Spring 2016

# Global and regional assessments of unsustainable groundwater use in irrigated agriculture

Danielle S. Grogan
*University of New Hampshire, Durham*, danielle.grogan@unh.edu

GLOBAL AND REGIONAL ASSESSMENTS OF UNSUSTAINABLE GROUNDWATER

USE IN IRRIGATED AGRICULTURE

BY

DANIELLE SARAH GROGAN

B.A. Smith College, 2009

M.S. Brown University, 2011

DISSERTATION

Submitted to the University of New Hampshire

In Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

In

Earth and Environmental Sciences

May, 2016

ProQuest Number: 10117670
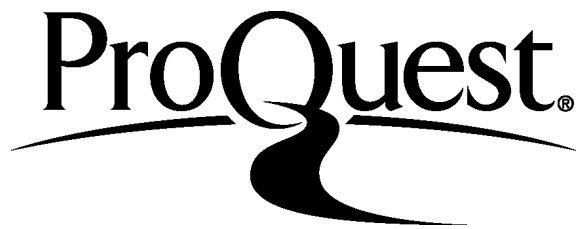
ProQuest 10117670

This dissertation has been examined and approved in partial fulfillment of the requirements for

the degree of Doctor of Philosophy in Earth and Environmental Sciences by:


Dissertation Co-Director, Dr. Stephen E. Frolking, Resaerch Professor

Earth Systems Research Center


Dissertation Co-Director, Dr. Richard B. Lammers, Research Assistant Professor

Earth Systems Research Center


Dr. Karen Fisher-Vanden

Professor of Environmental and Resource Economics, The Pennsylvania State University


Dr. Dominik Wisser

Senior Scientist at the Center for Development Research (ZEF), University of Bonn and

Affiliated Research Assistant Professor, University of New Hampshire


Dr. Wilfred M. Wollheim, Assistant Professor

Department of Natural Resources & The Environment


On April 5, 2016


Original approval signatures are on file with the University of New Hampshire Graduate School.

**Acknowledgements**

First and foremost I would like to thank my advisors Steve Frolking and Richard Lammers for their support and guidance, and most of all for their unwavering encouragement. They have shared not only their expertise, but their enthusiasm for scientific research. One of my greatest aspirations is to one day be the kind of advisor that they have been to me. I am also grateful to my committee members: Karen Fisher-Vanden, Dominik Wisser, and Wil Wollheim. They have each contributed to this dissertation, often pushing me to develop either a greater breadth or depth to each chapter.

I want to thank everyone in the Water Systems Analysis Group. In particular the technical expertise of Alex Prusevich and Stanley Glidden has been invaluable. In addition to my fellow students in the Water Systems Analysis Group, I have been lucky enough to work with Esha Zaveri, a graduate student at the Pennsylvania State University. Esha was brave enough to embark on a fully cross-disciplinary research project with me, and I am grateful for all she has taught me, as well as her remarkable perseverance.

Finally, I'd like to thank my family: my parents, sisters, and wonderful husband. The love, support, and confidence they have given me has made my work towards a PhD not only a success, but a joy.

TABLE OF CONTENTS

ABSTRACT

GLOBAL AND REGIONAL ASSESSMENTS OF UNSUSTAINABLE GROUNDWATER

USE IN IRRIGATED AGRICULTURE

by

Danielle Sarah Grogan

University of New Hampshire, May, 2016


Groundwater is an essential input to agriculture world-wide, but it is clear that current rates of groundwater use are unsustainable in the long term. This dissertation assesses both current use of groundwater for country- to global-scale agriculture, and looks at the future of groundwater. The focus is on 1) quantifying food directly produced as a result of groundwater use across spatially-varying agricultural systems, 2) projecting future groundwater demands with consideration of climate change and human decision-making, and 3) understanding the system dynamics of groundwater re-use through surface water systems. All three are addressed using a process-based model designed to simulate both natural and human-impacted water systems.

Irrigation can significantly increase crop production. Chapter 1 combines a hydrology model (WBM) with a crop model to quantify current crop production that is directly attributed to groundwater irrigation in China. Unsustainably-sourced groundwater – defined as groundwater extracted in excess of recharge – accounted for a quarter of China's crop production, and had significant spatial variability. Climate variability and groundwater demand magnified one another in hot and dry years, causing increased irrigation demand at the same time as limited surface water supplies.

