Estimating net irrigation across the North China Plain through dual modelling of evapotranspiration.

Authors:

Julian Koch[1,*], Wenmin Zhang[2], Grith Martinsen[1], Xin He[1,3], Simon Stisen[1]

Affiliations:

1) Department of Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark

2) School of Geography, Nanjing Normal University, Nanjing, China

3) Department of Water Resources, China Institute of Water Resources and Hydropower Research, Beijing, China

Submitted to: Water Resources Research, 26/02/2020

*Corresponding author address:

Julian Koch - Geological Survey of Denmark and Greenland (GEUS), Department of Hydrology - Øster Voldgade 10, Copenhagen, Denmark - E-mail: juko@geus.dk

Key points:

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

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 brings forward a novel framework to quantify net irrigation at monthly timescale at a spatial resolution of 1 km^2 providing unprecedented spatial and temporal detail. Net irrigation refers to the evaporative loss of irrigation water. The study is conducted in the Haihe River Basin (HRB) in China encompassing the North China Plain (NCP), a global hotspot 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 mm (15.2 km^3) for NCP and 108 mm (18.6 km^3) 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 of the HRB was evaluated with GRACE data and it was found that the net irrigation estimates could close the water balance gap. Annual winter wheat classifications reveal an increasing crop area with a trend of 2200 km^2 yr^{-1}. This trend is not accompanied by a likewise increasing trend in irrigation, which suggests an increased WUE in the NCP. The proposed framework can easily be scaled up or transferred to other regions and support decision makers to tackle irrigation induced water crises and 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 water cycle, there exists limited knowledge of the applied irrigation amounts. This study brings forward a novel approach to estimate net irrigation, i.e. the evaporative loss of irrigated 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 km^2 spatial resolution. Our analysis suggest an average annual net irrigation of 108 mm (18.6 km^3). The results are evaluated against annual winter wheat classification maps as well as satellite based total water storage data (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 land produces 40% of the global food on just 20% of the total agricultural land (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 prolong 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).

Despite the tangible affect irrigation has on the freshwater resources (Döll et al., 2014), it is also considered an important anthropogenic climate forcing (Cook et al., 2015; Kang & Eltahir, 2019). Irrigation alters the water and energy exchange between land surface and atmosphere leading to a cooling of the land surface as well as increasing atmospheric water vapor that modulates cloud cover and precipitation (Kang & Eltahir, 2018). It has been shown that irrigation has regionally dampened the potential warming caused by the greenhouse gas emissions (Thiery 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, 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 U.S. 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 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 at relevant spatio-temporal scales. In recent years, the literature on this topic is 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 follows the assumption to balance available water supply with plant and atmospheric water demand and, 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, predict management decisions and 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 et al. (2020) and Kumar et al. (2015), both accounted irrigation to differences between remotely sensed SM and SM modelled by hydrologic models without irrigation schemes. Other recent work suggests 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 (~50km^2) hindered to derive irrigation amounts at relevant spatial scales for regional analysis. Further, limitations of SM were highlighted by Escorihuela & Quintana-Seguí (2016) who compared various global SM satellite products in the context of irrigation quantification and conclude that the 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 assumption such as depth of soil, water capacity of the soil layer and other empirical parameters are necessary which introduce additional uncertainties.

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 non-precipitation 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 non-precipitation sources, such as irrigation, and sinks, such as drainage in the U.S.. 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; 2014; 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 need for this hydrologic model 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.

Here we present a novel methodology designed to quantify monthly net irrigation amounts that account for the evaporative loss of irrigated water at 1 km^2 spatial resolution. 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 (mm) at a spatial scale that is relevant to regional water management. Moreover, a remote sensing based ET model can be setup to be particularly tailored to the region of interest. A key novelty of this study is that the applied hydrologic model, which is used as ET baseline without irrigation, is specifically calibrated to perform well for rainfed conditions.

The study site is the Haihe River basin (HRB) in China. The HRB comprises the North Chin Plain (NCP), which is a global hotspot 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% are sustained by groundwater abstraction (Huang et al., 2015; Pan et al., 2017; H. 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 spatio-temporal scales that are required to tackle the water crisis.

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 at 1 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 as well as with land use maps.

2. Study area and data

2.1. Haihe River Basin

The Haihe River Basin (HRB) covers an area of approximately 320,000 km^2 and encompasses mountainous regions in the west and north and lowlands in the east and south. The lowlands refer to the North China Plain (NCP), which covers approximately 140,000 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; Huanhuan Qin et al., 2019). The HRB is dominated by a monsoon climate with an average annual rainfall of around 475 mm (2002 - 2016) of which 70% to 85% occurs during the summer months (June - September). Agricultural 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 and water requirements are therefore to a large degree met by rainfall. In contrast, the winter wheat growing season spans over the dry season and crop water requirements depend heavily on irrigation. The flood irrigation technique is widely applied in NCP and typically takes place at few occasions during the winter crop season (Qin et al., 2013). The average annual cropland Normalized Difference Vegetation Index (NDVI) in Figure 1 indicates high values in the NCP as a result of the applied two crop growing season. However the spatial variability of NDVI within the NCP suggests that the two stage crop rotation system and thereby irrigation is not applied uniformly. 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.

