances in mapping historic irrigation extents, few methodologies exist to estimate continuous irrigation amounts. In recent years, the literature on this topic has been growing quickly and the common ground of the published studies on irrigation quantification is that they rely on satellite remote sensing data. Retrievals of soil moisture (SM) or evapotranspiration (ET) are either used in stand-alone remote sensing approaches with auxiliary climate data or in conjunction with hydrologic models that either have an internal irrigation scheme or not.

Approaches to model irrigation dynamically in hydrologic models follow the assumption to balance available water supply with plant and atmospheric water demand, which is often based on simplified deficit rules applying predefined thresholds (Ozdogan et al., 2010). This framework is associated with large uncertainties due to the difficulties to correctly estimate plant water demand and to predict management decisions and due to challenges related to the land cover maps that identify irrigated croplands (Lawston et al., 2015; Wisser et al., 2008).

From the SM perspective, Brocca et al. (2018) used remotely sensed SM to invert the soil water balance equation to calculate irrigation at monthly timescale. Other recent SM-based studies aiming at quantifying irrigation amounts were conducted by Zaussinger et al. (2019), Zohaib and Choi (2020), and Kumar et al. (2015), all accounted irrigation to differences between remotely sensed SM and SM modeled by hydrologic models without irrigation schemes. Other recent works suggest to estimate irrigation through data assimilation of satellite-based SM in hydrologic models (Abolafia-Rosenzweig et al., 2019; Felfelani et al., 2018). Jalilvand et al. (2019) found that the low spatial resolution of global SM products ( $\sim 50 \text{ km}^2$ ) hindered to derive irrigation amounts at relevant spatial scales for regional analysis. Further, limitations of SM were highlighted by Escorihuela and Quintana-Seguí (2016) who compared various global SM satellite products in the context of irrigation quantification and conclude that the downscaled Soil Moisture and Ocean Salinity (SMOS) product was the only one able to detect an irrigation signal. Moreover, in order to convert volumetric SM into a corresponding water column depth, various assumptions such as depth of soil, water capacity of the soil layer, and other empirical parameters are necessary which introduce additional uncertainties. Even though Sentinel-1 provides SM at higher resolution, its application for irrigation quantification, especially at regional scale for water resources management, has not been reported in the literature yet.

In the literature, studies deriving irrigation quantities based on ET have not emerged at the same fast pace as this is currently the case for SM-based applications. The notion to infer regions where nonprecipitation sources, such as irrigation, significantly affect ET fluxes, by comparing prognostic hydrologic models without irrigation schemes with diagnostic remote sensing retrievals, has been applied in just a few studies. Hain et al. (2015) applied this framework to locate nonprecipitation sources, such as irrigation, and sinks, such as drainage in the United States. For the first case, the satellite-based ET retrievals show a systematic positive bias when being compared to hydrologic models that do not explicitly account for irrigation. Opposed, a systematic negative bias can be accounted to drainage. The same approach has been applied by Romaguera et al. (2012), Romaguera, Krol, et al. (2014), and Romaguera, Salama, et al. (2014) at European scale as well as for other study sites in East Africa and China. These studies highlighted the need to correct the irrigation amounts with the ET bias over rainfed agriculture. The purpose for this correction is comprehensible, but it remains disputable if the bias can be assumed constant in space. Van Dijk et al. (2018) presented an alternative ET-based approach to assess irrigation at globe scale using a hydrologic model without an irrigation scheme. Satellite-based land surface temperature (LST) was assimilated, and it was assumed that any

increase in ET was due to irrigation. It is questionable if the baseline model represents plausible rainfed conditions and if assimilating LST realistically affects ET.

The study site is the Haihe River Basin (HRB) in China. The HRB comprises the North China Plain (NCP), which is a global hot spot of prolonged groundwater depletion induced by irrigation agriculture (Taylor et al., 2013; Thenkabail et al., 2009). It is indisputable that irrigation is the main driver of the observed decline in groundwater heads, as agriculture consumes 70% of the total water use in the HRB of which 70% is sustained by groundwater abstraction (Huang et al., 2015; Pan et al., 2017; Shen et al., 2015). The water crisis of the NCP has been intensively studied, and the emerging environmental and societal risks have been clearly identified (Cao et al., 2013; Huanhuan Qin et al., 2019; Zheng et al., 2010). Despite the eminence and the awareness that irrigation agriculture is the principal driver of the groundwater depletion, there exist little knowledge on historic irrigation amounts at spatiotemporal scales that are required to tackle the water crisis.

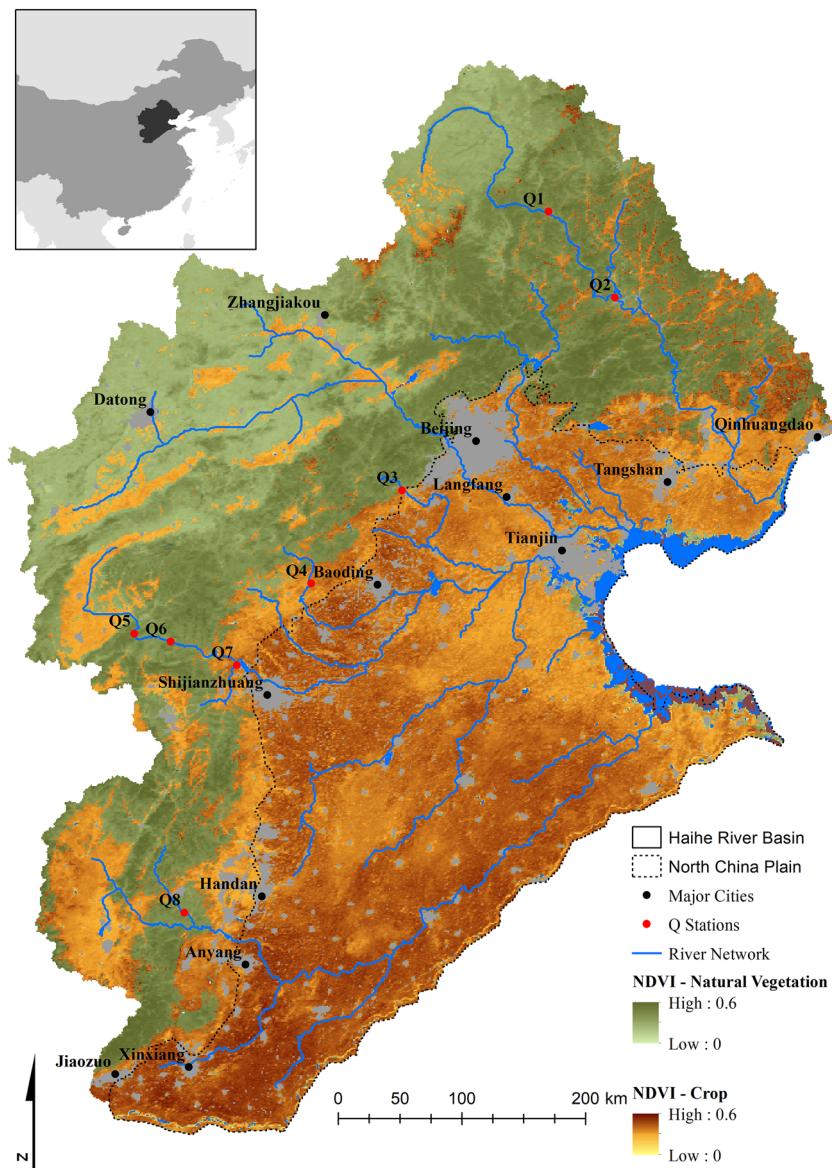
In this study, we present a methodology that is designed to quantify monthly net irrigation amounts that account for the evaporative loss of irrigation water at  $1 \text{ km}^2$  spatial resolution. We build upon the existing body of literature relating irrigation to systematic ET residuals. We target a spatial resolution of  $1 \text{ km}^2$ , which we regard as a relevant spatial resolution for water resources mapping and management at regional scale ( $>100,000 \text{ km}^2$ ), such as the NCP. For irrigation design and optimization at farm scale, other remote sensing sources, such as Landsat or Sentinel, are more appropriate. The ET-based approach is favored over the alternative SM-based approach, because it has the advantage of providing a direct estimate of water loss due to irrigation (millimeter) at a spatial scale that is relevant for regional water management. A key novelty of this study is that the applied hydrologic model, which is used as rainfed ET baseline without irrigation, is specifically calibrated to perform well for rainfed conditions, thus representing natural conditions. Systematic errors between ET derived from the hydrologic model and the remote sensing data have to be accounted for. In our opinion, calibrating the hydrologic model to account for any systematic bias is a more scientifically sound approach in comparison to simply assessing the average bias over rainfed areas to correct the hydrologic model (Romaguera et al., 2012). Furthermore, we present a method to estimate the predictive uncertainty of the irrigation quantifications, by investigating the variance of ET residuals over rainfed cropland. These two points mark a clear novelty to the field of irrigation quantification. Ground truth observations of irrigation amounts are rare, especially at larger scales, which hinders a direct evaluation. However, our study brings forward alternative satellite-based data sources that indirectly relate to water consumption and crop yield in the evaluation of the irrigation estimates, such as total water storage anomalies, land cover maps, and net primary productivity (NPP).

The four main objectives of this paper are as follows: (1) to set up a remote sensing-based ET model for the HRB, (2) to set up a hydrologic model without irrigation scheme and to specifically calibrate it for rainfed conditions, (3) to quantify monthly net irrigation amounts as well as predictive uncertainty at  $1 \text{ km}^2$  spatial resolution for a 15 year period, and (4) to evaluate the derived net irrigation amounts against satellite-based total water storage data, land use maps, and water use efficiency (WUE) data.

## 2. Study Area and Data

### 2.1. HRB

The HRB covers an area of approximately  $320,000 \text{ km}^2$  and encompasses mountainous regions in the west and north and lowlands in the east and south. The lowlands refer to the NCP, which covers approximately  $140,000 \text{ km}^2$  of the HRB (Figure 1). The western boundary of NCP are the Taihang Mountains and the Bohai Sea in the East. The NCP is home to over 135 million people, including the megacities Beijing and Tianjin, and produces around 30% of China's wheat and 20% of its maize (Guo & Shen, 2015; Pan et al., 2017; Qin et al., 2019). The HRB is dominated by a monsoon climate with an average annual rainfall of around  $475 \text{ mm yr}^{-1}$  (2002–2016) of which 70% to 85% occurs during the summer months (June–September). Agriculture is the major land use in the NCP, covering over 80% of the land, and the cropland is cultivated with a rotation system consisting of winter wheat (October–June) and summer maize (June–October). The summer maize growing season coincides with the rainy season (June–September with monthly precipitation larger than  $50 \text{ mm mo}^{-1}$ ), and water requirements are therefore to a large degree met by rainfall. In contrast, the winter wheat growing season spans over the dry season (October–May with monthly precipitation lower than  $50 \text{ mm mo}^{-1}$ ) and crop water requirements depend heavily on irrigation. The flood irrigation technique



**Figure 1.** Map of the Haihe River Basin (HRB) containing the NCP domain. The depicted river network represents the natural drainage system. The map of mean annual NDVI is differentiated into cropland and natural vegetation based on a MODIS land cover classification. Based on this classification, urban areas are shown in gray, waterbodies in blue, and barren soil in brown. The name of the eight discharge stations corresponds to the IDs in Table 1. The top left panel indicates the HRB in dark gray and China in medium gray.

is widely applied in NCP and typically takes place at few occasions during the winter crop season (Qin et al., 2013). The spatial variability of the average annual normalized difference vegetation index (NDVI) over cropland in Figure 1 indicates that the two-stage crop rotation system and thereby irrigation is not applied uniformly. Higher mean NDVI values can be expected for grids with two crop cycles in comparison to grids with just a single crop cycle, which explain the spatial variability. Following Shen et al. (2015), at least 70% of the total irrigation is sustained by groundwater abstractions which puts winter wheat cultivation at the center of the NCP water crisis.

## 2.2. MODIS

For the present study, a broad range of satellite remote sensing-based data sets were acquired. A prime data source were the MODIS (Moderate Resolution Imaging Spectroradiometer) instruments onboard Terra and Aqua satellites. The list of variables includes NDVI, leaf area index (LAI), fraction of absorbed

photosynthetically active radiation (FAPAR), LST, emissivity, albedo, NPP, and land cover. Details are given in Table S1 (ST1) in the supporting information to this article. The approximate nighttime and daytime overpass times for Terra and Aqua are 11 p.m. and 11 a.m. and 1 a.m. and 1 p.m., respectively. Missing nighttime LST observations have been filled using linear interpolation. If not already available at 1 km, all variables were resampled to 1 km, based on the average, for further analysis. MODIS quality flags were used to only extract high-quality observation. In order to get robust timeseries and thereby deal with missing data, we first calculated the average annual climatology for each grid for the MODIS data sets based on data from 2002 to 2016. This processing step was applied to NDVI, LAI, FAPAR, albedo, and emissivity. In the following, the relative deviation between the actual observations and the coinciding climatology was calculated by division of the first with the latter. Subsequently, the deviation was interpolated in time for the missing observations using linear interpolation. A Gaussian filter was applied to the interpolated deviations with the purpose to smooth the timeseries. Lastly, the smoothed timeseries represented the relative deviation of a given year to the climatology and could simply be multiplied with the climatology to obtain a full timeseries for a given year. With this processing of the MODIS data, we obtained robust and complete timeseries. The climatology was used as reference, but multiplying it with the smoothed deviations allowed to differentiate between the years and thereby adjusting the climatology respectively. Daytime LST was excluded from this analysis, because cloud-free LST cannot be used to interpolate for days with cloud cover.