Human decisions about water resource management can impact both the demand and sustainability of groundwater use.  Chapter 2 takes an interdisciplinary approach to projecting India's future (to 2050) groundwater demands, combining hydrology and econometric modeling. The econometric model projects how humans make decisions to expand or contract the irrigated land area of crops in response to climate change.  Even in areas with precipitation increases, human decisions to expand irrigated areas led to increasing demands for groundwater.  We additionally assessed the potential impact of a large water infrastructure project to alleviate groundwater demands in India, and found that maximum alleviation (up to 16%) was dependent upon the storage volume and location of new reservoirs.

One proposed method for reducing the world's demand for groundwater is to increase the efficiency of agricultural water use.  However, these same inefficiencies cause a portion of extracted groundwater to enter surface water systems; it can then be reused, creating a complex system in which groundwater demand does not linearly decline with increased water use efficiency. Chapter 3 quantifies the amount of groundwater that enters surface water systems, the number of times this water is reused for agriculture, and the minimum amount of groundwater required by current agricultural systems in the hypothetical scenario of perfect irrigation efficiency.

INTRODUCTION

Will there be enough food to support earth's growing population?  This is a very basic question, but answering it requires understanding both the natural and human-controlled systems that underlie global agriculture.  Water resources is one of these key systems: agricultural production places high demands on the freshwater supply, and these demands will only increase as the population grows (Vörösmarty et al., 2000a; Postel 2000; Foley et al., 2011).  Currently, 40% of global food production comes from irrigated agriculture (Abdulluah, 2006), which accounts for ~70% of all freshwater withdrawals (Shiklomanov, 2000; Cai and Rosegrant, 2002).  Up to one third of these withdrawals are considered nonrenewable or unsustainable (Vörösmarty et al., 2005; Wisser et al., 2010; Wada et al., 2012), indicating that under current agricultural practices, water availability will not be able to keep up with increasing demands.  Most of these unsustainable withdrawals are from groundwater (Aeschbach-Hertig and Gleeson, 2012), and many of the world's major groundwater aquifers are rapidly depleting due to unsustainable groundwater pumping (Aeschbach-Hertig and Gleeson, 2012; Rodell et al., 2009; Gleeson et al., 2012; Wada et al 2012).

Unsustainable groundwater is defined broadly as groundwater extracted in excess of recharge (Aeschbach-Hertig and Gleeson, 2012; Rodell et al., 2009; Gleeson et al., 2012; Wada et al 2012).  While several estimates of global groundwater extraction and depletion have been made (e.g., Döll and Siebert, 2002; Wisser et al., 2010; Wada et al., 2012), fewer analyses have linked these water extractions directly to the production of food (Siebert and Döll, 2010).  Tying irrigation water to food production is important because different combinations of crops, climate,

soil conditions, and management lead to significantly different crop water use productivities (Cai al., 2011) – i.e., the amount of crop biomass or food calories produced by a unit of water can vary widely.

Key considerations when studying global food security and water resources are: current agricultural production, arable land, freshwater availability, population growth, and climate change. India and China are excellent focal points for groundwater and agriculture assessments because they represent extreme cases for all these key considerations. Currently, China and India are the top agricultural producers in the world, accounting for 19% and 10% of total global food production, respectively, and are also the most populous countries in the world (FAOSTAT). Asian countries have the lowest ratio of arable land to population at 0.13 hectares per person (United Nations Population Division), with little room to increase agricultural area. This means that to produce more food to keep up with its growing population, Asia must rely on increasing the productivity of existing cropland. Water is the ultimate constraint on food production in both India and India (Cheng et al., 2009; Amarasinghe et al., 2005; Rosegrant et al., 2002), and the Upper Ganges aquifer in India and the North China Plain aquifer are two of the most rapidly-depleting groundwater reservoirs in the world (Aeschbach-Hertig and Gleeson, 2012; Gleeson et al., 2012).