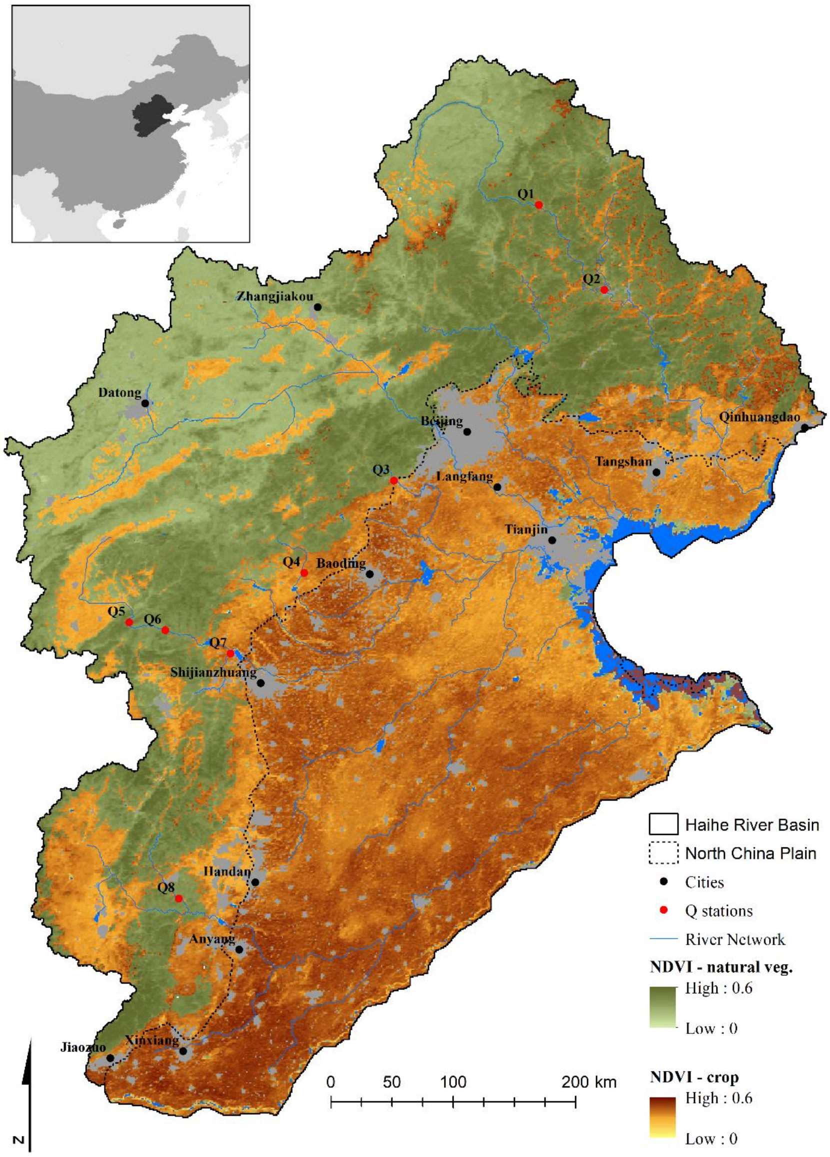

Figure 1. Map of the Haihe River basin (HRB) containing the North China Plain (NCP) domain. The depicted river network represents the natural drainage system. The map of mean annual Normalized Difference Vegetation Index (NDVI) is differentiated into cropland and natural vegetation based on a MODIS land cover classification. Based on this classification, urban areas are shown in grey, 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 grey and China in medium grey.

2.2. MODIS

For the present study, a broad range of satellite remote sensing based datasets were acquired. A prime data source were the MODIS instruments (Moderate Resolution Imaging Spectroradiometer) onboard Terra and Aqua satellites. Normalized Difference Vegetation Index (NDVI) data were obtained from the 16-day MOD13A2.006 and MYD13A2.006 products at 1km resolution. The MCD15A2H.006 product was used to acquire data on leaf area index (LAI) and fraction of absorbed photosynthetically active radiation (FAPAR) at 8-day interval and 500 m resolution. Land surface temperature (LST) datasets at daytime and nighttime were assembled from the daily 1 km products of MOD11A1.006 and MYD11A1.006. The approximate nighttime and daytime overpass times for Terra and Aqua are 11 p.m., 11 a.m. and 1 a.m. and 1 p.m., respectively. Missing nighttime LST observations have been filled using linear interpolation. Opposed to LST, emissivity was sufficient at coarser temporal resolution and therefore emissivity was acquired from the 8-day MOD11A2.006 and MYD11A2.006 products at 1 km spatial resolution. The 16-day MCD43A3.006 product was used to retrieve albedo at 500 m. An annual land cover classification at 500 m was obtained from MCD12Q1.006. If not already available at 1 km, all variables were resampled to 1km 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 datasets 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 of all variables. 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.

2.3. ERA-Interim

ERA-Interim is a global reanalysis dataset of atmospheric and land surface variables provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Data is available at 3 hourly temporal resolution at a spatial resolution of 0.75 degrees. We acquired daily shortwave downward radiation and daily mean, minimum and maximum temperature data. Furthermore, daytime LST data were processed to gap fill daily MODIS LST.

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 degree spatial resolution since April 2002. The monthly TWSA is relative to the baseline average over January 2004 to December 2009. As suggested by Landerer & Swenson (2012), the scaling coefficients were applied to the TWSA data to account for attenuated smalls scale mass variations in the 1 degree processing. For this study, GRACE data were averaged in space for the entire HRB (30 1-degree grids) as well as for the three solutions (CSR, JPL and GFZ) to a single timeseries.

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. The land surface model applied in this study only simulates natural flow conditions and therefore catchments that are excessively managed, by i.e. 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 km^2 and 18,000 km^2 (Table 1). The data coverage varied significantly among the stations where two had full timeseries of nine years with daily observations while others had only a few hundred measurements spread out over several years.

Table 1. Overview of the eight discharge stations that were applied in the model calibration. The IDs correspond to the ones in the map (Figure 1). The observartion period refers to the years were data is avaiable and the data coverage is with respect to the stated period. n is the number of daily discharge observations.

ID	Name	Area (km^2)	Longitude / Latitude	Observation Period	Data Coverage
Q1	Goutaizi (古太子)	2050	117.03 E / 41.35 N	Jan/2006 – Dec/2014	100 % (n=3287)
Q2	Sandaohezi (三道河子)	18234	117.7 E / 40.97 N	Jan/2006 – Dec/2016	17.2 % (n=691)
Q3	Zhangfang (张坊)	3041	115.68 E / 39.57 N	Jan/2006 – Dec/2014	6.3 % (n=206)
Q4	Zhongtangmei (中唐梅)	3562	114.88 E / 38.88 N	Jan/2006 – Dec/2014	13.9 % (n=455)
Q5	Jishengqiao (济胜桥)	11874	113.06 E / 38.38 N	Jan/2006 – Dec/2010	100 % (n=1826)
Q6	Xiao Jue (小觉)	14051	113.43 E / 38.23 N	Jan/2006 – Dec/2010	100 % (n=1826)
Q7	Pingshan (平山)	6268	114.12 E / 38.15 N	Jan/2006 – Dec/2010	99.9 % (n=1824)
Q8	Kuangmenkou (狂门口)	4932	113.47 E / 38.15 N	Jan/2006 – Dec/2014	100 % (n=3287)

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 October 1 and June 30 in the following 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 and split randomly into equal training and validation subsets. The overall accuracy of the binary maps (winter wheat and non-winter wheat) varied slight across the years with an average of around 98%.

3. Methods

3.1. Remote sensing evapotranspiration model

In this study, we apply the PT-JPL model to estimate daily actual evapotranspiration (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 land surface temperature (LST) as a proxy for the soil moisture 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). Subsequently the potential levels are reduced to their actual levels using various constraints. The constraints reflect the plant physiological status and soil moisture 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):

$$ET = ET_c + ET_s. \qquad \text{(eq.1)}$$

Canopy transpiration is calculated based on three physiological constraints:

$$ET_c = f_g \cdot f_T \cdot f_M \cdot PET_c, \qquad \text{(eq.2)}$$

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 soil moisture constraint (f_{SM}):

$$ET_s = f_{SM} \cdot PET_{SM}. \qquad \text{(eq.3)}$$

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.

Table 2. Equations used to calculate the biophysical constraints for the PT-JPL model. f_{APAR} is the fraction of absorbed photosincally active radiation, f_{IPAR} is the fraction of intercepted photosincally active radiation, calculated by the NDVI relationship proposed by Myneni & 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 95th percentile in this study, ATI is the apparent thermal inertia index and in this study, ATI_{max} and ATI_{min} related to the 95th and 5th percentiles, respectively.