### 2.3. ERA-Interim

ERA-Interim is a global reanalysis data set of atmospheric and land surface variables provided by the European Center for Medium-Range Weather Forecasts (ECMWF). Data is available at 3-hourly temporal resolution at a spatial resolution of  $0.75^{\circ}$ . We acquired daily shortwave downward radiation and daily mean, minimum, and maximum air temperature data. Furthermore, daytime LST data were processed to gap fill daily MODIS LST. For this purpose, the ERA-Interim LST data was resampled to  $1 \text{ km}^2$  using a spline interpolation.

### 2.4. Grace

Monthly total water storage anomalies (TWSA) from the Gravity Recovery and Climate Experiment (GRACE) satellite mission were acquired from level-2 release 05 spherical harmonics from the Center for Space Research (CSR), Jet Propulsion Laboratory (JPL), and GeoForschungsZentrum (GFZ) solutions. GRACE data are available at  $1^{\circ}$  spatial resolution since April 2002. The monthly TWSA is relative to the baseline average over January 2004 to December 2009. As suggested by Landerer and Swenson (2012), scaling coefficients were applied to the TWSA data to account for attenuated small scale mass variations in the  $1^{\circ}$  processing. The coefficients are multiplicative factors for each grid cell that minimize the difference between GRACE's smoothed and unfiltered monthly water storage variations. For this study, GRACE data were averaged in space for the entire NCP ( $29 1^{\circ}$  grids) as well as for the three solutions (CSR, JPL, and GFZ) to a single timeseries. In order to visualize uncertainties related to the GRACE data, three sources of uncertainty were combined, namely, the variability between the three solutions, the time constant GRACE measurement errors, and the associated leakage errors.

### 2.5. Discharge

Discharge measurements were available at eight stations across the HRB. These stations were selected due to their relatively undisturbed flow conditions. The discharge data was previously employed by Davidsen et al. (2015) and Martinsen et al. (2019) to optimize hydrologic models (rainfall-runoff models). The hydrologic model applied in this study only simulates natural flow conditions, and therefore, catchments that are excessively managed, by, that is, diversions and reservoirs, cannot be used for evaluation purposes. As shown by Figure 1, the eight stations are located upstream the NCP, in the areas where anthropogenic influences are less dominant. The size of the upstream area of the selected stations varied between 2,000 and 18,000  $\text{km}^2$  (Table 1). The data coverage varied significantly among the stations where two had full timeseries of 9 years with daily observations while others had only a few hundred measurements spread out over several years.

### 2.6. Annual Winter Wheat Classification

Annual winter wheat maps between 2002 and 2016 were derived by analyzing all available Landsat-5, Landsat-7, and Landsat-8 surface reflectance Tier 1 data between 1 October and 30 June in the following

**Table 1**

Overview of the Eight Discharge Stations That Were Applied in the Model Calibration

ID	Name	Area (km <sup>2</sup> )	Longitude/latitude	Observation period	Data coverage
Q1	Goutaizi (古太子)	2,050	117.03 E/41.35 N	Jan/2006–Dec/2014	100% ( <i>n</i> = 3,287)
Q2	Sandaohanzi (三道河子)	18,234	117.7 E/40.97 N	Jan/2006–Dec/2016	17.2% ( <i>n</i> = 691)
Q3	Zhangfang (张坊)	3,041	115.68 E/39.57 N	Jan/2006–Dec/2014	6.3% ( <i>n</i> = 206)
Q4	Zhongtangmei (中唐梅)	3,562	114.88 E/38.88 N	Jan/2006–Dec/2014	13.9% ( <i>n</i> = 455)
Q5	Jishengqiao (济胜桥)	11,874	113.06 E/38.38 N	Jan/2006–Dec/2010	100% ( <i>n</i> = 1826)
Q6	Xiao Jue (小觉)	14,051	113.43 E/38.23 N	Jan/2006–Dec/2010	100% ( <i>n</i> = 1826)
Q7	Pingshan (平山)	6,268	114.12 E/38.15 N	Jan/2006–Dec/2010	99.9% ( <i>n</i> = 1824)
Q8	Kuangmenkou (狂门口)	4,932	113.47 E/38.15 N	Jan/2006–Dec/2014	100% ( <i>n</i> = 3,287)

Note. The IDs correspond to the ones in the map (Figure 1). The observation period refers to the years were data is available, and the data coverage is with respect to the stated period. *n* is the number of daily discharge observations.

year. The data were processed in Google Earth Engine (GEE) to monthly enhanced vegetation index (EVI) maps. The training and validation data were collected from multiple reference sources that were comprised of GEE, Sentinel-2, and Landsat, which were used in combination with the monthly EVI time series in a Random Forests (RF) Classifier. The winter wheat labeled data used to train the RF model were extracted manually from the sources above by drawing polygons over known winter wheat fields. Then, points were placed randomly within the polygons. The labeled data also included other land cover types that were sampled based on the same approach as winter wheat. The labeled data was split randomly into equal training and validation subsets and applied to train a RF classifier. The overall accuracy of the binary maps (winter wheat and nonwinter wheat) varied slightly across the years with an average of around 96%.

### 3. Methods

#### 3.1. Remote Sensing ET Model

In this study, we apply the Priestly-Taylor JPL (PT-JPL) model to estimate daily actual ET (Fisher et al., 2008). In particular, we use the PT-JPL thermal model, developed by García et al. (2013) who extended the traditional PT-JPL model to incorporate LST as a proxy for the SM control on ET. PT-JPL initially estimates potential ET (PET) for soil ( $PET_s$ ) and canopy ( $PET_c$ ) based on the approach by Priestley and Taylor (1972) where the net radiation is split between soil and canopy based on the LAI (Norman et al., 1995).

$$R_{n_s} = R_n \cdot e^{(-k_{R_n} \cdot LAI)} \quad (1)$$

$$R_{n_c} = R_n - R_{n_s} \quad (2)$$

$$PET = \alpha_{PT} \cdot \frac{\Delta}{\Delta + \gamma} \cdot (R_n - G), \quad (3)$$

where  $R_n$  is the net radiation ( $\text{Wm}^{-2}$ ),  $R_{n_s}$  is the net soil radiation ( $\text{Wm}^{-2}$ ),  $R_{n_c}$  is the net canopy radiation ( $\text{Wm}^{-2}$ ),  $k_{R_n} = 0.6$  (Norman et al., 1995),  $\alpha_{PT} = 1.26$  is the P-T coefficient,  $\Delta$  is the slope of the saturation to vapor pressure curve ( $\text{Pa K}^{-1}$ ),  $\gamma = 0.066 \text{ kPa C}^{-1}$  is the psychrometric constant, and  $G$  is the soil heat flux that is considered negligible at daily scale ( $\text{Wm}^{-2}$ ).

$PET_s$  is calculated by using  $R_{n_s}$  in Equation 3, and correspondingly,  $PET_c$  uses  $R_{n_c}$  as energy input. Subsequently, the potential levels are reduced to their actual levels using various constraints. The constraints reflect the plant physiological status and SM availability and act as multipliers that can vary between 0 and 1. Finally, total actual ET is expressed as the sum of actual canopy transpiration ( $ET_c$ ) and actual soil evaporation ( $ET_s$ ):

$$ET = ET_c + ET_s. \quad (4)$$

Canopy transpiration is calculated based on three physiological constraints:

**Table 2**  
*Equations Used to Calculate the Biophysical Constraints for the PT-JPL Model*

Constraint	Description	Equation	Reference
$f_g$	Green canopy fraction	$f_{APAR}/f_{IPAR}$	Fisher et al. (2008)
$f_T$	Plant temperature constraint	$1.1814 \cdot \left[ 1 + e^{(T_{opt} - 10 - Ta_m)} \right]^{-1}$	Potter et al. (1993)
$f_M$	Plant moisture constraint	$f_{APAR}/f_{APARmax}$	Fisher et al. (2008)
$f_{SM}$	Soil moisture constraint	$\frac{ATI - ATI_{min}}{ATI_{max} - ATI_{min}}$	Verstraeten et al. (2006)

Note.  $f_{APAR}$  is the fraction of absorbed photosynthetically active radiation,  $f_{IPAR}$  is the fraction of intercepted photosynthetically active radiation, calculated by the NDVI relationship proposed by Myneni and Williams (1994),  $T_{opt}$  is the optimum temperature for plant growth (25°C),  $Ta_m$  is the daily mean air temperature (°C),  $f_{APARmax}$  is the maximum  $f_{APAR}$ , which was set to the 95<sup>th</sup> percentile in this study, and ATI is the apparent thermal inertia index and, in this study, ATI<sub>max</sub> and ATI<sub>min</sub> related to the 95th and 5th percentiles, respectively.

$$ET_c = f_g \cdot f_T \cdot f_M \cdot PET_c, \quad (5)$$

where  $f_g$  represents the green canopy fraction,  $f_T$  is the plant temperature constraint, and  $f_M$  captures the plant moisture constraint.

Soil evaporation is calculated by considering a single SM constraint ( $f_{SM}$ ):

$$ET_s = f_{SM} \cdot PET_s. \quad (6)$$

Equations for the four applied biophysical constraints are stated in Table 2, and more details of the PT-JPL thermal model can be found in García et al. (2013) and Moyano et al. (2018). LST and albedo are used to calculate the apparent thermal inertia (ATI) term used in  $f_{SM}$ . ATI requires a nighttime and a daytime LST observation. We calculated PT-JPL with ERA-Interim and MODIS LST. On days when MODIS did not provide a clear sky LST observation, ERA-Interim was used for gap filling.

### 3.2. Hydrologic Model

This study applies version 5.9 of the multiscale Hydrologic Model (mHM, Samaniego et al., 2019). mHM is a grid based spatially distributed hydrologic model that accounts for key hydrologic processes and includes a routing scheme (Kumar et al., 2013, Samaniego et al., 2010, Thober et al., 2019). A multiparameter regionalization technique enables mHM to consolidate three different spatial scales: meteorological forcing at coarse scale, an intermediate model scale, and morphological data at a fine scale. In case of the HRB basin, forcing data is resampled to 4 km resolution based on the mean function, the morphological data, such as the digital elevation model (DEM) or soil texture is used at 500 m resolution and the model is executed at 4 km resolution during calibration and at 1 km for a final production run after calibration. Effective parameters at the modeling scale are regionalized through nonlinear transfer functions which link spatially distributed basin characteristics at finer scale by means of global parameters which can be determined through calibration. This regionalization framework has the advantage of providing seamless parameter fields to mHM (Samaniego et al., 2017). Following the work presented by Demirel et al. (2018), a dynamic scaling enables mHM to downscale potential ET to the model resolution using the monthly climatology of LAI, in similar fashion to the well-known crop coefficient. In order to apply mHM to simulate baseline ET conditions that are purely driven by precipitation, the MODIS-based LAI data had to be corrected to remove the effect of irrigation. This was achieved by reducing LAI of irrigated cropland, in the months from October to May, to the average LAI of rainfed cropland of that particular month. mHM simulates actual ET as a fraction of potential ET depending on the SM conditions. The applied Feddes reduction factor is a function of SM, field capacity, and wilting point (Feddes et al., 1976). For this study, MSWEP v2 was used as precipitation forcing (Beck et al., 2017, 2019), daily mean air temperature was acquired from ERA-Interim, and potential ET was used from the PT-JPL model. Soil texture data was obtained from the Harmonized Soil World database (FAO & IIASA, 2009) which contains around 300 soil classes in the HRB. The DEM was obtained from the NASA's Shuttle Radar Topography Mission (Farr et al., 2007). The MODIS MCD15A2H.006 LAI product was utilized to derive the monthly climatology maps.

### 3.3. Model Calibration

#### 3.3.1. Calibration Strategy

A multiobjective and multivariable calibration framework has been designed to yield robust model performance of the hydrologic model to simulate reliable baseline rainfed conditions. More precisely, baseline rainfed conditions refer to the natural hydrological cycle where precipitation is the sole source of water, that is, irrigation is not present. Observations of the two most important water balance outflows, namely, discharge ( $Q$ ) and ET, were used in the calibration of the hydrologic model of the HRB. A global optimizer scheme within PEST (Doherty, 2005) that is based on a covariance matrix adaptation estimation strategy (CMA-ES; Hansen & Ostermeier, 2001) was applied to calibrate mHM parameters. CMA-ES utilizes a stochastic sampling approach based on updated probability fields and has proven very efficient in hydrologic model calibrations (Arsenault et al., 2014; Stisen et al., 2018).

#### 3.3.2. Objective Functions

Two objective functions were applied in the calibration to address the overall magnitude of ET and  $Q$  as well as to optimize the spatial pattern performance of ET. For the magnitude of ET, the mean absolute error (MAE) was used.

$$MAE = \frac{\sum_{i=1}^n |x_i - y_i|}{n} \quad (7)$$

For the magnitude of  $Q$ , the mean error (ME) was applied.

$$ME = \frac{\sum_{i=1}^n x_i - y_i}{n}, \quad (8)$$

where  $x_i$  is the observation,  $y_i$  is the simulation, and  $n$  is the number of data. MAE and ME have an optimal value of 0 and vary from 0 to positive infinity. For the spatial pattern performance of ET, the multi-component Spatial Efficiency (SPAET; Koch et al., 2018, and Demirel et al., 2018) metric was selected.