Human management of water resources has the potential to both exacerbate and alleviate unsustainable groundwater demand. For example, poor management of water delivery through canal systems paired with government subsidies for well installation and pumping costs in India (Badiani, Jessoe and Plant, 2012) have lead some farmers to favor (unsustainable) groundwater use over surface water use for irrigation even when surface water is available (Minor Irrigation Census, 1993). On the other hand, constructed infrastructure such as large reservoirs, small

tanks and ponds, and canals all work to bolster the supply of surface water. Globally, large dams constructed in the last century have created ~7,000 km$^3$ of surface water storage (ICOLD, 2011), and small tanks and ponds have significant potential to increase water availability for crop production (Wisser et al, 2010). Faced with declining groundwater resources and growing populations, many countries are planning large-scale water transfer projects (Ghassemi and White, 2007; Adhikari et al., 2009). India has begun construction on its National River Linking Project, which plans to transport 178 km$^3$ yr$^{-1}$ of water by connecting 37 rivers and building ~3,000 storage dams (Amarasinghe, Shah and Malik, 2009); China similarly has plans to transport 45 km$^3$ yr$^{-1}$ of water from the water-rich southern China Yangtze river basin to the arid northeast (Zhang, 2009). Increasing irrigation efficiency – i.e., decreasing runoff, recharge, and non-beneficial evaporation from irrigation systems – has also been proposed as a potential solution to groundwater stress (Gleick, 2001; Wada, 2012; Simmons et al., 2015). However, several studies have shown that irrigation efficiency improvements can lead to unintended increased water use (Contor and Taylor, 2013; Ward and Pulido-Velazquez, 2008) due to "Jevon's Paradox" which states that as technology leads to improved resource use efficiency, people respond by increasing consumption of the resource (Sorrell, 2009).

This dissertation addresses two major questions:

1) How much food production is directly attributable to irrigation with unsustainable groundwater in China and India?

2) What is the potential for human actions – through farmer-level decision making, large-scale infrastructure development, and irrigation efficiency increases – to alter groundwater demands for agriculture?

Chapters 1 and 2 address the first question. Chapter 1 uses a combined hydrology and crop modeling approach to assess how unsustainable groundwater contributes to China's current agricultural production. Chapter 2 quantifies the impact of removing unsustainable groundwater as an irrigation water source on India's current agricultural production. Question 2 is addressed in both Chapter 2 and Chapter 3. In Chapter 2, India's future unsustainable groundwater demand is projected in a combined hydrology-econometrics modeling system that represents not only the physical system, but also human decision making. Chapter 2 contributes to the emerging field of socio-hydrology (Baldassare et al., 2013), which focuses on the interactions between people and water resources, including not only large-scale infrastructure projects and policy but also farmer-level decisions. Chapter 2 additionally assesses the ability of different implementations of India's National River Linking Project to alleviate unsustainable groundwater demand. Lastly, Chapter 3 addresses Question 2 by assessing the groundwater implications of increasing irrigation efficiency on a global scale. Chapter 3 aims to understand the system dynamics of groundwater use and reuse, and does not project future demands but rather contributes new understanding to current global agricultural water use.

Each chapter was written in the format of individual, peer-reviewed journal articles. As of this date (March 2016), Chapter 1 is published in the journal Science of the Total Environment with coauthors Fan Zhang, Alexander Prusevich, Richard Lammers, Dominik Wisser, Stanley Glidden, Changsheng Li and Steve Frolking. Chapter 2 is in review at the

journal Environmental Research Letters with coauthors Esha Zaveri (equal first authorship), Karen Fisher-Vanden, Steve Frolking, Richard Lammers, Douglas Wrenn, Alexander Prusevich and Robert Nicholas. Chapter 3 will be submitted shortly to the journal Science Advances with coauthors Dominik Wisser, Alexander Prusevich Richard Lammers, and Steve Frolking. The citations for these papers are:

1. Grogan D S, Zhang F, Prusevich A, Lammers R B, Wisser D, Glidden S, Li C, and Frolking S (2015) Quantifying the link between crop production and mined groundwater irrigation in China *Sci. Total Environ.* **511** 161-75.