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)

3.2. Hydrologic model

This study applies version 5.9 of the multiscale Hydrologic Model (mHM). mHM is a grid based spatially distributed hydrologic model that accounts for key hydrologic processes and includes a routing scheme (Samaniego et al., 2010). A multi-parameter 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 model, forcing data is resampled to 4 km resolution, 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 5 km scale during calibration and at 1 km for a final production run after calibration. Effective parameters at the modelling 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. 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 as computed by PT-JPL. Soil texture data was obtained by 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

A multi-objective and multi-variable calibration framework has been designed to yield robust model performance during rainfed conditions. ET simulated by the rainfed hydrologic model without irrigation scheme is used as baseline in the net irrigation estimation. For this purpose, land cover was classified into natural vegetation, rainfed cropland and irrigated cropland. Initially, the MODIS land cover product was used to differentiate between cropland and natural vegetation. 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 0.075 per month 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 affected by 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. The concentrated patch of rainfed crops in the center of NCP coincides well with very sandy soils, which may explain the absence of irrigation in this area. Observations of the two most important water balance outflows, namely discharge (Q) and ET were considered. The observed ET data, obtained from the PT-JPL model covering the years 2002 to 2016, has been used in several ways. For the three land cover classes, rainfed cropland, all cropland and natural vegetation the monthly MAE was calculated. The MAE of ET over natural vegetation was used as calibration target for all months throughout the years. The MAE associated to rainfed cropland was utilized during the winter wheat growing season (October - May) and the MAE 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 this calibration design, the hydrologic model was calibrated exclusively against ET under rainfed conditions to minimize the influence of irrigation. In order to target the calibration on the spatial pattern performance, the multi-component Spatial Efficiency

(SPAEF: Koch et al. (2018)) metric was applied to the multi-year average monthly simulated and observed ET maps. SPAEF was used to assess the simulated spatial patterns under rainfed conditions for the months March to October. The ET patterns during the four winter months 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 applied on 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. mHM incorporates a LAI driven scaling function to estimate a PET multiplier, in similar fashion to the well-known crop coefficient, and the MODIS based LAI data had to be corrected to remove the effect of irrigation. This was achieved by reducing LAI of irrigated cropland, using the above-mentioned NDVI based mask, in the months from October to May, to the average LAI of rainfed cropland of that particular month. A global optimizer scheme within PEST (Doherty, 2005) that is based on a covariance matrix adaptation estimation strategy (CMA-ES) was applied to calibrate mHM parameters. For Q, the mean absolute error (MAE) for each of the eight stations (Figure 1) was used as objective function. Achieving the best possible accuracy of Q dynamics is not at the center of this study and therefore a simple water balance objective function, as the MAE, has been applied. The set of objective functions was weighted as follows, 40 % was allocated to the MAE for the 8 discharge stations, 20 % to the SPAEF applied to eight mean monthly ET maps (March - October), 20 % to the MAE of all cropland ET (June – September, 15 years), 10 % to the MAE of the rainfed cropland (October – May, 15 years) and 10 % to the MAE of natural vegetation (January – December, 15 years). The weighting has been implemented with respect to the residuals as obtained from the initial parameter set. Net irrigation estimation.

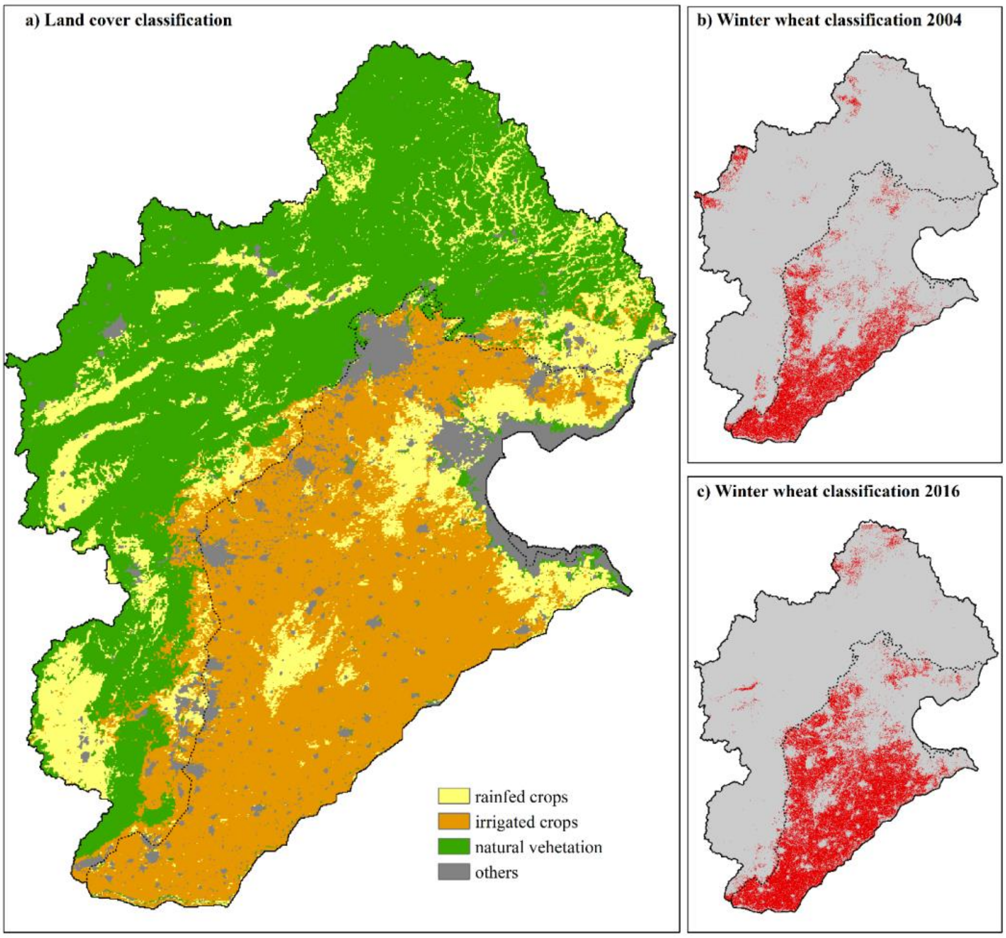

Figure 2. The map in a) depicts the land cover classification applied in the ET calibration to differentiate between irrigated cropland, rainfed cropland and natural vegetation. Two examples of the winter wheat classification maps by Zheng et al. (2020) are shown for 2004 (b)) and 2016 (c)). NCP domain indicated with dashed line.

Net irrigation ($netIrr$) amounts are quantified at monthly time scale at 1 km^2 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 irrigated water. With the absence of irrigation in the hydrologic model it can be assumed that mHM systematically underestimates ET at times of irrigated crop growth as compared to PT-JPL:

$$netIrr = ET_{PT-JPL} - ET_{mHM}. \qquad \text{(eq.4)}$$

Negative $netIrr$ estimations, caused by to overestimations of ET_{mHM}, are conceivable, especially in periods of high precipitation. Therefore, we investigate two hypotheses to quantify net irrigation. The first (h1), neglects negative residuals in equation 4 whereas the second (h2) takes both, positive and negative residuals into consideration. h2 can be considered a conservative estimate of irrigation. In case the hydrologic model overestimates ET with respect to the remote sensing based model, h2 can yield unrealistic negative irrigation amounts.

Nevertheless, h2 is included to shed light on some of the uncertainties related to an approach based on the residual 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-June), whereas the latter covers the summer crops (June-September).

4. Results

4.1. Remote sensing model

The PT-JPL model was used to estimate daily ET at 1 km^2 spatial resolution across the entire HRB from 2002 to 2016. As stated in the method section, most of the MODIS derived inputs to the PT-JPL model underwent a processing step using the multi-year 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. Afterwards 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 intra-year variability.

Daytime LST data from three different sources, namely MODIS Terra, MODIS Aqua and ECMWF ERA-Interim were acquired to calculate the soil moisture constraint in the PT-JPL model. All of the above use the same MODIS based nighttime LST data to calculate the apparent thermal inertia (Table 2). 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. 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. Overall, the bias between PT-JPL forced by Terra LST and Aqua LST is 0.02 mm d^{-1} using only grids with coinciding observations. The bias between ECMWF derived ET and Terra and Aqua is 0.03 mm d^{-1} and 0.04 mm d^{-1}, respectively.

The resulting daily ET dataset is a combination of the three PT-JPL models forced with the above-mentioned LST datasets. For the final dataset, 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 non-precipitation source of water.