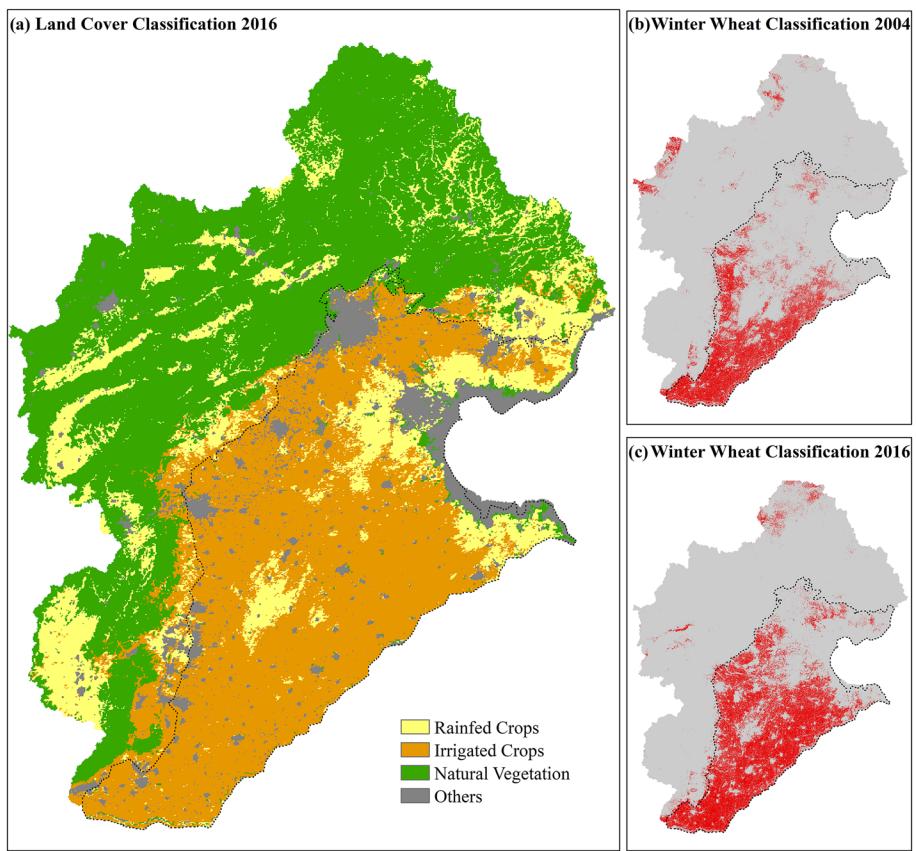
$$SPAET = 1 - \sqrt{(\alpha - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \quad (9)$$

$$\alpha = \rho(x, y) \text{ and } \beta = \left( \frac{\sigma_x}{\mu_x} \right) / \left( \frac{\sigma_y}{\mu_y} \right) \text{ and } \gamma = \frac{\sum_{j=1}^n \min(K_j, L_j)}{\sum_{j=1}^n K_j}$$

where  $\alpha$  is the Pearson correlation coefficient between observed ( $x$ ) and simulated data ( $y$ ),  $\beta$  is the fraction of coefficient of variations representing spatial variability, and  $\gamma$  is the percentage of histogram intersection for an observed histogram  $K$  and a simulated histogram  $L$ , each one containing  $n$  bins, that is, herein 100 bins. SPAET has an optimal value of 1 and varies from 1 to negative infinity.

In order to calibrate the hydrologic model for rainfed conditions, irrigated areas had to be identified and excluded from the objective functions targeting ET. For this purpose, land cover was classified into natural vegetation and cropland based on the MODIS land cover product. In a next step, the MODIS NDVI climatology was analyzed to further split the cropland into irrigated and rainfed grids. For this, two constraints were applied: (1) if the NDVI slope during spring (February to May) was below 0.03 per 16 days and (2) the maximum NDVI of that period was below 0.35, a cropland grid was classified as being rainfed. The resulting classification is depicted in Figure 2 and underlines that large parts of the NCP cropland is irrigated. The map resembles the irrigation classifications shown by Mo et al. (2005) and Guo and Shen (2015). In the coastal plain, irrigation is not feasible from the shallow aquifers due to saltwater intrusion of the aquifers. This classification is applied for the entire simulation period of 2002 to 2016, which disregards any potential expansion of irrigation practices. However, this was found to be the most robust way to obtain a meaningful classification, as NDVI data from single years may be subject to noise. Further, the threshold values used for the two constraints were selected based on expert judgment, which may pose an additional limitation to the classification.

Table 3 states details of the four ET related observation groups and the  $Q$  related observation group. With respect to ET, for the three land cover classes, rainfed cropland, all cropland, and natural vegetation the



**Figure 2.** The map in (a) depicts the land cover classification applied in the ET calibration to differentiate between irrigated cropland, rainfed cropland, and natural vegetation. Two examples of the winter wheat classification maps are shown for 2004 (b) and 2016 (c). NCP domain indicated with dashed line.

MAE of ET was calculated. The MAE of ET over natural vegetation was used for all months throughout the years. The MAE associated to ET over rainfed cropland was utilized during the winter wheat growing season (October–May) and the MAE of ET for all cropland was used during the monsoon months (June–September). Urban areas, water bodies, and barren soil were excluded in the ET calibration. With the given objective functions, the hydrologic model was calibrated exclusively against rainfed ET. SPAEF was calculated for the monthly climatology of the months March to October for simulated and observed ET. The ET patterns during the four winter months (November–February) are characterized by a very low variance, which disqualifies them for a meaningful spatial pattern calibration. In the summer crop season from June to September, SPAEF was calculated for the combined area of natural vegetation and cropland. For the remaining months, irrigation is expected to significantly affect ET, and thereby, SPAEF was solely calculated for the combined area of natural vegetation and rainfed cropland. For Q, the ME for each of the eight stations (Figure 1) was used as objective function. Achieving the best possible accuracy of Q dynamics is not the focus of this study, and therefore, a simple water balance objective function, as the

**Table 3**  
*Overview of the Applied Calibration Scheme*

Observation group	Objective function	Months	Spatial mask	Temporal aggregation	Weight (%)
ET: Spatial Pattern	SPAEF	March–October	June–September: C_all and NV October–May: C_rain and NV	Monthly ET climatology	20
ET: Cropland Rainfed (C_rain)	MAE	October–May	C_rain	Monthly ET	10
ET: Cropland All (C_all)	MAE	June–September	C_all	Monthly ET	20
ET: Natural Vegetation (NV)	MAE	all months	NV	Monthly ET	10
Discharge (Q)	ME	all months	-	Daily Q	40

ME, has been applied. Anthropogenic water diversion affects the discharge of large parts of the HRB; therefore, the eight stations used for calibration are located in the mountainous region where predominantly natural flow conditions can be expected.

In PEST, a total error phi ( $\phi$ ), representing the sum of the squared residuals of all objective functions is used in the calibration.

$$\phi = \sum_{i=1}^l ((1 - SPAEF_i) \cdot \omega_{SPAEC_i})^2 + \sum_{j=1}^m ((0 - MAE_j) \cdot \omega_{MAE_j})^2 + \sum_{k=1}^n ((0 - ME_k) \cdot \omega_{ME_k})^2, \quad (11)$$

where  $l$  is the number of months for which SPAEF is calculated,  $m$  is the number of months for which MAE is calculated, and  $k$  is the number of discharge stations for which ME is calculated.  $\omega$  are the weights that are assigned to the residuals. The residuals can vary depending on the range of the objective function and the magnitude of the variable. This requires a weighting scheme that harmonizes these differences (Stisen et al., 2018). The weighting of each objective function has been implemented with respect to the residuals as obtained from the initial parameter set and defined in a way so that the total error phi equals to one with the initial parameter set. The weights are indicated in Table 3 and are not adjusted during calibration.

A local sensitivity analysis was conducted in order to identify the most sensitive parameters with respect to the defined objective functions. The initial parameter values were perturbed one at a time, and 30 parameters were found to be sensitive. An overview is given in Table S2 (ST2) in the supporting information to this article.

The purpose of the hydrologic model calibration is to obtain robust rainfed ET for the entire 15 years of the study period. All available Q and ET data from this period (2002–2016) were used in the calibration. The observational data set has not been split into a calibration and validation set because we do not aim at extrapolating the model in space or time, which would require a validation test. Nevertheless, in order to test the robustness of the calibration design, an alternative calibration with a 9 year calibration period (2002–2010) and a 6 year validation period (2011–2016) has been conducted and results are presented in the supporting information Figure S1 (SF1). From this analysis, we concluded that the model performance does not deteriorate in the validation period, which underlines the robustness of the chosen calibration strategy.

### 3.4. Net Irrigation Estimation

#### 3.4.1. Irrigation Amounts

Net irrigation ( $netIrr$ ) amounts are quantified at monthly time scale at 1 km<sup>2</sup> spatial resolution based on the ET residuals of PT-JPL and mHM. Net irrigation refers to the water column depth of the evaporative loss of irrigation water. With the absence of irrigation in the hydrologic model, representing the baseline with purely rainfed conditions, it can be assumed that mHM systematically underestimates ET at times of irrigated crop growth as compared to PT-JPL:

$$netIrr = ET_{PT-JPL} - ET_{mHM}. \quad (12)$$

Given the uncertainties in precipitation forcing as well as the uncertainties in the remote sensing-based ET data, overestimations of  $ET_{mHM}$  are conceivable during periods of high precipitation. This results in negative  $netIrr$  estimations. We hypothesize that the irrigation signal is easier to detect from ET residuals during dry periods than during wet periods where precipitation is high. Therefore, we investigate two hypotheses to quantify net irrigation. The first (H1) neglects negative residuals in Equation 11 whereas the second (H2) takes both positive and negative residuals into consideration. H2 can be considered a conservative estimate of irrigation. H2 is included to shed light on some of the uncertainties related to an approach based on the residuals of two independent ET estimates, each associated with their own uncertainties. In the analysis, net irrigation is separated into a winter and a summer fraction. The first corresponds to the winter wheat growing season (October–May), whereas the latter covers the summer crops (June–September).

The proposed framework yields net irrigation estimates that are only valid within the geographical domain of the hydrological model that was subject to calibration and within the time period where calibration data is available. We do not suggest to use this framework to extrapolate in space or in time. However, the framework can be applied in other regions, where satellite-based ET estimates and suitable hydrologic models

are available. Importantly, calibrating the hydrologic model to perform well under rainfed conditions is essential when transferring the framework to other regions.

### 3.4.2. Irrigation Uncertainty

We anticipate that potential systematic biases between ET mHM and ET PT-JPL over rainfed grids can effectively be removed through the proposed calibration design. Nevertheless, it can be expected that residuals of rainfed ET will remain after calibration. These can be linked to uncertainties in precipitation forcing, uncertainties in the ET calibration data set, and mHM parameters and model structure. We propose that the variance of the remaining residuals for rainfed cropland grids can be utilized to estimate the predictive uncertainty of the net irrigation quantification. This assumption is based on the idea that the variance of the ET residuals for rainfed cropland grids is valid for the ET residuals for irrigated cropland grids. This cannot be validated directly, but based on the parameter regionalization scheme in mHM, we expect similar model behavior within the same land cover due to seamless parameter distributions that are physically consistent.

## 4. Results

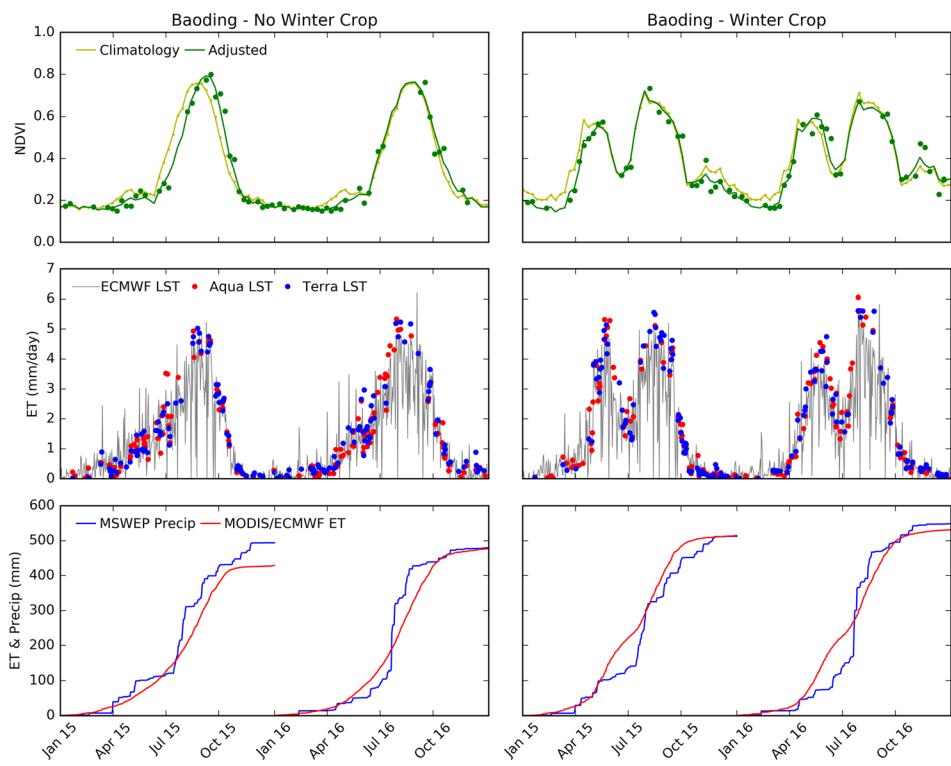
### 4.1. Remote Sensing Model

The PT-JPL model was used to estimate daily ET at 1 km<sup>2</sup> spatial resolution across the entire HRB from 2002 to 2016. As stated in section 3, most of the MODIS derived inputs to the PT-JPL model underwent a processing step using the multiyear climatology to obtain robust daily timeseries with full coverage. Figure 3 exemplifies this processing for NDVI at two grids near the city of Baoding. One of the grids exhibits the typical dual crop rotation system constituted by winter wheat and summer maize, while the other one is characterized by just a single summer crop. In NCP, winter wheat is sown in October, after which the plant goes through dormancy until spring, peaks in May and is typically harvested in June. Afterward, the summer cropping cycle begins which ends with harvest in September. These well-studied crop dynamics are captured accordingly by the NDVI timeseries in Figure 3. The climatology-based processing provides realistic dynamics compared to simple interpolation techniques, which are prone to errors, and allows to differentiate intrayear variability.