2. Zaveri E, Grogan D S, Fisher-Vanden K, Frolking S, Lammers R B, Wrenn D H, Prusevich A, and Nicholas R E (2016) Inivisible water, visible impact: Groundwater use and Indian agriculture under climate change. *Environmental Research Letters, in review.*

3. Grogan D S, Wisser D, Prusevich A, Lammers R B, Frolking S (2016) The use and reuse of unsustainable groundwater: A global budget. *In prep.*

CHAPTER I:

QUANTIFYING THE LINK BETWEEN CROP PRODUCTION AND MINED

GROUNDWATER IRRIGATION IN CHINA

**Abstract**

In response to increasing demand for food, Chinese agriculture has both expanded and

intensified over the past several decades.  Irrigation has played a key role in increasing crop

production, and groundwater is now an important source of irrigation water.  Groundwater

abstraction in excess of recharge (i.e., groundwater mining) has resulted in declining

groundwater levels and could eventually restrict groundwater availability.  In this study we used

a hydrological model, WBMplus, in conjunction with a process based crop growth model,

DNDC, to evaluate Chinese agriculture's recent dependence upon mined groundwater, and to

quantify mined groundwater-dependent crop production across a domain that includes variation

in climate, crop choice, and management practices.  This methodology allowed for the direct

attribution of crop production to irrigation water from rivers and reservoirs, shallow (renewable)

groundwater, and mined groundwater.  Simulating 20 years of weather variability and circa year

2000 crop areas, we found that mined groundwater fulfilled 20% - 49% of gross irrigation water

demand, assuming all demand was met.  Mined groundwater accounted for 15% - 27% of

national total crop production. There was high spatial variability across China in irrigation water

demand and crop production derived from mined groundwater.  We find that climate variability

and mined groundwater demand do not operate independently, but rather magnify one another in

hot and dry years with increased irrigation demand and limited surface water supply for irrigation.

## Introduction

Increasing global demand for food over the past several decades has forced agriculture to expand into water-scarce regions and increase irrigation water use substantially (Molden et al., 2007). With little additional land available for agricultural expansion except in tropical rainforests, future increases in crop production will likely rely on increases in irrigation and intensification, both in China and globally (Molden et a., 2007). Historically, China's agriculture was concentrated in the wetter southern half of the country, but significant expansion over the past 50 years has led to over 50% of current national crop production occurring in the dry northern regions (Ma, 2006). Irrigated agriculture has expanded significantly in China in the past 75 years, increasing by more than 35 million hectares since 1939 to 51 Mha of planted land and 79 Mha of harvested land in 2000 (Calow et al., 2009; Portmann et al., 2010). Groundwater exploitation has underpinned the agricultural intensification of northern China since the 1990's, where groundwater accounts for up to 40% of irrigation water (Wada et al. 2012). Declining groundwater levels are threatening to limit the irrigation water supply for China's crop production (Kang et al., 2009; Aeschbach-Hertig and Gleeson, 2012; Syed et al., 2008). There has been a 15 m drop in groundwater levels in the North China Plain since 1960 (Calow et al., 2009), and the current rate of groundwater depletion across China is approximately 1m per year (Aeschbach-Hertig and Gleeson, 2012). This heavy reliance on groundwater for irrigation is driven largely by lack of sufficient surface water supplies (Wisser et al., 2008; Wada et al., 2012), and Northern China is now considered to be a region of physical water scarcity, i.e., more

than 75% of river discharge is abstracted (Molden et al., 2007). Global multi-model projections of irrigation water availability show significant reduction in Northern China for irrigation potential from renewable surface water by 2100 due to climate change (using a scenario of high greenhouse gas emissions (RCP8.5) (Elliott et al., 2013). Aquifer depletion could also significantly decrease irrigation water availability in the future. Despite the importance of groundwater and groundwater depletion for the future of Chinese agriculture, it is currently unknown how much food is produced as a direct result of irrigation with non-renewable groundwater mining.

Large-scale surface water balance models can simulate the use of both surface water and groundwater for irrigation. Several model-based estimates of irrigation water demand in China have been made, ranging from 220 to 850 $km^3$ $yr^{-1}$, circa 2000 (Wisser et al., 2008), with most estimates in the range of 350 – 500 $km^3$ $yr^{-1}$ (Döll and Siebert, 2002; Siebert and Döll, 2007; Liu and Yang, 2010; Wada et al., 2012). The proportion of irrigation water demand fulfilled by mined groundwater pumping is less well constrained. Groundwater (both renewable and mined in excess of recharge) provides up to 40% of China's irrigation water, and model results from Wada et al. (2012) show that 20 $km^3$ $yr^{-1}$ (5% of irrigation demand) is drawn from non-renewable groundwater.