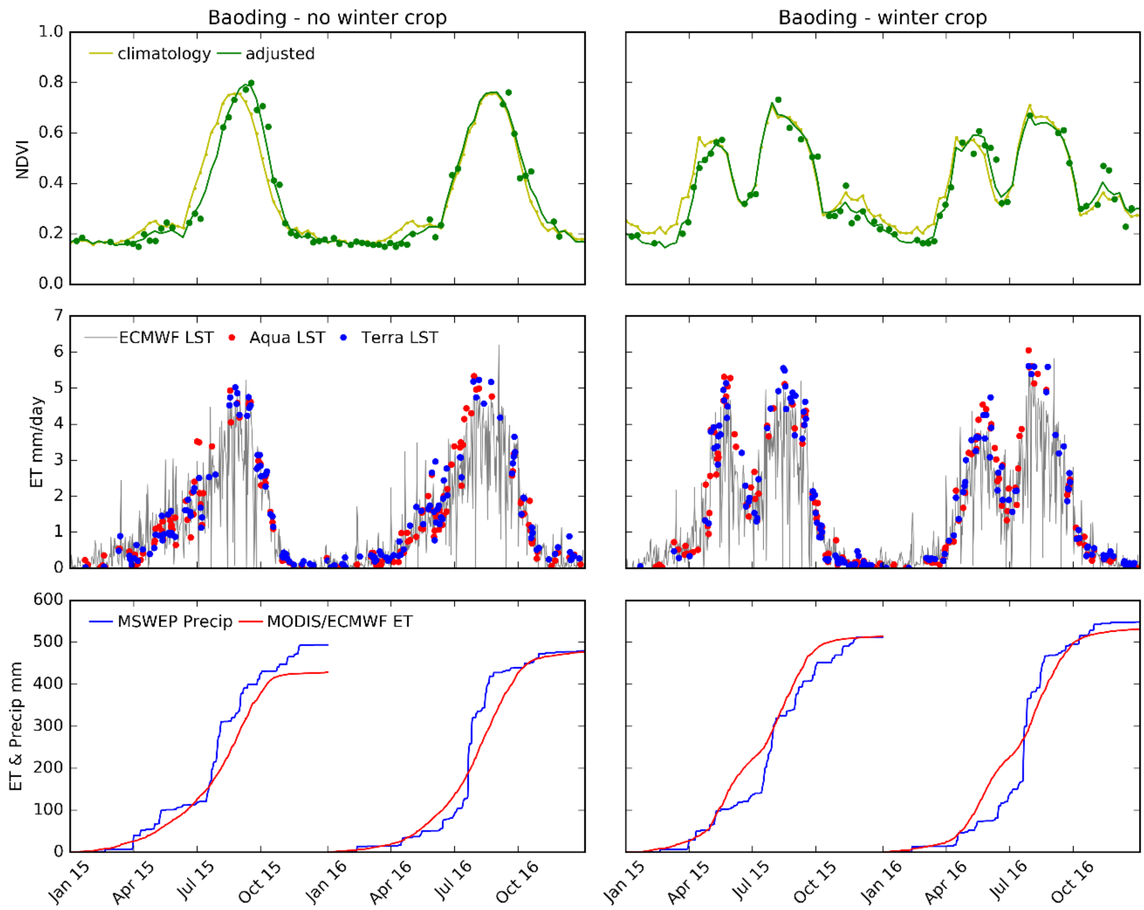

Figure 3. Example timeseries for two 1 km^2 grid cells over two years (2015 - 2016) close to the city of Baoding, located in NCP. The first depicts an example of rainfed crops without winter crop in the left column and the second showcases the crop rotation system of irrigated winter crop and rainfed summer crop in the right column. The first row shows the multi-year NDVI climatology which is adjusted to capture actual observations (green points) while providing a robust interpolation on days of missing data. The middle row contains ET based on the PT-JPL model using land surface temperature (LST) from Aqua, Terra and ECMWF. The cumulative density functions of precipitation and ET are shown in the bottom row. In case of ET, data obtained with ECMWF LST was used to gap fill days without MODIS observations.

Figure 4 depicts the average annual remote sensing based ET pattern for the HRB for the years 2002 to 2016 at 1 km2 spatial resolution. The average annual ET for the HRB and NCP domains are 483 and 511 mm/year, respectively. The highest ET fluxes are found in the mountainous regions north and west of the NCP that are covered by forest. The spatial pattern of ET in the NCP does to a large degree reflect patterns of agriculture. ET is generally high in the southern part of the NCP towards 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 Mountains. Low ET corresponds well to the areas of rainfed cropland as classified in Figure 2 and areas of low NDVI in Figure 1.

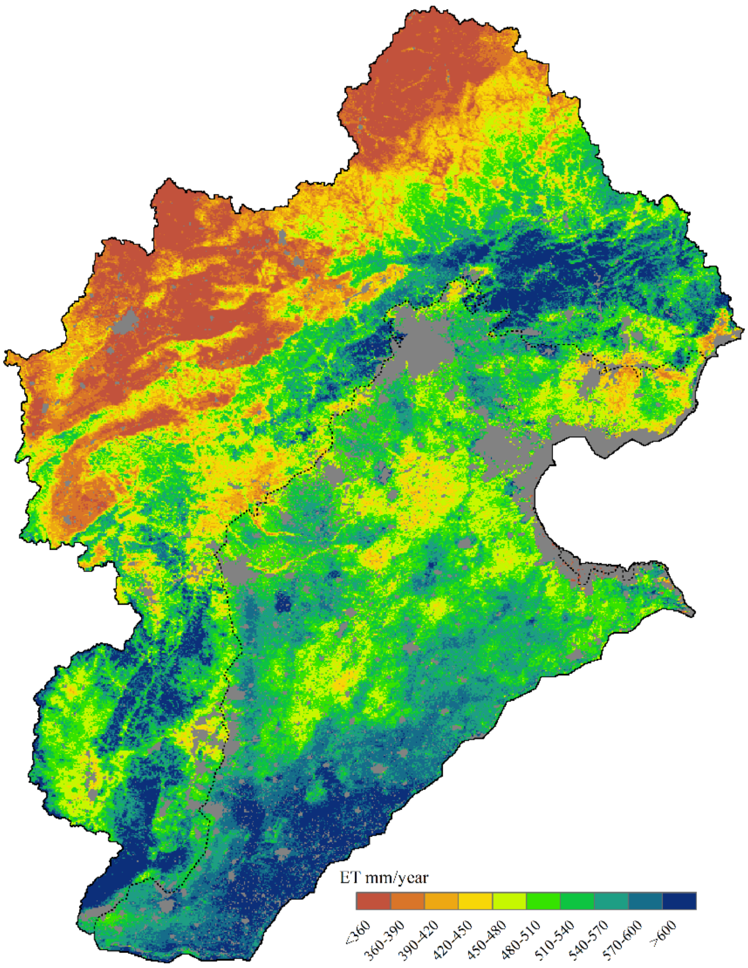

Figure 4. Multi-year average (2002-2016) of remote sensing based ET using the PT-JPL model. NCP domain indicated with dashed line.

4.2. Hydrologic model

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. The natural ET variability is driven by climate seasonality with an annual range from 10 mm/month in winter to 100 mm/month in summer, which is represented accordingly by mHM. 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 rainfed cropland was 9.6 mm for the summer months and 10.6 mm for the winter months. The MAE for natural vegetation was 11.6 mm. 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.