Daytime LST data from three different sources, namely, MODIS Terra, MODIS Aqua, and ECMWF ERA-Interim, were acquired to calculate the SM constraint in the PT-JPL model. All of the above use the same MODIS-based nighttime LST data to calculate the ATI (Table 2) and the upward longwave radiation term to estimate net radiation. The results are illustrated in Figure 3, and differences are entirely due to different actual soil evaporation terms, as the canopy transpiration term is not affected by LST as well as differences in net radiation. Overall, the bias between PT-JPL forced by Terra LST and Aqua LST is 0.02 mm d<sup>-1</sup> using only grids with coinciding observations. MODIS LST is only available at clear sky days, and therefore, the derived ET values do not show the same abrupt fluctuations as the ECMWF based ET timeseries, which is a result of low available energy during cloudy days. On some individual days, this results in daily ET rates below 1 mm d<sup>-1</sup> during the summer months, which can be considered unrealistic. The bias on cloud-free days between ECMWF derived ET and Terra and Aqua is 0.03 and 0.04 mm d<sup>-1</sup>, respectively. This underlines that the effect of the proposed gap filling approach using ECMWF based LST is negligible, despite differences in spatial resolution and the fact that the ECMWF LST data is based on reanalysis contrary to MODIS data which are purely observational based.

The resulting daily ET data set is a combination of the three PT-JPL models forced with the abovementioned LST data sets. For the final data set, MODIS LST-based ET was always favored over ERA-Interim LST. In case both, Terra and Aqua provide a LST observation for the same day; the average of the two ET retrievals was used. Annual cumulative distribution functions are plotted in the bottom row of Figure 3. The effect of irrigation becomes evident in the winter wheat example, where the high ET rates in spring are not sustained by the available precipitation, which is a strong indication for an additional nonprecipitation source of water.

Figure 4 depicts the average annual ET pattern obtained from PT-JPL and mHM for the years 2002 to 2016 at 1 km<sup>2</sup> spatial resolution. Based on PT-JPL, the average annual ET for the HRB and NCP domains are 483 and 511 mm yr<sup>-1</sup>, respectively. The highest ET fluxes are found in the mountainous regions north and west of the NCP that are covered by forest. Agricultural activities control the spatial variability of ET in the NCP. ET is generally high in the southern part of the NCP toward the Yellow River. Another region of high ET is the so-called Piedmont Plain, which is the part of NCP that is located on the foothills of the Taihang



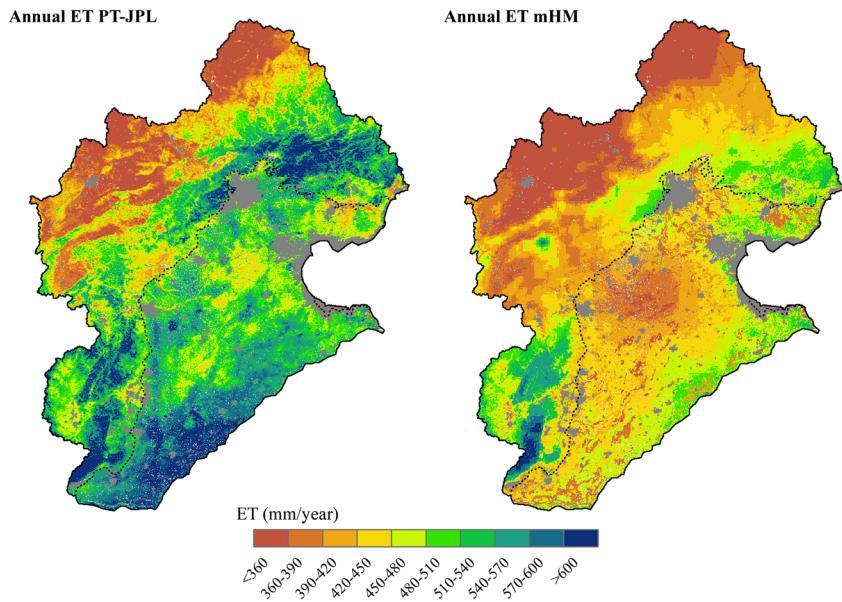
**Figure 3.** Example timeseries for two  $1 \text{ km}^2$  grid cells for 2 years (2015–2016) close to the city of Baoding, located in NCP. The first depicts an example of rainfed crops without winter crop in the left column, and the second showcases the crop rotation system of irrigated winter crop and rainfed summer crop in the right column. The first row shows the multiyear NDVI climatology which is adjusted to capture actual observations (green points). The middle row contains ET based on the PT-JPL model using LST from Aqua, Terra, and ECMWF. The cumulative density functions of precipitation and ET are shown in the bottom row.

Mountains. Low ET corresponds well to the areas of rainfed cropland as classified in Figure 2 and areas of low NDVI in Figure 1.

#### 4.2. Hydrologic Model

The average annual ET based on mHM for the HRB and NCP domains are  $416$  and  $417 \text{ mm yr}^{-1}$ , respectively, and the spatial pattern is depicted in Figure 4. ET is higher in the southeastern boundary of the HRB, which is a consequence of higher precipitation. Other regions of high ET are in the forested parts of the Taihang Mountains. The mHM model is configured to simulate rainfed ET, disregarding the effect of irrigation. This becomes apparent when comparing the annual ET maps of mHM and PT-JPL for the NCP in Figure 4. The comparison of monthly ET aggregated to NCP for PT-JPL, and mHM is shown in Figure 5. The hydrologic model has been calibrated for rainfed conditions and does clearly not reflect the effect of winter wheat irrigation. There is a systematic mismatch between ET simulated by the hydrologic model and the remote sensing-based model during the spring months. According to the mHM simulation, the natural ET variability, which is driven by climate seasonality, has an annual range from  $10 \text{ mm mo}^{-1}$  in winter to  $100 \text{ mm mo}^{-1}$  in summer. Years 2002 and 2014 were characterized by low precipitation, which likely entailed extended summertime irrigation that could explain the underestimations of summertime ET of mHM for the respective years. After calibration, the average monthly MAE of ET over rainfed cropland was  $4.8 \text{ mm mo}^{-1}$  for the summer months and  $10.0 \text{ mm mo}^{-1}$  for the winter months. The MAE of ET for natural vegetation was  $11.6 \text{ mm mo}^{-1}$ . The spatial pattern metric SPAEF has an optimal value of 1, and the calibrated rainfed ET patterns for the multiyear averages of the months March till October varied between 0.14 (July) and 0.71 (September) with an average of 0.5.

The ME at the eight discharge stations was used as calibration target and the resulting Q performance is stated in Table 4. Q data was included in the calibration to get the overall water balance in place. Nevertheless,



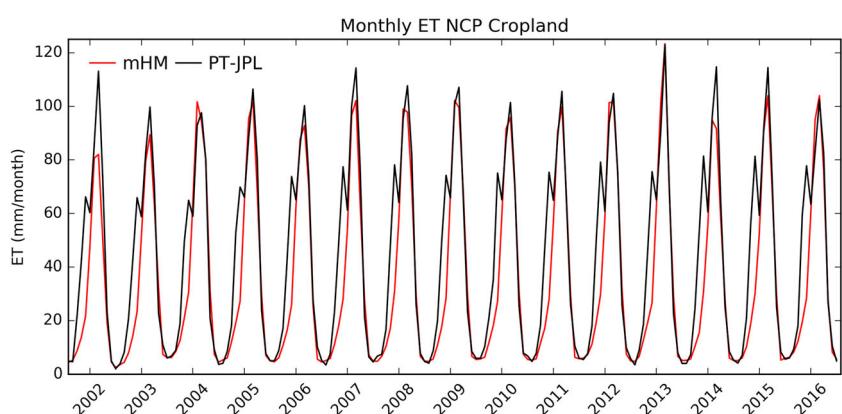
**Figure 4.** Multiyear average (2002–2016) of ET based on PT-JPL (left panel) and mHM (right panel). NCP domain indicated with dashed line.

some stations are associated with large errors that may be a result of anthropogenic interference and uncertainties in the precipitation forcing or Q observations.

Figure 6 provides an overview of the key annual water balance components of NCP. Despite two drought years in 2002 and 2014, there is no notable trend in precipitation. Based on PT-JPL, ET more or less equals precipitation, considering that not all available water is likely to evaporate and transpire but will also generate recharge and runoff, this is a possible indicator for the extensive irrigation scheme. The variance across the 15 years of annual ET simulated by mHM is much larger than for PT-JPL. Thereby, irrigation counteracts precipitation variability, keeping ET more constant than it would be under natural conditions. All water balance components of NCP and HRB are stated in Table S3 (ST3) in the supporting information to this article. For the HRB, discharge amounts to approximately 10% of precipitation and recharge constitutes 8.5%.

#### 4.3. Net Irrigation Estimation

Irrigation was quantified based on the ET residuals from a hydrologic model and a remote sensing model following two hypotheses. Hypothesis H1 neglects any negative residuals (ET overestimations by mHM) whereas hypothesis H2 takes both, positive and negative residuals into consideration. Logically, annual



**Figure 5.** Timeseries of monthly ET obtained for all cropland (irrigated and rainfed) in NCP by the remote sensing model (PT-JPL) and the calibrated hydrologic model (mHM).

**Table 4**  
*Discharge Performance Obtained Through Calibration*

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Average Observed ( $\text{m}^3 \text{s}^{-1}$ )	1.6	12.9	6.1	6.0	2.0	8.7	7.7	2.0
Average Observed ( $\text{mm yr}^{-1}$ )	24.0	22.2	63.5	53.4	5.4	19.4	38.5	12.5
ME ( $\text{m}^3 \text{s}^{-1}$ )	-1.2	0.3	-4.9	-4.5	4.5	-1.6	-0.4	3.4
ME (% of observed)	-79	2	-80	-74	221	-18	-5	174

Note. The station names refer to the ones in Figure 1 and to the IDs in Table 1. The first two rows state the observed discharge. The residuals were calculated by subtracting the observed from the simulated discharge.

net irrigation based on H1 is larger than H2 with an average of  $126 \text{ mm yr}^{-1}$  ( $15.2 \text{ km}^3 \text{ yr}^{-1}$ ) for NCP and  $108 \text{ mm yr}^{-1}$  for HRB ( $18.6 \text{ km}^3 \text{ yr}^{-1}$ ).

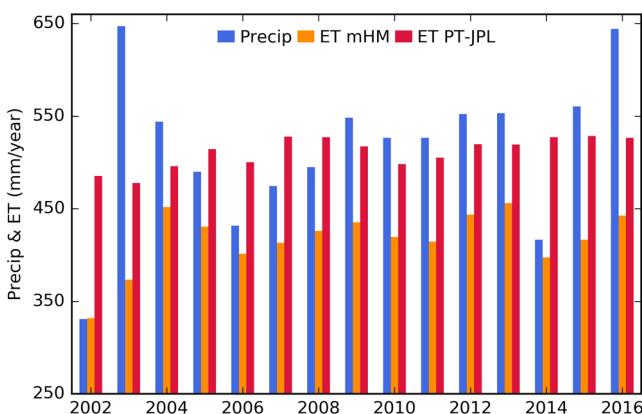
Figure 7 illustrates the spatial pattern of mean annual net irrigation. Irrigated agriculture is concentrated along the Piedmont Plain and the southern part of NCP, which corresponds well with areas of high ET based on Figure 4. The spatial resolution of  $1 \text{ km}^2$  reveals many interesting details on the irrigation pattern, such as the absence of irrigation agriculture along the broad riverbeds intertwined in the Piedmont plain. The irrigation analysis is performed at monthly timescale which allows to separate the irrigation activities into a summer and a winter fraction. Following the results presented in Figure 7, the majority of irrigation takes place during the winter wheat cropping period between October and May. Summertime irrigation is generally lower and limited to the center part of NCP where large-scale fruit orchards are located. On average, 77% of the annual net irrigation takes place during the winter cropping season. At monthly scale, May is the month with the largest fraction of annual irrigation (33%), followed by April (21%) and August (10%).

The differences between the two hypotheses are investigated in Figure 7 to analyze the uncertainties related to neglecting negative ET residuals in H1. The uncertainties are large along the southern and northeastern boundaries of the NCP and the Piedmont Plain. However, a majority of the uncertainty (58%) can be attributed to the net irrigation during the summer months. During summer, precipitation is high which makes it more challenging to isolate the irrigation signal. Since irrigation is not equally divided between summer and winter cropping season, the irrigation uncertainty is 11.2% for the winter season and 48.0% for the summer season relative to the irrigation estimates using H1.