Water supply alone does not determine food production; cropped areas, crop choice, soil quality, and management practices all contribute (Tilman et al., 2002; Foley et al., 2012). Agriculture's vulnerability to changes in water supply will necessarily also depend on these factors, which all vary spatially across China. Crop water productivity, the crop yield gained from one unit of water, also varies spatially, even within individual watersheds (Cai et al., 2011). Global-scale studies of unsustainable water supplies and food production have used an empirical

method to determine the relationship between water use and crop yields (Siebert and Döll, 2010), but this method is not suited for higher-resolution analysis (Siebert and Döll, 2010). An alternative approach is to employ a process-based crop growth model that can capture these important spatially-variable production factors (Müller et al, 2013).

In this study, we used a hydrological model in conjunction with a process based crop growth model to evaluate Chinese agriculture's dependence upon mined groundwater, and quantify mined groundwater-dependent crop production across a domain that includes variation in climate, crop choice, and management practices. We computed irrigation water demand, as well as sources of water supply, across 20 years of climate variability. Using these two models allowed us to quantify the amount of food produced as a direct result of irrigation with groundwater mined in excess of recharge. We also defined an index of vulnerability to loss of mined groundwater resources that is a function of the amount of mined groundwater required for irrigation and the productivity of a crop irrigated from that water source.

## **Methods**

We used two models to simulate irrigation water demand, irrigated and rainfed crop yields, and crop production due to mined groundwater. A grid-based water balance model (WBMplus, Wisser et al., 2010) calculated daily fluxes and storage of water between and within different water storage pools (Figure 1.1). WBMplus was used to estimate the irrigation water demand of different crop types based on weather variables, soil properties, and crop parameters, and tracked the sources of irrigation water available to meet that demand (e.g., Wisser et al., 2008). DNDC (Li et al., 1992; 2007), a process-based crop growth and agroecosystem biogeochemistry model, was used to simulate fully-irrigated and rainfed crop yields for individual crops and multi-

cropping systems for all counties in China.  WBMplus provided an estimate of the irrigated crop

area dependent upon mined groundwater.  By applying DNDC's difference between irrigated

crop yields and rainfed crop yields to these areas, we estimated the portion of total crop yields

directly resulting from groundwater mining.


**Water Balance Model**

WBMplus computed a daily water balance for each 0.5 degree grid cell.  Water was input

through precipitation and irrigation, and outputs were evapotranspiration, runoff, and shallow

rechargeable groundwater (Figure 1.1). Water was stored as soil moisture and in the shallow

groundwater pool. Surface runoff and baseflow from shallow groundwater were transported

downstream through the STN-30p river network (Vörösmarty et al., 2000b) taking into account

the storage of water in large reservoirs. A detailed description of WBMplus's fundamental

processes is given by Wisser et al. (2010).  Here we describe WBMplus's method of irrigation

and crop water use, and WBMplus's updated method of simulating operation of large reservoirs.


Irrigation and crop ET

Irrigation was simulated by abstracting water from rivers, reservoirs, and groundwater,

then moving that water to the soil water pool. The amount of water abstracted was determined by

an irrigation water demand that was based on root zone soil moisture, crop evapotranspiration,

and irrigation efficiency. Evapotranspiration depended on each crop's planting date, growing

season length, growth stages, rooting depth, and a crop water use coefficient. Crop coefficients,

growth stages, rooting depths, and depletion factors were from Siebert and Döll (2010), and the

coefficient method used was based on FAO recommendations (Allen, 1998). Each crop's evapotranspiration was calculated as:

$$E_c = K_c ET_0 \qquad (1.1)$$

Where $E_c$ (mm day$^{-1}$) is the crop's evapotranspiration, $Kc$ (-) is a dimensionless crop- dependent and growth stage-dependent coefficient for crop $c$, and $ET_0$ (mm day$^{-1}$) is a reference-surface potential evapotranspiration. $ET_0$ was calculated using the Hamon method (Hamon, 1963; Vörösmarty et al., 1998).