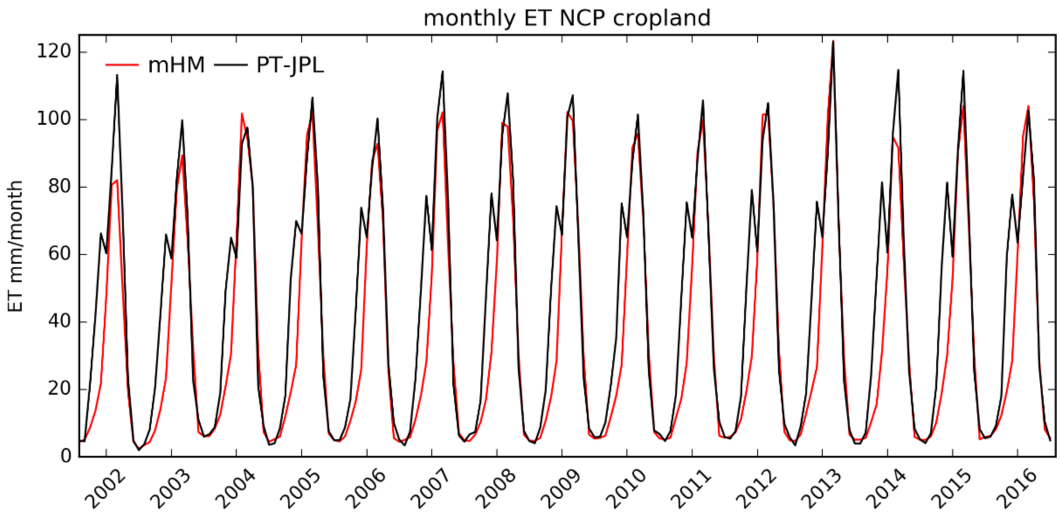

Figure 5. Timeseries of monthly ET obtained for all cropland (irrigated and rainfed) in NCP by the remote sensing model (PT-JPL) and the calibrated hydrologic model (mHM).

The MAE at the eight discharge stations was used as calibration target and the resulting Q performance is stated in Table 3. Q data was included in the calibration to get the overall waterbalance in place. Nevertheless, some stations are associated with large errors that may be a result of anthropogenic interference and uncertainties in the precipitation forcing or Q observations.

Table 3. Discharge performance obtained through calibration. 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.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
observed (m^3/s)	1.6	12.9	6.1	6.0	2.0	8.7	7.7	2.0
observed (mm/year)	24.0	22.2	63.5	53.4	5.4	19.4	38.5	12.5
residual (m^3/s)	-1.2	0.3	-4.9	-4.5	4.5	-1.6	-0.4	3.4
residual (% of observed)	-79	2	-80	-74	221	-18	-5	174

Table 4 provides an overview of the annual water balance components of NCP and HRB. 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, given that not all available water is likely to evaporate, but will also generate recharge and runoff, this is a strong indicator for the extensive irrigation scheme. For the HRB, discharge amounts to approximately 10 % of precipitation and recharge constitutes 8.5 %. 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.

Table 4. Overview over annual and average water balance components for the NCP and the HRB. All values are stated in mm yr^{-1}. Precipitation was obtained from MSWEP v2. Simulated discharge (surface runoff), groundwater recharge (percolation from bottom soil layer) and potential ET (PET) were taken from the calibrated mHM model. ET is given for the remote sensing model (PT-JPL) and calibrated hydrologic model (mHM). Irrigation was calculated based on the two hypothesis defined in section 3.3.

NCP	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	avg
Precipitation	331	647	544	490	432	474	495	548	527	526	552	553	416	560	644	516
Discharge	15	94	121	75	50	49	61	67	90	73	104	111	36	77	160	79
Recharge	18	79	100	64	45	45	55	60	72	64	83	89	35	64	123	66
PET	728	694	727	751	729	765	748	754	736	732	744	763	725	733	715	736
ET PT-JPL	485	478	496	514	500	528	527	517	498	505	519	519	527	528	527	511
ET mHM	332	373	452	430	401	413	426	435	419	414	444	456	397	416	442	417
Irrigation h1	169	132	92	120	125	145	128	117	112	124	117	112	156	140	125	128
Irrigation h2	154	105	44	84	99	115	101	82	78	91	76	63	130	112	84	94
HRB	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	avg
Precipitation	358	554	480	442	404	440	473	450	492	472	508	520	416	511	623	476
Discharge	13	56	69	44	32	29	37	37	50	44	63	68	27	43	103	48
Recharge	13	48	56	37	27	26	33	32	41	38	50	55	23	37	80	40
PET	901	880	910	935	927	959	952	948	939	949	940	967	927	925	925	932
ET PT-JPL	468	458	463	480	466	501	498	488	484	480	483	492	496	495	495	483
ET mHM	349	383	442	418	392	396	429	400	414	419	432	466	403	413	480	416

Irrigation h1	139	111	72	99	100	133	103	120	109	103	96	82	125	116	83	106
Irrigation h2	119	76	21	62	74	105	69	89	70	61	50	26	92	81	15	67

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. Table 4 contains the results for annual estimated irrigation amounts for h1, which neglects any negative residuals (ET overestimations by mHM) whereas h2 takes both, positive and negative residuals into consideration. Logically, annual net irrigation based on h1 is larger than h2 with an average of 126 mm (15.2 km^3) for NCP and 108 mm for HRB (18.6 km^3).

Figure 6 illustrates the spatial pattern of mean annual net irrigation. Irrigation 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 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 6, 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 were 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 hypothesis is investigated in Figure 6 to analyze the uncertainties related to neglecting negative ET residuals in h1. The uncertainties are large along the southern and northeastern boundary 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.