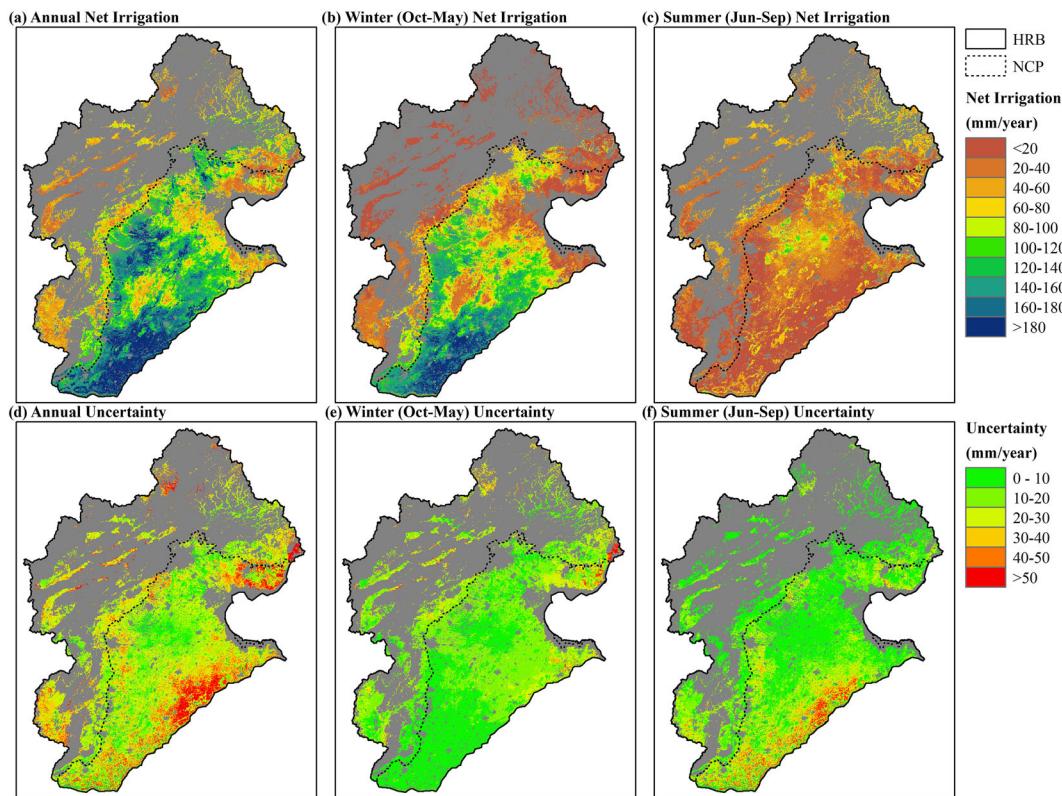
Figure 8 further investigates the interannual variability of net irrigation, the partitioning between winter and summer crop as well as the differences between H1 and H2. For NCP, the annual net irrigation varies between  $89.4$  ( $10.7 \text{ km}^3 \text{ yr}^{-1}$ ) and  $168.8 \text{ mm yr}^{-1}$  ( $20.3 \text{ km}^3 \text{ yr}^{-1}$ ). The variability partly relates to precipitation anomalies, as the two driest years, 2002 and 2014, show the largest irrigation. However, this dependency seems to be only valid to the summer crop irrigation, which is in phase with the monsoon precipitation and therefore more dependent on precipitation. The differences between H1 and H2 are largely controlled by

summer irrigation, which underlines that winter irrigation amounts are estimated with a higher certainty. Following H2, some years (that is, 2004, 2012, 2013, and 2016) have an overall negative summer irrigation as consequence of a systematic overestimation of ET in mHM. These results emphasize the larger uncertainties related to the summertime irrigation quantification, when precipitation is high and overall energy limited conditions are present, in comparison to the wintertime assessment, when precipitation is low and ET is limited by the water availability. In the 4 years with negative summer net irrigation (H2), values range between  $-1.5$  and  $-11.4 \text{ mm}$  which corresponds to less than 4% of summertime ET.

Figure 9 depicts the monthly climatology for net irrigation (H1) for NCP. With  $31$  and  $47 \text{ mm mo}^{-1}$ , the months April and May, respectively, stand out as the months with the highest irrigation water use. The summer cropping season (June–September) varies between  $4$  and  $10 \text{ mm mo}^{-1}$ . In addition to the uncertainty analysis related to the two different hypotheses (Figure 7), Figure 9 depicts the predictive uncertainty of monthly



**Figure 6.** Overview over annual precipitation (MSWEP v2) and ET based on PT-JPL and mHM for the NCP domain.

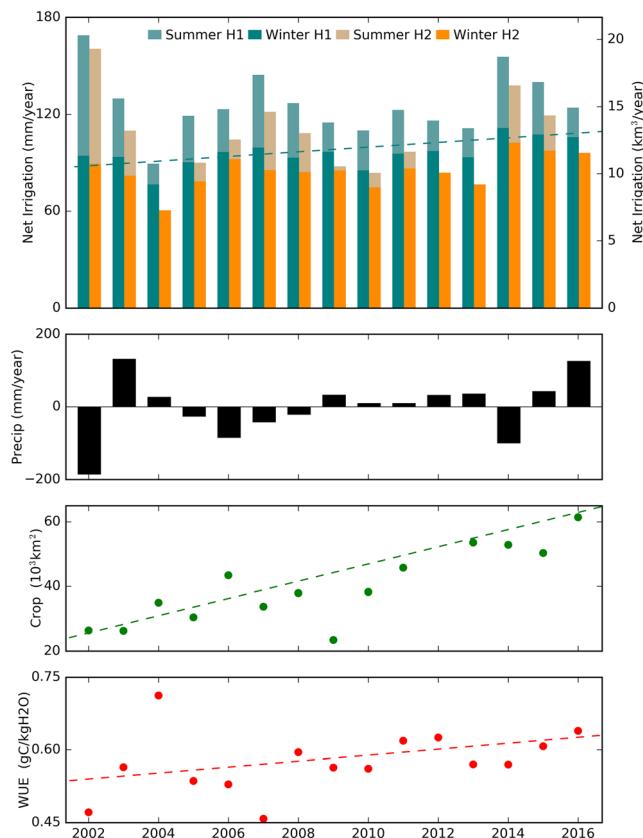


**Figure 7.** (a) Average annual net irrigation (2002–2016) as obtained from H1 (section 3.3). (b and c) Split (a) into a winter and a summer fraction, respectively. (d–f) The differences between the two hypotheses (H1 and H2) to investigate uncertainties of (a), (b), and (c), respectively.

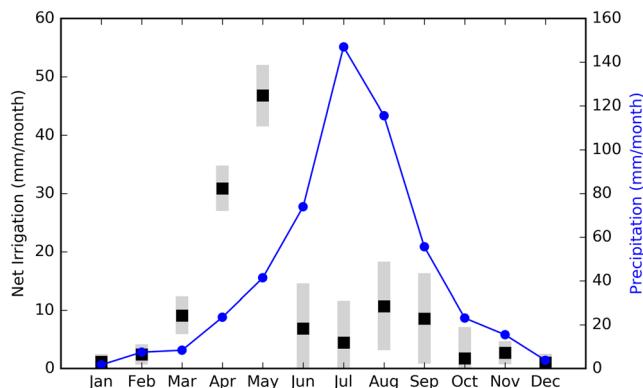
net irrigation quantified based on H1 following the method outline in section 3.4.2. The confidence bars represent  $\pm$  one standard deviation which is derived from the variability of the ET residuals over rainfed cropland. The rainfed residuals have a mean of  $3.1 \text{ mm mo}^{-1}$  and a standard deviation of  $10.4 \text{ mm mo}^{-1}$ . This low bias underlines that the systematic residual between ET mHM and ET PT-JPL was removed through calibration. The remaining variability of the residuals can be interpreted as a consequence of uncertainties in the hydrologic model as well as uncertainties in ET PT-JPL. The predictive uncertainty is most pronounced during the summer cropping season (June–September) which coincides with the Monsoon period with high precipitation. This is supported by the coefficient of variation on the aggregated values for the summer and winter cropping seasons. From this, we can derive that the standard deviation is 7.4% of the mean for the winter cropping season and 43.8% for the summer cropping season. For average annual conditions, the coefficient of variation suggests that the standard deviation is 39.5% of the mean. Evidently, most of this uncertainty can be related to the summer irrigation.

#### 4.4. Net Irrigation Evaluation

Figure 8 also contains the development of winter wheat cultivation area in NCP, which is characterized by a clear increasing trend of  $2,200 \text{ km}^2 \text{ yr}^{-1}$ . This trend does not entail a clearly increasing trend of winter wheat irrigation amounts, which suggests that irrigation water use may have become more efficient. This is supported by the analysis of MODIS derived NPP data. Annual WUE was calculated using the ET obtained by PT-JPL. The annual NPP data over NCP suggests an increasing trend of  $4.1 \text{ g C m}^{-2} \text{ yr}^{-1}$  for NCP. After conversion to WUE, an increasing trend of  $0.005 \text{ g C kg}^{-1} \text{ H}_2\text{O yr}^{-1}$  can be attested to the data (Figure 8). The detailed winter wheat classification maps are a valuable source to evaluate the net irrigation estimates spatially. Figure 10 depicts the winter wheat classification for the years 2004 and 2016, as already shown in Figure 2. Based on the selected years, the area used for winter wheat expanded from approximately 38,000 to 65,000  $\text{km}^2$  which marks an increase of 70%. The continuous winter net irrigation estimates are classified into four classes for better visual comparison with the winter wheat classification. Overall, the



**Figure 8.** (top row) Annual analysis of net irrigation based on the two hypotheses (H1 and H2) splits up into a winter and a summer fraction. The net irrigation water depth is in relation to the NCP area. The dashed line indicates the fitted linear trend of winter irrigation based on H1 with slope =  $1.2 \text{ mm yr}^{-1}$ . (second row) Annual precipitation anomalies with respect to the 2002 to 2016 mean. (third row) Annual winter wheat area, with fitted linear trend (dashed line) with slope =  $2,200 \text{ km}^2 \text{ yr}^{-1}$ . (bottom row) Annual WUE, with fitted linear trend (dashed line) with slope =  $0.005 \text{ g C kg}^{-1} \text{ H}_2\text{O yr}^{-1}$ .



**Figure 9.** Climatology of estimated net irrigation based on H1 (2002–2016). Black squares indicate mean monthly net irrigation, and gray boxes indicate the uncertainty, quantified as  $\pm$  one standard deviation. Monthly precipitation climatology, in blue, is shown on the second y axis.

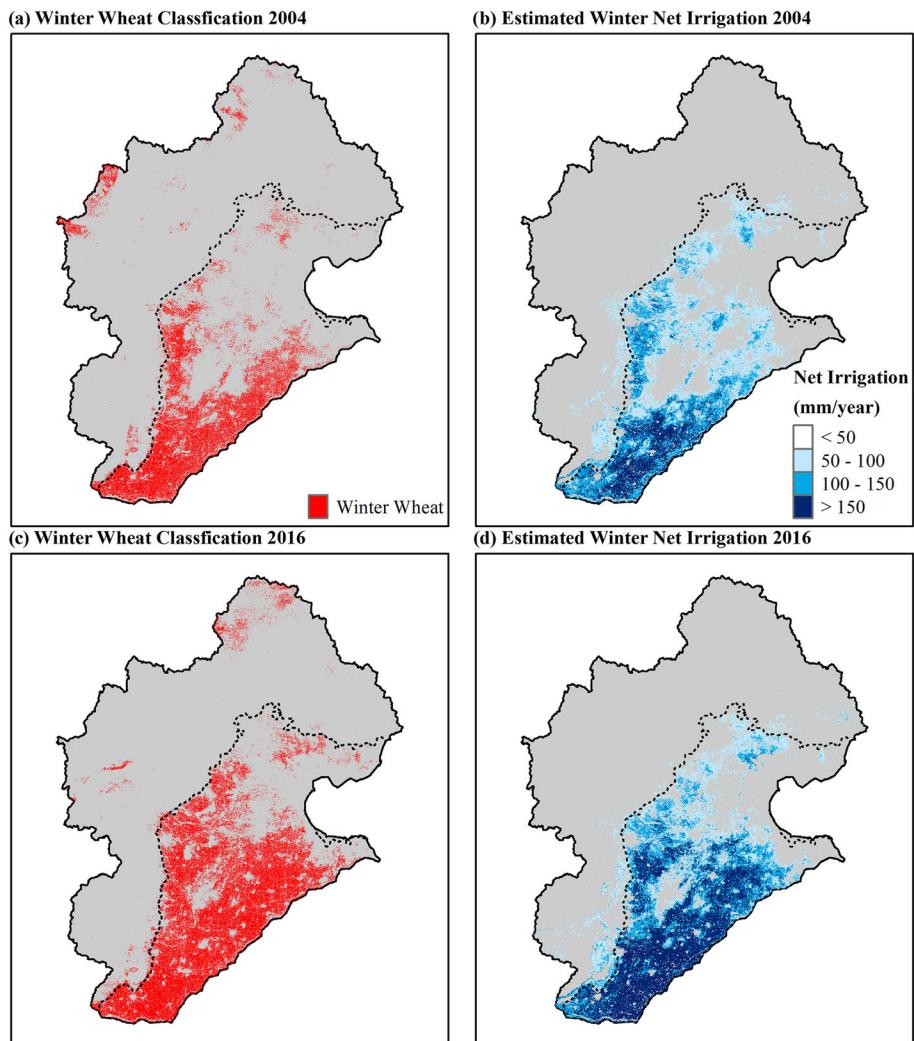
spatial expansion is well represented between the two approaches. Winter wheat and thereby irrigation expands distinctly in the eastern and northern parts of NCP, and in general, the cropping area becomes denser. Based on the three selected thresholds in Figure 10, greater than 50, 100, and  $150 \text{ mm yr}^{-1}$ , the irrigated areas expand from  $83,000$ ,  $39,000$ , and  $12,000 \text{ km}^2$  in 2004 to  $98,000$ ,  $66,000$ , and  $35,000 \text{ km}^2$  in 2016, which marks an increase of 18%, 69%, and 190%, respectively. The best visual agreement with the winter wheat classification maps is obtained with the second threshold, namely, greater than  $100 \text{ mm yr}^{-1}$ .