Crop evapotranspiration and percolation removed water from the soil moisture pool. When water inputs to the soil moisture pool brought the soil moisture stock above field capacity, the excess was diverted in equal proportions to surface runoff and percolation. Percolated water moved to the shallow groundwater storage pool (implemented as a linear reservoir); water from the groundwater storage pool flowed to the river network as baseflow. A daily soil moisture accounting was done in irrigated areas with inputs from precipitation, irrigation water, and snowmelt, and with evapotranspiration and percolation as outputs from the soil storage. If soil moisture fell below a crop-dependent threshold, $C_t$ (mm), irrigation water was applied to bring the soil moisture up to field capacity. The crop-dependent threshold for soil moisture was calculated as:

$$C_t = C_s * R_d (F_{cap} - W_{pu}) \qquad (1.2)$$

where $C_s$ (-) is a crop-specific scalar that represents a crop's inability to remove all water from the soil, $R_d$ (mm) is the crop's rooting depth, $F_{cap}$ (-) is the field capacity of the soil, and $W_{pt}$ (-) is the wilting point of the soil. Crop-specific parameters $C_s$ and$R_d$ were from Siebert and Döll (2010), and $F_{cap}$ and $W_{pt}$ are from FAO/UNESCO (2003).

When soil moisture was below the crop-specific threshold, the crop had a positive

irrigation demand (otherwise demand was zero). For each grid cell, a net irrigation demand, $I_{net}$

(mm day$^{-1}$), was calculated daily as the area-weighted water demand of all the crops (26 crop

types; taken from Portmann et al., 2010) in the grid cell. Daily net irrigation demand was

calculated as:

$$I_{net} = \sum_{c=1}^{n} A_i * a_i * I_c \qquad (1.3)$$

where $A_I$ (-) is the portion of the grid cell's area equipped for irrigation, $a_c$ (-) is the portion of the

irrigated area containing crop $c$ (Table 1.1), $I_c$ (mm day$^{-1}$) is the irrigation demand of crop $c$, and

$n$ is the number of crops.

Rice paddies require additional irrigation water due to inundation-induced percolation.

We assumed that irrigation water was applied to rice paddies to maintain a 50 mm flooding depth

throughout the growing season.  To achieve this, irrigation water was applied on the first day of

the paddy rice-growing season to fill the soil moisture pool 50 mm above field capacity, and

each subsequent day water was applied to account for percolation plus evapotranspiration losses

minus precipitation gains.  Rice paddy percolation was assumed to occur at a constant rate that is

determined by the soil drainage class (e.g., Wisser et al., 2008), estimated spatially from the

FAO/UNESCO soil map of the world (FAO/UNESCO, 2002).

In each grid cell, water for irrigation could be withdrawn from large reservoir storage, if

present, rivers flowing through the grid cell, shallow renewable groundwater, and mined

groundwater (modeled as a distinct pool of water, Figure 1.1).  These water sources defined the

water volume available for irrigation.  WBMplus applied sufficient irrigation water to each grid

cell to fulfill the irrigation demand, bringing the soil in the grid cell's irrigated area up to field

capacity. This water was first removed from the (renewable) shallow groundwater pool, then

from the river discharge and reservoirs. If these water sources were not sufficient to fulfill

demand, then mined groundwater was added from a limitless pool to make up the difference.

While the shallow groundwater pool was recharged through infiltration, the mined groundwater

pool received no recharge. The mined groundwater pool represented groundwater abstractions in

excess of recharge. We assumed that all irrigation demand was met, though in practice this may

not always be the case.

Delivery of water from an irrigation water source to the irrigated field is inefficient. An

efficiency factor, $E_{eff}$ (0.34 for China (Döll and Siebert, 2002), was applied to account for these

losses. Gross irrigation demand, $I_{gross}$ (mm day$^{-1}$) was:

$$I_{gross} = \frac{I_{net}}{E_{eff}} \tag{1.4}$$

Only the efficient portion of irrigation water (i.e., the water equal to irrigation demand) was

added to the soil moisture pool. The daily "inefficient" water ($I_{gross}$–$I_{net}$) was split three ways and

returned to other pools or fluxes in each grid cell: evaporation, groundwater water recharge, and

surface runoff. The amount of inefficient water that evaporated, $I_{evap}$, is:

$$I_{evap} = \begin{cases} PET - AET & if \left(I_{gross} - I_{net}\right) \geq (PET - AET) \\ I_{gross} - I_{net} & if \left(I_{gross} - I_{net}\right) < (PET - AET) \end{cases} \tag{1.5}$$

If ($I_{gross}$–$I_{net}$) >PET–AET, then the remaining inefficient water was divided evenly between

groundwater recharge and river discharge.