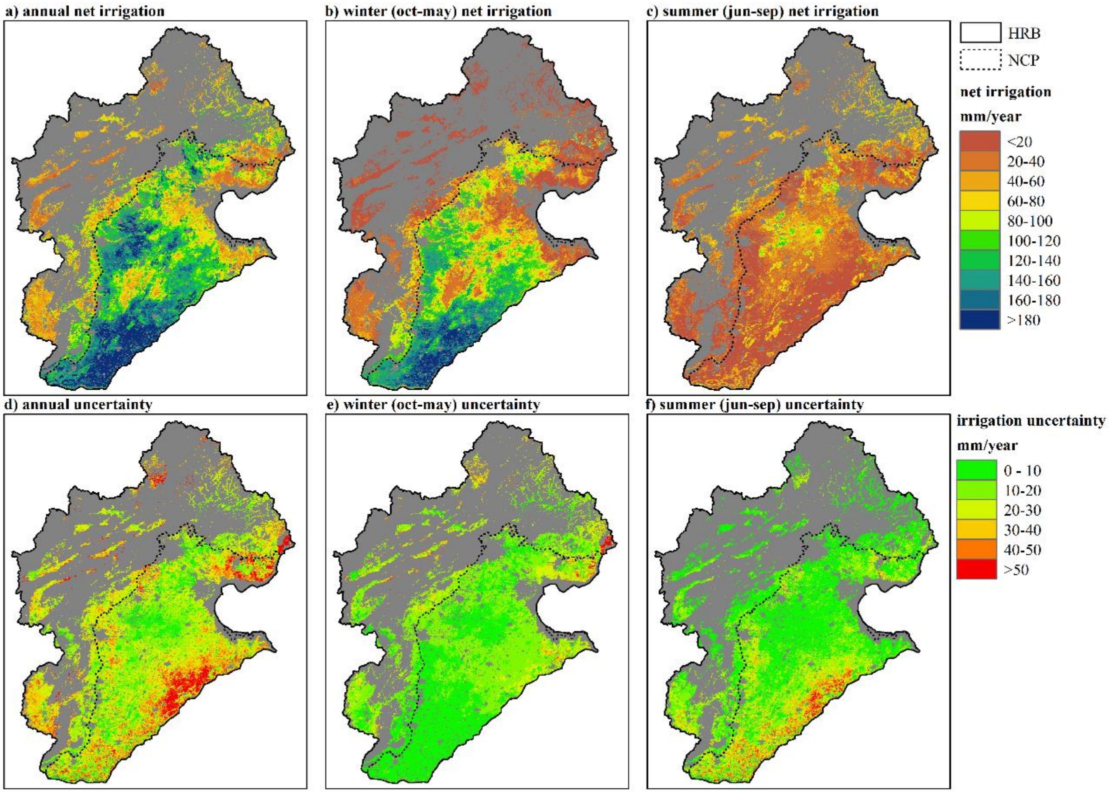

Figure 6. a) Average annual net irrigation (2002-2016) as obtained from h1 (section 3.3). b) and c) split a) into a winter- and a summer fraction, respectively. d), e) and f) depict the differences between the two hypotheses (h1 - h2) to investigate uncertainties of a), b) and c), respectively.

Figure 7 further investigates the inter-annual 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 mm (10.7 km^3) and 168.8 mm (20.3 km^3). 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 (i.e. 2004, 2012, 2013 and 2016) have an overall negative summer irrigation as consequence of a systematic overestimation of ET in mHM. These results are unrealistic and emphasize the larger uncertainties related to the summertime irrigation quantification, when precipitation is high, in comparison to the wintertime assessment, when precipitation is low.

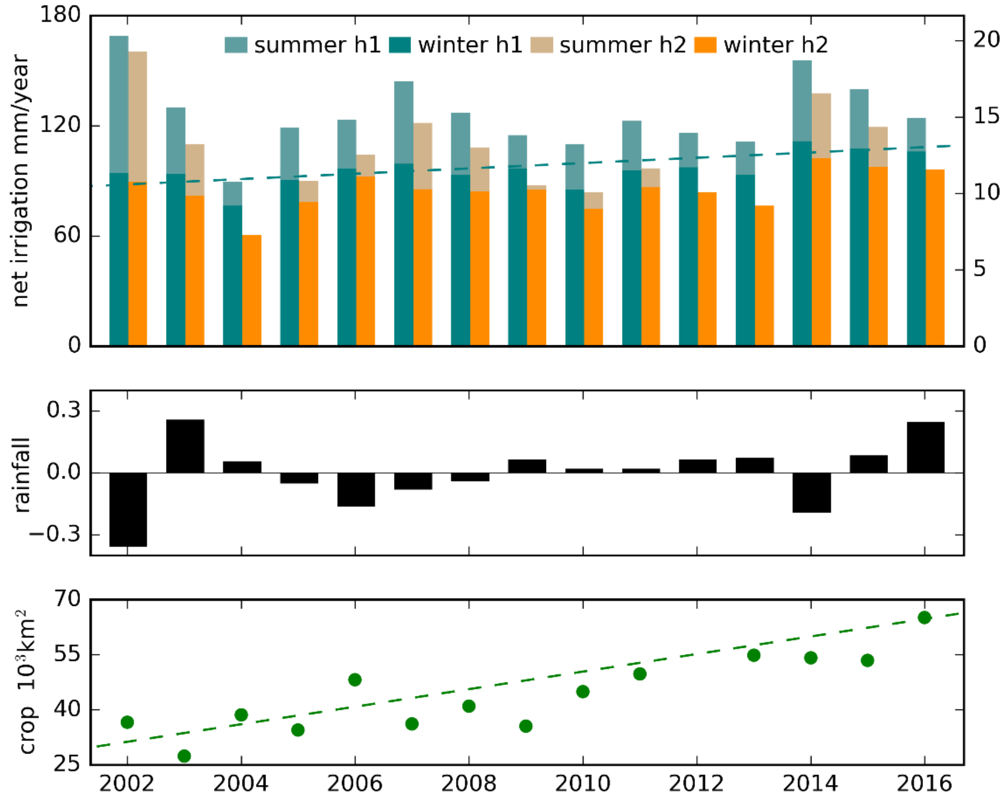

Figure 7. Top row: Annual analysis of net irrigation based on the two hypothesis (h1 and h2) split up into a winter- and a summer fraction. The dashed line indicates the fitted linear trend of winter irrigation based on h1. Middle row: Annual precipitation anomalies with respect to the 2002 to 2016 mean. Bottom row: Annual winter wheat area, as classified by Zhang et al. (2020), with fitted linear trend (dashed line).

4.4. Net irrigation evaluation

Figure 7 also contains the development of winter wheat cultivation areas in NCP, which is characterized by a clear increasing trend of 2200 km^2 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. The detailed winter wheat classification maps are a valuable source to evaluate the net irrigation estimates spatially. Figure 8 depicts the winter wheat classification for the years 2004 and 2016, as already shown in Figure 2, but zoomed into NCP. Based on the selected years, the area used for winter wheat expanded from approximately 38,000 km^2 to 65,000 km^2 which marks an increase of 70 %. The continuous winter net irrigation estimates are classified into three classes for better visual comparison with the winter wheat classification. Overall, the spatial expansion is well represented between the two approaches. Winter wheat and thereby irrigation expands drastically in the Eastern and Northern part of NCP and in general the cropping area becomes more compact. Based on the three selected thresholds in Figure 8, greater than 50 mm yr^{-1}, 100 mm yr^{-1} and 150 mm yr^{-1}, the irrigated areas expand from 83,000 km^2, 39,000 km^2 and 12,000 km^2 in 2004 to 98,000 km^2, 66,000 km^2, 35,000 km^2 in 2016, which marks an increase of 18 %, 69 %, 190 %, respectively. The best agreement with the winter wheat classification maps is obtained with the second threshold, namely, greater than 100 mm yr^{-1}.