The GRACE-based TWSA clearly support the observed groundwater depletion of the NCP with a decreasing trend of  $-11.8 \text{ mm yr}^{-1}$  (Figure 11). The monthly TWSA simulated by mHM were calculated based on hydrologic state variables at the land surface, in the soil layers, and in the subsurface. In order to make the mHM storage simulations comparable to GRACE, the same time mean baseline period (2004 to 2009) was used to obtain the mHM anomalies. The storage anomalies of mHM possess a slight positive trend ( $3.7 \text{ mm yr}^{-1}$ ). Thereby, mHM does not follow the observed GRACE signal, which constitutes that the negative trend in the GRACE data ( $-11.8 \text{ mm yr}^{-1}$ ) cannot be attributed to climate variability. The fitted trend lines cross around the years 2006 and 2007 which is a result of the way the anomalies are calculated, i.e. using the period of 2004 to 2009 as baseline. mHM has an overall dampened TWSA amplitude in comparison to GRACE, due to the absence of groundwater withdrawals for irrigation as well as a general simplified groundwater description. The comparison of GRACE and mHM is further hampered by the absence of key processes in mHM controlling water storage, such as reservoirs, wetlands, and water diversion. The trend-corrected TWSA climatology of mHM and GRACE underlines the effect of extensive irrigation activities (Figure 11b). Both have a clear increasing trend in the summer monsoon months from June to August driven by high precipitation. The TWSA data disagrees more in the spring months where GRACE shows a clear negative trend induced by groundwater abstractions for irrigation, whereas this is not captured by the purely rainfed mHM setup. These findings are further supported by the total water storage change (TWSC) calculated as the difference of TWSA in a particular month and the TWSA of the subsequent month (Figure 11c). Agreement between mHM and GRACE can be found in the summer months that are mainly driven by precipitation whereas a strong disagreement can be attested to the spring months where GRACE possess negative TWSC that are not represented by mHM. The springtime discrepancies can be alleviated by taking the estimated net irrigation amounts into consideration and subtracting them from the mHM based TWSC.

## 5. Discussion

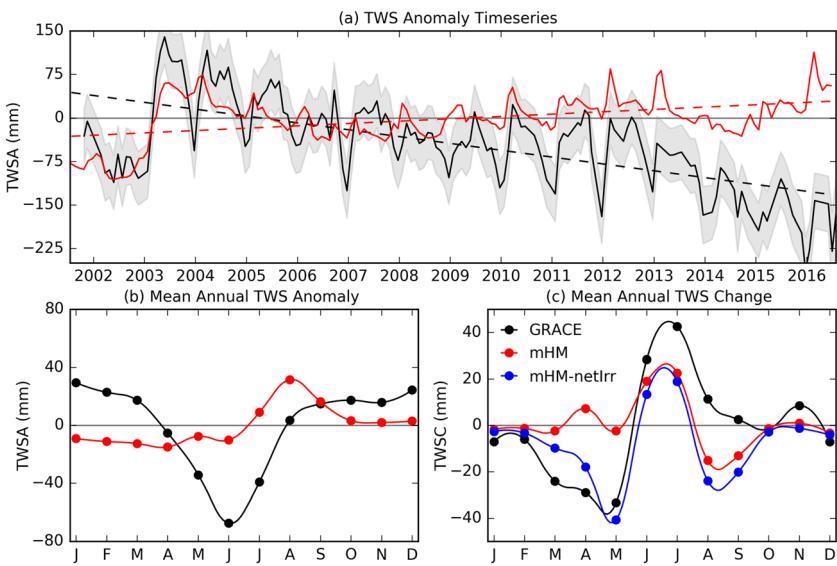
### 5.1. Irrigation in the NCP

There exists a broad variety of NCP irrigation studies in the water resources as well as in the agronomy literature that can be utilized to evaluate our results. However, direct comparison is not always trivial due to deviating study periods, spatial resolutions that range from plant scale studies to administrative unit scale, but more importantly, the term irrigation can have different notions, such as optimal crop irrigation water requirement, the actual applied irrigation to the field, or the net irrigation as the



**Figure 10.** Spatial evaluation of the net irrigation estimates for NCP. The winter wheat classification is shown as reference for the years 2004 and 2016 in panels (a) and (c), respectively. The estimated winter irrigation is classified in four classes for the same years (panels b and d) to be comparable with the binary winter wheat map.

actual evaporative loss. Yang et al. (2010) applied agronomic crop models and reported an overall irrigation water requirement of  $16.5 \text{ km}^3 \text{ yr}^{-1}$  in 2001 with the highest requirements along the Piedmont Plain and the southern part of NCP. Moreover, April and May were found to be the months accounting for the largest fraction of the annual irrigation, with 18% and 25%, respectively. These findings are in very good agreement with our analysis. Likewise, average agricultural water use for the HRB was estimated to be around  $17.7 \text{ km}^3 \text{ yr}^{-1}$  by Shen et al. (2015). Yin et al. (2020) compared ET residuals obtained from two soil water balance equations, where one included an irrigation scheme and the other did not, to calculate annual irrigation at  $10 \text{ km}^2$  over China. Results suggest a range between 50 and 200  $\text{mm yr}^{-1}$ , which is in fair agreement with our results. Hu et al. (2016) estimated average annual irrigation as the residual term of the soil water balance equation which amounted to  $317 \text{ mm yr}^{-1}$  with irrigation to ET ratios of about 0.5. Our irrigation estimates are half of their reported values. To solve the soil water balance equation, Hu et al. (2016) interpolated in situ SM data for NCP and groundwater recharge was estimated and interpolated based on in situ tracer experiments, which may have introduced large uncertainties. A GRACE based water balance analysis yielded annual ET of  $521 \text{ mm yr}^{-1}$  which was compared to three land surface models (GLDAS) without irrigation schemes (Pan et al., 2017). GRACE based ET was 12% higher than the GLDAS models, which is in good agreement with our analysis, where PT-JPL based ET is 13.8% higher than mHM. An integrated subsurface-surface hydrologic model was applied by Qin et al. (2013) where irrigation



**Figure 11.** The monthly total water storage anomalies (TWSA) based on GRACE and mHM for NCP are shown in (a) with their respective fitted linear trends as dashed lines. The gray shaded envelop presents the uncertainty ( $\pm$  one standard deviation) associated to GRACE. The average TWSA are shown in (b) based on the trend-removed monthly data in (a). The average TWSC based on the data in (a) is illustrated in (c) including a scenario where mHM was corrected for net irrigation (netIrr).

amounts and frequencies were prescribed based on literature. Annual NCP irrigation was estimated to be around  $180 \text{ mm yr}^{-1}$ , which is slightly higher than our estimates. For a similar model setup, irrigation was reported to be  $290 \text{ mm yr}^{-1}$  for an irrigation district within NCP (Shu et al., 2012). The two abovementioned studies specified actual irrigation amounts applied to the field, and return flows have to be considered before being directly comparable to net irrigation estimations. Despite the deviation to our findings, we regard our approach more trustworthy as it is more observational based compared to the simple deficit rules applied in purely model based approaches. Nevertheless, our proposed approach is not purely observational based as it relies on a hydrologic model as well. Based on the literature review, our irrigation estimates at  $1 \text{ km}^2$  provide critical information at high spatial and temporal resolution for large-scale water resources mapping and management. Such data can build an important asset in future research as boundary condition of groundwater models investigating depleting aquifers (Cao et al., 2013), input to water resources management scenarios (Huanhuan Qin et al., 2019), or calibrating irrigation parameters in land surface models (Lei et al., 2015), which build an important boundary condition to regional climate models. The importance of irrigation for the NCP water crisis has been well discussed in literature. However, recent NCP studies also highlight the interactions between irrigation and the atmosphere, resulting in a cooling of the land surface (Yang et al., 2020) or increasing the risk of heatwaves due to increases in humidity (Kang & Eltahir, 2018). This further promotes the importance of our work, as more detailed knowledge on irrigation may help explain the complex micro-climatic interactions.

Based on the MODIS-based NPP and ET from PT-JPL, we could draw the conclusion that the irrigation WUE has improved in NCP since the early 2000s. This was supported by the annual winter wheat classification, which conveyed a clear increase while the annual irrigation amounts were quite stagnant. Fang et al. (2020) found a significant trend in winter wheat ET of  $1.28 \text{ mm yr}^{-1}$  due to anthropogenic influence, which can be supported by our irrigation results as seen by the trend line in Figure 8. However, this trend does not correspond to the doubling in winter wheat cultivation area, which supports the increase in WUE. Mo et al. (2017) studied trends in ET and gross primary productivity for NCP and found increasing WUE in the winter wheat growing season. Similar findings were presented by Liu et al. (2015) who analyzed WUE trends over entire China. More locally, these findings were supported by Zhang et al. (2017) and Lu et al. (2016) for detailed yield and ET records at agronomic research sites in NCP.

## 5.2. Irrigation Uncertainties

At the core of the irrigation quantifications lies the dual modeling of ET using a rainfed hydrologic model and a remote sensing-based ET model, both of which are subject to uncertainties. PT-JPL was used for the latter, and generally, it has been reported that PT-JPL provides accurate ET estimations, especially under semiarid conditions (Fisher et al., 2008; García et al., 2013; McCabe et al., 2019). Based on PT-JPL, ET was estimated to be  $511 \text{ mm yr}^{-1}$  for NCP and  $483 \text{ mm yr}^{-1}$  for HRB. Ground truth ET observations to evaluate the PT-JPL model were not available. Despite that, the difference in scale between flux tower or lysimeter ET data and remote sensing data remains a challenge. Based on various modelling approaches, annual ET rates ranging from  $480$  to  $600 \text{ mm yr}^{-1}$  for NCP and from  $500$  to  $560 \text{ mm yr}^{-1}$  for HRB have been reported in the literature (Guo & Shen, 2015; Hu et al., 2016; Li et al., 2008; Li et al., 2013; Mo et al., 2005; Pan et al., 2017; Qin et al., 2013), which underlines the general plausibility of the PT-JPL results. Based on in situ eddy covariance ET observations at several agricultural sites in NCP, daily ET reached approximately  $6 \text{ mm d}^{-1}$  during the peak of the cropping season (Guo & Shen, 2015; Lei & Yang, 2010; Shu et al., 2011) which is in good agreement with the daily PT-JPL dynamics (Figure 3).

We believe that the choice of hydrologic model is less crucial than the choice of precipitation forcing to the model for the estimation of rainfed ET, as long as the hydrologic model can be set up not to simulate irrigation. We applied mHM due to its favorable regionalization scheme which enables the simulation of physically meaningful spatial patterns of hydrological states and fluxes (Demirel et al., 2018; Samaniego et al., 2017). MSWEP v2 was used as precipitation forcing, which recently has been reported to be accurate for China (Xu et al., 2019). The HRB total water storage trend of the GLDAS models was found to be  $2.7 \text{ mm yr}^{-1}$  (Pan et al., 2017), which is in good agreement with the  $3.7 \text{ mm yr}^{-1}$  predicted by mHM. The ability of mHM to simulate rainfed ET was ensured by means of the proposed calibration strategy. Uncertainties may arise due to fact that minor irrigation also takes place during the summer crop season, which was assumed to be rainfed in our calibration design. This simplification in combination with uncertain precipitation forcing may result in overestimations of ET in mHM, which we addressed by applying two hypotheses for the estimation of net irrigation. Applying MAE and SPAEF as ET objective functions allowed to remove any potential bias between ET mHM and ET PT-JPL. The remaining variability of ET residuals for rainfed grids was employed to quantify the predictive uncertainty of the irrigation estimates. Here we assume that the variance of rainfed cropland can be applied to irrigated cropland. The mHM parameter regionalization schemes generate seamless parameter fields that are physically consistent which strengthens this assumption. However, a limitation may lie in that fact that we apply a time invariant classification of rainfed and irrigated crops.

Comparing the two hypotheses to quantify net irrigation as well as the monthly predictive uncertainty of H1 revealed that winter crop irrigation could be estimated with a higher certainty than summer crop irrigation. This relates to the fact that it is easier to isolate the irrigation signal during dry periods in comparison to wet periods where precipitation is the dominating source. Overall, net irrigation quantification based on ET residuals is more certain during water limited conditions than energy limited conditions. This finding are in good agreement with Jalilvand et al. (2019) who drew very similar conclusions for the irrigation quantification over a semiarid region in Iran.

In principle, different ET products, such as GLEAM, ALEXI, PML, or MOD16, and different P products, such as TRMM, CHIRPS, or PERSIANN, are available and could be applied as alternative reference and forcing data sets in the hydrologic model calibration. Future research of irrigation quantification should utilize ensembles of ET and P products to investigate uncertainties related to the irrigation quantifications. Each combination of P and ET should be calibrated, which may diminish the effect of potential biases between the products. Supporting information SF2 shows an ensemble of selected P and ET products aggregated to monthly and annual scales for cropland and rainfed cropland in NCP. Not surprisingly, differences among the ET products are larger than among the P products. The effect of irrigation, that is, yielding elevated ET rates, is notable for all ET products but varies in characteristic. Therefore, further research is needed on how to evaluate the suitability of a given ET product to be applied to quantify net irrigation.

The proposed approach estimates net irrigation, that is, the evaporative loss of irrigation water, which will naturally be smaller than the actual irrigation applied to the fields (Van Dijk et al., 2018). Flood irrigation is typically practiced in NCP (Cao et al., 2013), and irrigation return flows can be significant. Shen

et al. (2015) reported that return flows constitute 15% of recharge to the shallow aquifer in HRB, which relates to approximately  $1.5 \text{ km}^3 \text{ yr}^{-1}$ . The advantage of estimating net irrigation is that the uncertain assumptions on return flows are not required.

### 5.3. Irrigation Management

The central government of China has from the 1950s to 2010 supported the development of the Chinese irrigation infrastructure (Liu et al., 2013). Ever since the water scarcity in several regions of China became increasingly evident, water scarcity alleviating measures, such as increasing WUE, has been given special attention in policies and guidelines. Across NCP, the cultivation of grain crops like the winter wheat summer maize crop rotation system is mostly done on family-run small parcels of land with an average size of 0.1 hectare (Chen et al., 2011), which complicates the implementation of water policies. Moreover, the political plans and guidelines are challenged by traditional means of flood irrigation that has long prevailed and is applied on more than 70% of irrigated land in China, according to Deng et al. (2006). In addition to tripling the investments in agricultural research from 7 billion renminbi in 2000 to 24.4 billion renminbi in 2009, the Chinese government has taken initiatives to transfer know-how on increasing WUE from experimental research fields to practice. The fundamental role of agriculture in China's economy and food security complicates economic reforms in agricultural water management and the sector is still subsidized, not realizing cost recovery of irrigation water supply (Shen & Wu, 2017). Our analysis suggested an increase in WUE across the NCP, which can be supported by the described efforts toward a more sustainable water resource management in eastern China. Despite the past advances of increasing WUE, groundwater abstraction is still unsustainable and groundwater tables decline by approximately  $4 \text{ cm yr}^{-1}$  across NCP with accelerating depletion rates since 2013 (Zhao et al., 2019).

## 6. Conclusions

This study brings forward a novel framework to estimate net irrigation amounts at regional scale for the HRB, encompassing the NCP, based on dual modeling of ET. The systematic differences between a rainfed hydrological model and a remote sensing-based model of ET provide realistic irrigation estimates at high spatiotemporal detail. We draw the following general conclusion from our work:

1. Calibrating the hydrological model for rainfed ET conditions contributes to the fidelity of the irrigation estimates, by removing potential biases between remote sensing-based ET and rainfed ET simulated by a hydrologic model.
2. The irrigation signal can be isolated with higher certainty during dry periods, whereas high precipitation leads to more ambiguous irrigation amounts in the wet periods. The predictive uncertainty of irrigation quantification during the wet period was found to be roughly 6 times larger than during the dry period.
3. Annual net irrigation is estimated to be 128 and  $106 \text{ mm yr}^{-1}$  for NCP and HRB, respectively, which constitutes approximately 25% of ET.
4. Summer irrigation is more sensitive to interannual precipitation variability, while winter irrigation is less affected.
5. GRACE-based total water storage data underline the plausibility of the quantified irrigation amounts.
6. Assessing winter wheat crop area and winter irrigation amounts imply increasing areas under irrigation accompanied by an increase in WUE, which is supported by MODIS-based NPP data.

## Data Availability Statement

The monthly ET data sets used to conduct the analysis are shared by the authors and made available via the Pangaea database (<https://doi.pangaea.de/10.1594/PANGAEA.914113>).

## References

- Abolafia-Rosenzweig, R., Livneh, B., Small, E. E., & Kumar, S. V. (2019). Soil moisture data assimilation to estimate irrigation water use. *Journal of Advances in Modeling Earth Systems*, 11, 3670–3690. <https://doi.org/10.1029/2019MS001797>
- Arsenault, R., Poulin, A., Côté, P., & Brissette, F. (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering*, 19, 1374–1384. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000938](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000938)
- Bazzi, H., Baghdadi, N., Ienco, D., El Hajj, M., Zribi, M., Belhouchette, H., et al. (2019). Mapping irrigated areas using Sentinel-1 Time series in Catalonia, Spain. *Remote Sensing*, 11, 1836. <https://doi.org/10.3390/rs11151836>

- Beck, H. E., Van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & De Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21, 589–615. <https://doi.org/10.5194/hess-21-589-2017>
- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., et al. (2019). MSWep v2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, 100, 473–500. <https://doi.org/10.1175/BAMS-D-17-0138.1>
- Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., & Fernández-Prieto, D. (2018). How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. *International Journal of Applied Earth Observation and Geoinformation*, 73, 752–766. <https://doi.org/10.1016/j.jag.2018.08.023>
- Cao, G., Zheng, C., Scanlon, B. R., Liu, J., & Li, W. (2013). Use of flow modeling to assess sustainability of groundwater resources in the North China Plain. *Water Resources Research*, 49, 159–175. <https://doi.org/10.1029/2012WR011899>
- Chen, X. P., Cui, Z. L., Vitousek, P. M., Cassman, K. G., Matson, P. A., Bai, J. S., et al. (2011). Integrated soil-crop system management for food security. *Proceedings of the National Academy of Sciences of the United States of America*, 108(16), 6399–6404. <https://doi.org/10.1073/pnas.1101419108>
- Cook, B. I., Shukla, S. P., Puma, M. J., & Nazarenko, L. S. (2015). Irrigation as an historical climate forcing. *Climate Dynamics*, 44, 1715–1730. <https://doi.org/10.1007/s00382-014-2204-7>
- Davidsen, C., Pereira-Cardenal, S. J., Liu, S., Mo, X., Rosbjerg, D., & Bauer-Gottwein, P. (2015). Using stochastic dynamic programming to support water resources management in the Ziya River Basin, China. *Journal of Water Resources Planning and Management*, 141, 04014086. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000482](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000482)
- Deines, J. M., Kendall, A. D., Crowley, M. A., Rapp, J., Cardille, J. A., & Hyndman, D. W. (2019). Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. *Remote Sensing of Environment*, 233, 111400. <https://doi.org/10.1016/j.rse.2019.111400>
- Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., & Stisen, S. (2018). Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model. *Hydrology and Earth System Sciences*, 22, 1299–1315. <https://doi.org/10.5194/hess-22-1299-2018>
- Deng, X. P., Shan, L., Zhang, H., & Turner, N. C. (2006). Improving agricultural water use efficiency in arid and semiarid areas of China. *Agricultural Water Management*, 80(1–3), 23–40. <https://doi.org/10.1016/j.agwat.2005.07.021>
- Doherty, J. (2005). *PEST: Model Independent Parameter Estimation. Fifth Edition of User Manual*. Brisbane: Watermark Numerical Computing.
- Escorihuela, M. J., & Quintana-Seguí, P. (2016). Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes. *Remote Sensing of Environment*, 180, 99–114. <https://doi.org/10.1016/j.rse.2016.02.046>
- Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., et al. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, 38, L03403. <https://doi.org/10.1029/2010GL046442>
- Fang, B., Lei, H., Zhang, Y., Quan, Q., & Yang, D. (2020). Spatio-temporal patterns of evapotranspiration based on upscaling eddy covariance measurements in the dryland of the North China Plain. *Agricultural and Forest Meteorology*, 281, 107844. <https://doi.org/10.1016/j.agrformet.2019.107844>
- FAO, & IIASA (2009). Harmonized world soil database. Food and Agriculture Organization, Rome, Italy & International Institute for Applied Systems Analysis, Vienna, Austria. Retrieved from <https://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45, RG2004. <https://doi.org/10.1029/2005RG000183>
- Feddes, R. A., Kowalik, P., Kolinska-Malinka, K., & Zaradny, H. (1976). Simulation of field water uptake by plants using a soil water dependent root extraction function. *Journal of Hydrology*, 31(1–2), 13–26. [https://doi.org/10.1016/0022-1694\(76\)90017-2](https://doi.org/10.1016/0022-1694(76)90017-2)
- Felfelani, F., Pokhrel, Y., Guan, K., & Lawrence, D. M. (2018). Utilizing SMAP soil moisture data to constrain irrigation in the Community Land Model. *Geophysical Research Letters*, 45, 12,812–892,902. <https://doi.org/10.1029/2018GL080870>
- Fisher, J. B., Tu, K. P., & Baldocchi, D. D. (2008). Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. *Remote Sensing of Environment*, 112, 901–919. <https://doi.org/10.1016/j.rse.2007.06.025>
- Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., et al. (2011). Solutions for a cultivated planet. *Nature*, 478(7369), 337–342. <https://doi.org/10.1038/nature10452>
- García, M., Sandholt, I., Ceccato, P., Ridler, M., Mougin, E., Kerfoot, L., et al. (2013). Actual evapotranspiration in drylands derived from in-situ and satellite data: Assessing biophysical constraints. *Remote Sensing of Environment*, 131, 103–118. <https://doi.org/10.1016/j.rse.2012.12.016>
- Guo, Y., & Shen, Y. (2015). Quantifying water and energy budgets and the impacts of climatic and human factors in the Haihe River Basin, China: 1. Model and validation. *Journal of Hydrology*, 528, 206–216. <https://doi.org/10.1016/j.jhydrol.2015.06.039>
- Hain, C. R., Crow, W. T., Anderson, M. C., & Yilmaz, M. T. (2015). Diagnosing neglected soil moisture source-sink processes via a thermal infrared-based two-source energy balance model. *Journal of Hydrometeorology*, 16, 1070–1086. <https://doi.org/10.1175/JHM-D-14-0017.1>
- Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2), 159–195. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/106365601750190398>
- Hu, X., Shi, L., Zeng, J., Yang, J., Zha, Y., Yao, Y., & Cao, G. (2016). Estimation of actual irrigation amount and its impact on groundwater depletion: A case study in the Hebei Plain, China. *Journal of Hydrology*, 543, 433–449. <https://doi.org/10.1016/j.jhydrol.2016.10.020>
- Huang, Z., Pan, Y., Gong, H., Yeh, P. J. F., Li, X., Zhou, D., & Zhao, W. (2015). Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. *Geophysical Research Letters*, 42, 1791–1799. <https://doi.org/10.1002/2014GL062498>
- Jalilvand, E., Tajrishy, M., Ghazi Zadeh Hashemi, S. A., & Brocca, L. (2019). Quantification of irrigation water using remote sensing of soil moisture in a semi-arid region. *Remote Sensing of Environment*, 231, 111226. <https://doi.org/10.1016/j.rse.2019.111226>
- Kang, S., & Eltahir, E. A. B. (2018). North China Plain threatened by deadly heatwaves due to climate change and irrigation. *Nature Communications*, 9, 2894. <https://doi.org/10.1038/s41467-018-05252-y>
- Kang, S., & Eltahir, E. A. B. (2019). Impact of irrigation on regional climate over Eastern China. *Geophysical Research Letters*, 46, 5499–5505. <https://doi.org/10.1029/2019GL082396>