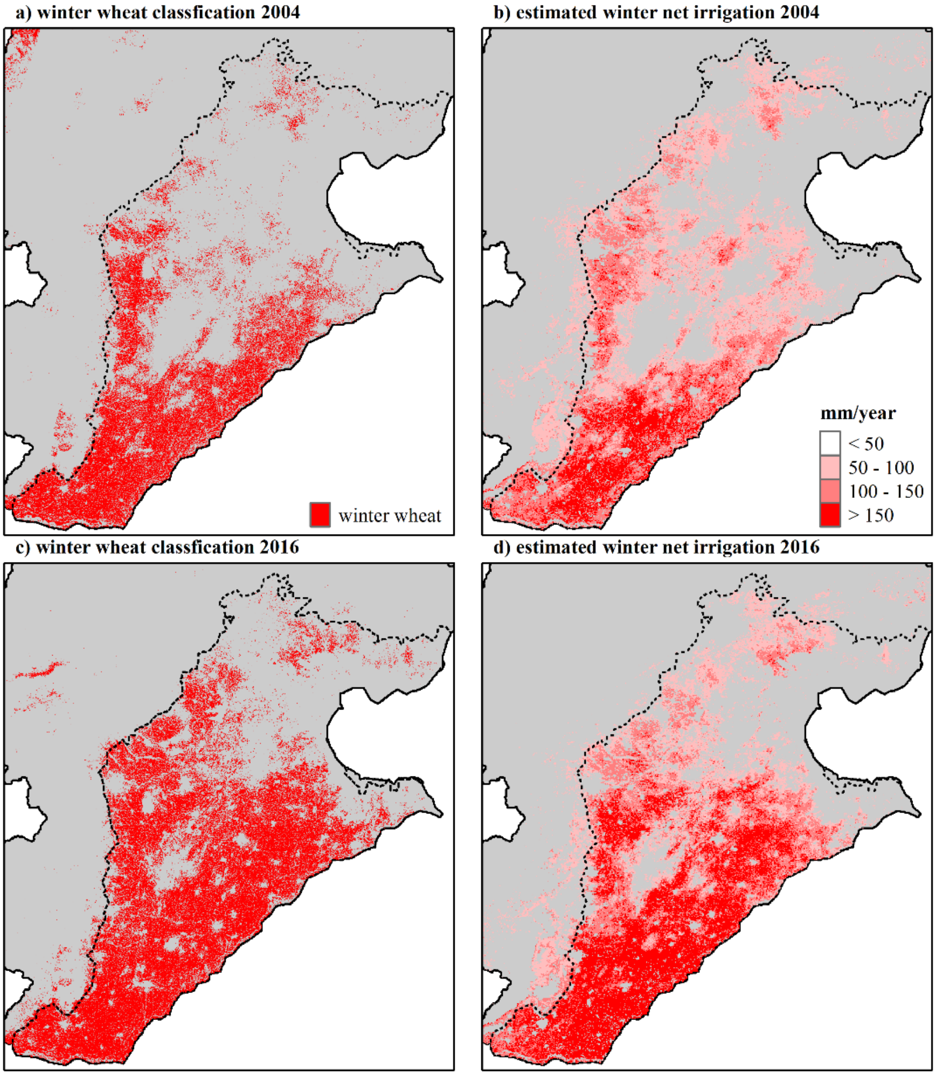

Figure 8. Spatial evaluation of the net irrigation estimates for NCP. The winter wheat classification is shown as reference for the years 2004 and 2016 in panels a) and c), respectively. The estimated winter irrigation is classified in four classes for the same years (panels b) and d)) to be comparable with the binary winter wheat map.

The GRACE based total water storage anomalies (TWSA) clearly support the observed groundwater depletion of the NCP with a decreasing trend of -9.0 mm yr^{-1} (Figure 9). 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. The storage anomalies of mHM possess a slight positive trend (3.6 mm yr^{-1}). Thereby, mHM does not follow the observed GRACE signal, which constitutes that the negative trend in the GRACE data cannot be attributed to climate variability. 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 underline the effect of extensive irrigation activities. Both have a clear increasing trend in the summer monsoon months from June to August driven by high precipitation. The TWSA data disagrees most 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. 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 TMSC 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.

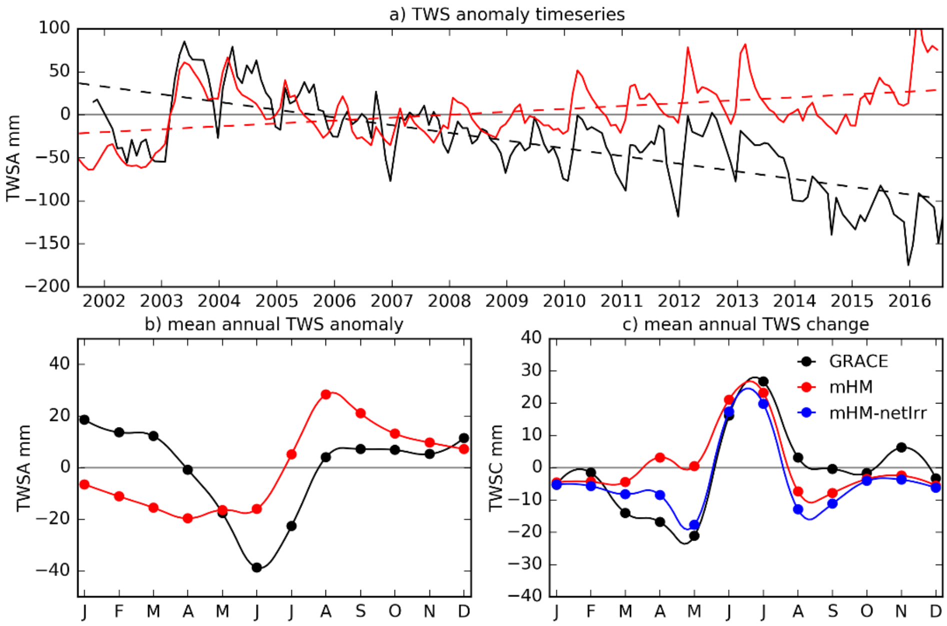

Figure 9. The monthly total water storage anomalies (TWSA) based on GRACE and mHM for the entire HRB are shown in a) with their respective fitted linear trends as dashed lines. The average TWSA are shown in b) based on the trend-removed monthly data in a). The average total water storage change (TWSC) based on the data in a) is illustrated in c) including a scenario where mHM was corrected for net irrigation (netIrr).

5. Discussion

5.1. Irrigation in the North China Plain

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 are not always trivial due to deviating study periods, spatial resolutions that vary 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 actual evaporative loss. Yang et al. (2010) applied agronomic crop models and reported an overall irrigation water requirement of 16.5 km^3 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 around 17.7 $km^3\ yr^{-1}$ by Shen et al. (2015). Hu et al. (2016) estimated average annual irrigation based as the residual term of the soil water balance equation which amounted to 317 mm yr^{-1} with irrigation to ET rations 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 soil moisture 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 mm yr^{-1} which was compared to three land surface models (GLADAS) 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 amounts and frequencies were prescribed based on literature. Annual NCP irrigation was estimated to be around 180 mm yr^{-1}, which is slightly higher than our estimates. For a similar model setup, irrigation was reported to be 290 mm yr^{-1} for an irrigation district within NCP (Shu et al., 2012). The two above-mentioned 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 hydrologic models. Based on the literature review, our irrigation estimates at 1 km^2 provide critical information at an unprecedented spatial and temporal resolution, which 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 have 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 (Q. 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 winter wheat classification we could draw the conclusion that the irrigation water use efficiency (WUE) must have improved in NCP since the early 2000s. Fang et al. (2020) found a significant trend in winter wheat ET of 1.28 mm yr^{-1} due to anthropogenic influence, which can be supported by our irrigation results as seen by the trend line in Figure 7. However, this trend does not correspond to the doubling in winter wheat cultivation area, which implies 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. 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 modelling 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 semi-arid conditions (Fisher et al., 2008; García et al., 2013; McCabe et al., 2019). In principle, different ET models are available, such as GLEAM, ALEXI or MOD16, and future research should utilize ensembles of remote sensing based ET data to investigate uncertainties related to the irrigation quantifications. Based on PT-JPL, ET was estimated to be 511 mm yr^{-1} for NCP and 483 mm yr^{-1} for HRB. Based on various approaches, annual ET rates ranging from 480 to 600 mm yr^{-1} have been reported in the literature (Guo & Shen, 2015; Hu et al., 2016; H. Li et al., 2008; X. Li et al., 2013; X. Mo et al., 2005; Pan et al., 2017; H. 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 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 2). Moreover we expect that the effect of a potential bias in the ET dataset to quantify net irrigation will be diminished by calibrating the hydrologic model against the ET data during rainfed conditions.

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 does not 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 mm yr^{-1} (Pan et al., 2017), which is in good agreement with the 3.6 mm yr^{-1} predicted by mHM. In future research, an ensemble of precipitation forcing could be utilized to quantify the uncertainty of the irrigation quantification. 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.