- Koch, J., Cüneyd Demirel, M., & Stisen, S. (2018). The SPAterial EFFiciency metric (SPAterial EFFiciency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models. *Geoscientific Model Development*, 11, 1873–1886. <https://doi.org/10.5194/gmd-11-1873-2018>
- Kumar, S. V., Peters-Lidard, C. D., Santanello, J. A., Reichle, R. H., Draper, C. S., Koster, R. D., et al. (2015). Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. *Hydrology and Earth System Sciences*, 19, 4463–4478. <https://doi.org/10.5194/hess-19-4463-2015>
- Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, 49, <https://doi.org/10.1029/2012WR012195>
- Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, 48, W04531. <https://doi.org/10.1029/2011WR011453>
- Lawston, P. M., Santanello, J. A., Zaichik, B. F., & Rodell, M. (2015). Impact of irrigation methods on land surface model spinup and initialization of WRF forecasts. *Journal of Hydrometeorology*, 16, 1135–1154. <https://doi.org/10.1175/JHM-D-14-0203.1>
- Lei, H., & Yang, D. (2010). Interannual and seasonal variability in evapotranspiration and energy partitioning over an irrigated cropland in the North China Plain. *Agricultural and Forest Meteorology*, 150(4), 581–589. <https://doi.org/10.1016/j.agrformet.2010.01.022>
- Lei, H., Yang, D., Yang, H., Yuan, Z., & Lv, H. (2015). Simulated impacts of irrigation on evapotranspiration in a strongly exploited region: A case study of the Haihe River basin, China. *Hydrological Processes*, 29, 2704–2719. <https://doi.org/10.1002/hyp.10402>
- Li, H., Zheng, L., Lei, Y., Li, C., Liu, Z., & Zhang, S. (2008). Estimation of water consumption and crop water productivity of winter wheat in North China Plain using remote sensing technology. *Agricultural Water Management*, 95(11), 1271–1278. <https://doi.org/10.1016/j.agwat.2008.05.003>
- Li, X., Gemmer, M., Zhai, J., Liu, X., Su, B., & Wang, Y. (2013). Spatio-temporal variation of actual evapotranspiration in the Haihe River Basin of the past 50 years. *Quaternary International*, 304, 133–141. <https://doi.org/10.1016/j.quaint.2013.02.027>
- Liu, J., Zang, C., Tian, S., Liu, J., Yang, H., Jia, S., et al. (2013). Water conservancy projects in China: Achievements, challenges and way forward. *Global Environmental Change*, 23, 633–643. <https://doi.org/10.1016/j.gloenvcha.2013.02.002>
- Liu, Y., Xiao, J., Ju, W., Zhou, Y., Wang, S., & Wu, X. (2015). Water use efficiency of China's terrestrial ecosystems and responses to drought. *Scientific Reports*, 5(1), 1–12. <https://doi.org/10.1038/srep13799>
- Lu, Y., Zhang, X., Chen, S., Shao, L., & Sun, H. (2016). Changes in water use efficiency and water footprint in grain production over the past 35 years: A case study in the North China Plain. *Journal of Cleaner Production*, 116, 71–79. <https://doi.org/10.1016/j.jclepro.2016.01.008>
- Martinsen, G., Liu, S., Mo, X., & Bauer-Gottwein, P. (2019). Optimizing water resources allocation in the Haihe River Basin under groundwater sustainability constraints. *Journal of Geographical Sciences*, 29, 935–958. <https://doi.org/10.1007/s11442-019-1638-6>
- McCabe, F. M., Miralles, G. D., Holmes, R. H. T., & Fisher, B. J. (2019). Advances in the remote sensing of terrestrial evaporation. *Remote Sensing*, 11, 1138. <https://doi.org/10.3390/rs11091138>
- Mo, X., Liu, S., Lin, Z., Xu, Y., Xiang, Y., & McVicar, T. R. (2005). Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain. *Ecological Modelling*, 183(2–3), 301–322. <https://doi.org/10.1016/j.ecolmodel.2004.07.032>
- Mo, X., Chen, X., Hu, S., Liu, S., & Xia, J. (2017). Attributing regional trends of evapotranspiration and gross primary productivity with remote sensing: A case study in the North China Plain. *Hydrology and Earth System Sciences*, 21, 295–310. <https://doi.org/10.5194/hess-21-295-2017>
- Moyano, M. C., Garcia, M., Palacios-Orueta, A., Tornos, L., Fisher, J. B., Fernández, N., et al. (2018). Vegetation water use based on a thermal and optical remote sensing model in the mediterranean region of Doñana. *Remote Sensing*, 10, 1105. <https://doi.org/10.3390/rs10071105>
- Myeni, R. B., & Williams, D. L. (1994). On the relationship between FAPAR and NDVI. *Remote Sensing of Environment*, 49, 200–211. [https://doi.org/10.1016/0034-4257\(94\)90016-7](https://doi.org/10.1016/0034-4257(94)90016-7)
- Norman, J. M., Kustas, W. P., & Humes, K. S. (1995). Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, 77(3–4), 263–293. [https://doi.org/10.1016/0168-1923\(95\)02265-Y](https://doi.org/10.1016/0168-1923(95)02265-Y)
- Ozdogan, M., & Gutman, G. (2008). A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US. *Remote Sensing of Environment*, 112(9), 3520–3537. <https://doi.org/10.1016/j.rse.2008.04.010>
- Ozdogan, M., Rodell, M., Beaudoin, H. K., & Toll, D. L. (2010). Simulating the effects of irrigation over the United States in a land surface model based on satellite-derived agricultural data. *Journal of Hydrometeorology*, 11(1), 171–184. <https://doi.org/10.1175/2009JHM1116.1>
- Pan, Y., Zhang, C., Gong, H., Yeh, P. J. F., Shen, Y., Guo, Y., et al. (2017). Detection of human-induced evapotranspiration using GRACE satellite observations in the Haihe River Basin of China. *Geophysical Research Letters*, 44, 190–199. <https://doi.org/10.1002/2016GL071287>
- Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, 7(4), 811–841. <https://doi.org/10.1029/93GB02725>
- Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100(2), 81–92. [https://doi.org/10.1175/1520-0493\(1972\)100<0081:OTAOSH>2.3.CO;2](https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2)
- Qin, H., Cao, G., Kristensen, M., Refsgaard, J. C., Rasmussen, M. O., He, X., et al. (2013). Integrated hydrological modeling of the North China Plain and implications for sustainable water management. *Hydrology and Earth System Sciences*, 17, 3759–3778. <https://doi.org/10.5194/hess-17-3759-2013>
- Qin, H., Zheng, C., He, X., & Refsgaard, J. C. (2019). Analysis of water management scenarios using coupled hydrological and system dynamics modeling. *Water Resources Management*, 33, 4849–4863. <https://doi.org/10.1007/s11269-019-02410-9>
- Rockström, J., Falkenmark, M., Lannerstad, M., & Karlberg, L. (2012). The planetary water drama: Dual task of feeding humanity and curbing climate change. *Geophysical Research Letters*, 39, L15401. <https://doi.org/10.1029/2012GL051688>
- Romaguera, M., Krol, M. S., Salama, M. S., Su, Z., & Hoekstra, A. Y. (2014). Application of a remote sensing method for estimating monthly blue water evapotranspiration in irrigated agriculture. *Remote Sensing*, 6, 10,033–10,050. <https://doi.org/10.3390/rs61010033>
- Romaguera, M., Krol, M. S., Suhyb Salama, M., Hoekstra, A. Y., & Su, Z. (2012). Determining irrigated areas and quantifying blue water use in Europe using remote sensing Meteosat Second Generation (MSG) products and Global Land Data Assimilation System (GLDAS) Data. *Photogrammetric Engineering and Remote Sensing*, 78(8), 861–873. <https://doi.org/10.14358/PERS.78.8.861>
- Romaguera, M., Salama, M. S., Krol, M. S., Hoekstra, A. Y., & Su, Z. (2014). Towards the improvement of blue water evapotranspiration estimates by combining remote sensing and model simulation. *Remote Sensing*, 6, 7026–7049. <https://doi.org/10.3390/rs6087026>

- Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46, W05523. <https://doi.org/10.1029/2008WR007327>
- Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Schweppe, R., Schäfer, D., et al. (2019). Mesoscale Hydrologic Model. <https://doi.org/10.5281/zenodo.1299584>
- Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., et al. (2017). Toward seamless hydrologic predictions across spatial scales. *Hydrology and Earth System Sciences*, 21, 4323–4346. <https://doi.org/10.5194/hess-21-4323-2017>
- Schwartz, F. W., Liu, G., & Yu, Z. (2020). HESS opinions: The myth of groundwater sustainability in Asia. *Hydrology and Earth System Sciences*, 24, 489–500. <https://doi.org/10.5194/hess-24-489-2020>
- Shen, D., & Wu, J. (2017). State of the art review: Water pricing reform in China. *International Journal of Water Resources Development*, 33, 198–232. <https://doi.org/10.1080/07900627.2016.1171743>
- Shen, H., Leblanc, M., Tweed, S., & Liu, W. (2015). Groundwater depletion in the Hai River Basin, China, from in situ and GRACE observations. *Hydrological Sciences Journal*, 60, 671–687. <https://doi.org/10.1080/02626667.2014.916406>
- Shu, Y., Stisen, S., Jensen, K. H., & Sandholt, I. (2011). Estimation of regional evapotranspiration over the North China Plain using geostationary satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 13(2), 192–206. <https://doi.org/10.1016/j.jag.2010.11.002>
- Shu, Y., Villholth, K. G., Jensen, K. H., Stisen, S., & Lei, Y. (2012). Integrated hydrological modeling of the North China Plain: Options for sustainable groundwater use in the alluvial plain of Mt. Taihang. *Journal of Hydrology*, 464–465, 79–93. <https://doi.org/10.1016/j.jhydrol.2012.06.048>
- Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation—A global inventory. *Hydrology and Earth System Sciences*, 14(10), 1863–1880. <https://doi.org/10.5194/hess-14-1863-2010>
- Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., & Scanlon, B. R. (2015). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, 19, 1521–1545. <https://doi.org/10.5194/hess-19-1521-2015>
- Stisen, S., Koch, J., Sonnenborg, T. O., Refsgaard, J. C., Bircher, S., Ringgaard, R., & Jensen, K. H. (2018). Moving beyond run-off calibration—Multivariable optimization of a surface-subsurface-atmosphere model. *Hydrological Processes*, 32, 2654–2668. <https://doi.org/10.1002/hyp.13177>
- Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., et al. (2013). Ground water and climate change. *Nature Climate Change*, 3, 322–329. <https://doi.org/10.1038/nclimate1744>
- Thenkabail, P. S., Biradar, C. M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., et al. (2009). Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. *International Journal of Remote Sensing*, 30(14), 3679–3733. <https://doi.org/10.1080/01431160802698919>
- Thenkabail, P. S., Schull, M., & Tural, H. (2005). Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. *Remote Sensing of Environment*, 95(3), 317–341. <https://doi.org/10.1016/j.rse.2004.12.018>
- Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J., & Samaniego, L. (2019). The multiscale routing model mRM v1.0: simple river routing at resolutions from 1 to 50 km. *Geoscientific Model Development*, 12(6), 2501–2521. <https://doi.org/10.5194/gmd-12-2501-2019>
- Van Dijk, A. I. J. M., Schellekens, J., Yebra, M., Beck, H. E., Renzullo, L. J., Weerts, A., & Donchyts, G. (2018). Global 5 km resolution estimates of secondary evaporation including irrigation through satellite data assimilation. *Hydrology and Earth System Sciences*, 22, 4959–4980. <https://doi.org/10.5194/hess-22-4959-2018>
- Verstraeten, W. W., Veroustraete, F., Van Der Sande, C. J., Grootaers, I., & Feyen, J. (2006). Soil moisture retrieval using thermal inertia, determined with visible and thermal spaceborne data, validated for European forests. *Remote Sensing of Environment*, 101(3), 299–314. <https://doi.org/10.1016/j.rse.2005.12.016>
- Vörösmarty, C. J., & Sahagian, D. (2000). Anthropogenic disturbance of the terrestrial water cycle. *Bioscience*, 50(9), 753–765. [https://doi.org/10.1641/0006-3568\(2000\)050\[0753:ADOTTW\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2000)050[0753:ADOTTW]2.0.CO;2)
- Wisser, D., Frölking, S., Douglas, E. M., Fekete, B. M., Vörösmarty, C. J., & Schumann, A. H. (2008). Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets. *Geophysical Research Letters*, 35, L24408. <https://doi.org/10.1029/2008GL035296>
- Xiang, K., Ma, M., Liu, W., Dong, J., Zhu, X., & Yuan, W. (2019). Mapping irrigated areas of Northeast China in comparison to natural vegetation. *Remote Sensing*, 11(7), 825. <https://doi.org/10.3390/rs11070825>
- Xu, Z., Wu, Z., He, H., Wu, X., Zhou, J., Zhang, Y., & Guo, X. (2019). Evaluating the accuracy of MSWEP V2.1 and its performance for drought monitoring over mainland China. *Atmospheric Research*, 226, 17–31. <https://doi.org/10.1016/j.atmosres.2019.04.008>
- Yang, Q., Huang, X., & Tang, Q. (2020). Irrigation cooling effect on land surface temperature across China based on satellite observations. *Science of the Total Environment*, 705, 135984. <https://doi.org/10.1016/j.scitotenv.2019.135984>
- Yang, Y., Yang, Y., Moiwo, J. P., & Hu, Y. (2010). Estimation of irrigation requirement for sustainable water resources reallocation in North China. *Agricultural Water Management*, 97(11), 1711–1721. <https://doi.org/10.1016/j.agwat.2010.06.002>
- Yin, L., Feng, X., Fu, B., Chen, Y., Wang, X., & Tao, F. (2020). Irrigation water consumption of irrigated cropland and its dominant factor in China from 1982 to 2015. *Advances in Water Resources*, 143, 103661. <https://doi.org/10.1016/j.advwatres.2020.103661>
- Zaussinger, F., Dorigo, W., Gruber, A., Tarpanelli, A., Filippucci, P., & Brocca, L. (2019). Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data. *Hydrology and Earth System Sciences*, 23(2), 897–923. <https://doi.org/10.5194/hess-23-897-2019>
- Zhang, X., Qin, W., Chen, S., Shao, L., & Sun, H. (2017). Responses of yield and WUE of winter wheat to water stress during the past three decades—A case study in the North China Plain. *Agricultural Water Management*, 179, 47–54. <https://doi.org/10.1016/j.agwat.2016.05.004>
- Zhao, Q., Zhang, B., Yao, Y., Wu, W., Meng, G., & Chen, Q. (2019). Geodetic and hydrological measurements reveal the recent acceleration of groundwater depletion in North China Plain. *Journal of Hydrology*, 575, 1065–1072. <https://doi.org/10.1016/j.jhydrol.2019.06.016>
- Zheng, C., Liu, J., Cao, G., Kendy, E., Wang, H., & Jia, Y. (2010). Can China cope with its water crisis?—Perspectives from the North China Plain. *Ground Water*, 48(3), 350–354. [https://doi.org/10.1111/j.1745-6584.2010.00695\\_3.x](https://doi.org/10.1111/j.1745-6584.2010.00695_3.x)
- Zhu, X., Zhu, W., Zhang, J., & Pan, Y. (2014). Mapping irrigated areas in China from remote sensing and statistical data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 4490–4504. <https://doi.org/10.1109/JSTARS.2013.2296899>
- Zohaib, M., & Choi, M. (2020). Satellite-based global-scale irrigation water use and its contemporary trends. *Science of the Total Environment*, 714, 136719. <https://doi.org/10.1016/j.scitotenv.2020.136719>