Comparing the two hypotheses to quantify net irrigation 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. This finding is important to take into consideration for transferring the proposed method to other regions.

The proposed approach estimates net irrigation, i.e. the evaporative loss of irrigated 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 km^3 yr^{-1}. The advantage of estimating net estimation is that uncertain assumptions on return flows are not required.

5.3. Irrigation management

The central government of China has from the 1950's 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. In 2010, the No. 1 Central Document for 2011 (Ministry of Agriculture of the People's Republic of China, 2010), issued by the Ministry of Agriculture, laid out an ambitious plan about the 'most stringent water management' to achieve sustainable use of water resources and promote water savings. The following year the Three Red Lines (Global Water Partnership, 2015), defined national targets for capping water use, increasing WUE and reducing water pollution. 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. More than 12,000 researcher-led demonstrations of soil- and crop management improvements were carried out across China and subsidies of around 1.5 billion renminbi were given to soil-testing of farm land in 2012 (Zhang et al., 2013). Since 2011 there has been a focus on water pricing reforms to promote water savings in the agricultural sector. 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 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 Haihe River Basin (HRB), encompassing the North China Plain (NCP), based on dual modelling of evapotranspiration (ET). The systematic differences between a rainfed hydrological model and a remote sensing based model of ET provide realistic irrigation estimates at unprecedented spatio-temporal 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.
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.
3. Annual net irrigation is estimated to be 128 mm and 106 mm for NCP and HRB, respectively, which constitutes approximately 25 % of ET.
4. Summer irrigation is more sensitive to inter-annual precipitation variability, while winter irrigation is less affected.
5. GRACE based total water storage data underline the plausibility of the quantified irrigation amounts.
6. Evaluation of winter irrigation coverage and amounts implies increasing areas under irrigation accompanied by an increase in water use efficiency.

Acknowledgments and data availability

The work has been carried out in connection with the project ”Managed Aquifer Recharge in the North China Plain” with financial support from the Danish Ministry of Foreign Affairs. The project grant (17-M08-GEU) is administrated and approved by the Danida Fellowship Center on behalf of the Danish Ministry of Foreign Affairs. Wenmin Zhang acknowledges the project grant of the Chinese National Postdoctoral Program for Innovative Talents (BX20190154). Xin He's work has been supported by the National Natural Science Foundation of China (NSFC), grant number 41701042. We would like to thank Monica Garcia and Gorka Mendiguren Gonzalez for sharing their PT-JPL code with us. Part of discharge data were provided by Dr. C. Davidsen who collected the data when he pursued his PhD degree at Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. We acknowledge ECMWF to provide the ERA-Interim data, NASA to provide MODIS products and Princeton Climate Analytics to share MSWEP v2 precipitation data. The monthly ET datasets used to conduct the analysis are shared by the authors and made available via the Pangaea database (https://issues.pangaea.de/browse/PDI-23229).

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*(11), 3670–3690. https://doi.org/10.1029/2019MS001797

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* . 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*. 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*. 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*. 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*. 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*. 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*. 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/https://doi.org/10.1016/j.rse.2019.111400

Demirel, Mehmet 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*. 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. In *Agricultural Water Management*. https://doi.org/10.1016/j.agwat.2005.07.021

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*. https://doi.org/10.5194/hess-22-4959-2018

Doherty, J. (2005). *PEST: Model Independent Parameter Estimation. Fifth Edition of User Manual*. Brisbane: Watermark Numerical Computing.

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T., & Eicker, A. (2014). Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resources Research*. https://doi.org/10.1002/2014WR015595

Escorihuela, M. J., & Quintana-Seguí, P. (2016). Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes. *Remote Sensing of Environment*. 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*. 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/https://doi.org/10.1016/j.agrformet.2019.107844

FAO, & IIASA. (2009). *Harmonized world soil database*. *Food and Agriculture Organization*. https://doi.org/3123

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*. https://doi.org/10.1029/2005RG000183

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*(23), 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*. 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*. https://doi.org/10.1038/nature10452

García, M., Sandholt, I., Ceccato, P., Ridler, M., Mougin, E., Kergoat, L., et al. (2013). Actual evapotranspiration in drylands derived from in-situ and satellite data: Assessing biophysical constraints. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2012.12.016

Global Water Partnership. (2015). *China's water resources management challenge: The "three red lines."*

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*. 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*(3), 1070–1086. https://doi.org/10.1175/JHM-D-14-0017.1

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*. 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*. 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*. 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*(1), 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*. https://doi.org/10.1029/2019GL082396

Koch, J., Cüneyd Demirel, M., & Stisen, S. (2018). The SPAtial EFficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models. *Geoscientific Model Development*, *11*(5). 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*. https://doi.org/10.5194/hess-19-4463-2015

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*. https://doi.org/10.1029/2011WR011453

Lawston, P. M., Santanello, J. A., Zaitchik, B. F., & Rodell, M. (2015). Impact of irrigation methods on land surface model spinup and initialization of WRF forecasts. *Journal of Hydrometeorology*. 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*. 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*. 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*. 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*. 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*. https://doi.org/10.1016/j.gloenvcha.2013.02.002

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*. 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*. 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* . https://doi.org/10.3390/rs11091138

Ministry of Agriculture of the People's Republic of China. (2010). No. 1 Central Document for 2011.

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*. https://doi.org/10.1016/j.ecolmodel.2004.07.032

Mo, Xingguo, 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*. 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*. https://doi.org/10.3390/rs10071105

Myneni, R. B., & Williams, D. L. (1994). On the relationship between FAPAR and NDVI. *Remote Sensing of Environment*. 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

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*. https://doi.org/10.1016/j.rse.2008.04.010

Ozdogan, M., Rodell, M., Beaudoing, 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. 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*. 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*. 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

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*. https://doi.org/10.5194/hess-17-3759-2013

Qin, Huanhuan, Zheng, C., He, X., & Refsgaard, J. C. (2019). Analysis of Water Management Scenarios Using Coupled Hydrological and System Dynamics Modeling. *Water Resources Management*, *33*(14), 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*. https://doi.org/10.1029/2012GL051688

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*. https://doi.org/10.14358/PERS.78.8.861

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*. https://doi.org/10.3390/rs61010033

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*. 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*. https://doi.org/10.1029/2008wr007327

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*. 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*(1), 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*(2), 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*. 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*. 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*. 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*. 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*. https://doi.org/10.5194/hess-19-1521-2015

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*. https://doi.org/10.1038/nclimate1744

Thenkabail, P. S., Schull, M., & Turral, 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/https://doi.org/10.1016/j.rse.2004.12.018

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*. https://doi.org/10.1080/01431160802698919

Thiery, W., Visser, A. J., Fischer, E. M., Hauser, M., Hirsch, A. L., Lawrence, D. M., et al. (2020). Warming of hot extremes alleviated by expanding irrigation. *Nature Communications*, *11*(1), 290. https://doi.org/10.1038/s41467-019-14075-4

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*. 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

Wisser, D., Frolking, 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*. 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* . 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*. 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/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*. https://doi.org/10.1016/j.agwat.2010.06.002

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*. https://doi.org/10.5194/hess-23-897-2019

Zhang, F., Chen, X., & Vitousek, P. (2013, May). Chinese agriculture: An experiment for the world. *Nature*. Nature Publishing Group. https://doi.org/10.1038/497033a

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*. 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*. 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*. 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*(11), 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/https://doi.org/10.1016/j.scitotenv.2020.136719