The sources of irrigation water, both the efficient and inefficient portions, were recorded

by the model. We assumed that the contribution of each water source to the total irrigation water

amount was the same for both the efficient and the inefficient portions.

*Large reservoirs*

Large reservoirs were represented in the model as river grid cells with large river water storage capacity.. Unlike other river grid cells, the rules for outgoing discharge were governed by reservoir operation rules. The model applied the same set of rules to all reservoirs. If the reservoir storage was below 80% full, then water was released as the log of the reservoir level, and if the reservoir storage was above 80% full, then water was released as the exponent of the reservoir level; release at 80% capacity was equal to annual mean river discharge in that grid cell.  Minimum allowed reservoir release was 20% of the 5-year average annual discharge (Prusevich et al., 2013).


### Crop Yield Model

The DeNitrification-DeComposition (DNDC) model, a process-based model of carbon and nitrogen biogeochemistry in agroecosystems, simulated irrigated and rainfed crop yields (kg C/ha/yr). The model simulates soil temperature and moisture regimes, soil carbon and nitrogen dynamics, crop growth and yield, nitrogen leaching, and emissions of trace gases for both individual crops and multi-cropping systems, and requires as inputs soil properties, daily weather, and crop management details (e.g., crop type and rotation, fertilization, irrigation, tillage, planting and harvest dates) (Li, 2007a). DNDC has been used extensively for studies of both Chinese and global agriculture (e.g., Zhang et al., 2002; Wang et al., 2008; Qiu et al., 2009; Li et al., 2010; Deng et al., 2011; Han et al., 2014).

Unlike the gridded spatial structure of WBMplus, DNDC modeled crop growth on a county-based (polygon) system. The model was run from 1981 to 2000 to capture 20 years of weather-driven crop yield and irrigation water demand variability for each county in China.

NASA's MERRA climate reanalysis product (Rienecker et al., 2011) was used for the temperature and precipitation inputs to DNDC, matching each county to the MERRA grid cell closest to the county polygon center.  The average county size in China is similar to the size of a MERRA grid cell (~2,500 km2). Data on crops grown in single- and multiple-crop rotations for each of China's ~2400 counties came from Qiu et al. (2003).  County soil properties used by DNDC (texture, bulk density, pH and carbon content) were from digitization of the Chinese Third National Soil Survey maps (Shi et al., 2004; Tang et al., 2006).  All crops in all counties were simulated with both full irrigation (no water stress) and no irrigation (rainfed) for all 20 years, to capture weather-driven interannual variability in rainfed and irrigated crop yields.  For the simulations presented here, to represent general changes in fertilization across China, we applied a constant increase (~2% yr$^{-1}$) in crop-specific fertilizer application rates across China from 1981 to 2000, and a step decrease (36%) in manure application rate in 1990 (*NBS,* 2008). The crop distribution represented circa 2000 conditions, and simulated crop yields were de-trended with respect to increases in fertilizer and decrease in manure application rates to achieve c.2000 crop yield rates for all 20 years of the simulation, so yield interannual variability resulted from weather effects only.  DNDC planting and harvest dates are from Cui et al. (1994), and crop-specific growth parameters are reported in Li (2007b).

*Data*

Both WBMplus and DNDC used daily climate drivers of temperature and precipitation from the MERRA NASA reanalysis product, years 1981 - 2000.  Soil and non-crop vegetation data were from the UNESCO/FAO soil map of the world (FAO/UNESCO, 2003), and the

GRanD database of global dams and reservoirs (Lehner et al., 2011) was used for the location and storage capacity of ~7000 large reservoirs globally.

The MIRCA2000 gridded crop maps and crop calendars were used to determine crop areas, planting month, and harvest months in WBMplus for each of 26 crop classes, and up to 4 subcrop types (Portmann et al., 2010). Crop class coefficients (Kc), rooting depths (Rd), and proportional lengths of growing stages were from Siebert and Döll (2010). Crop calendars (for planting and harvesting dates) were downscaled from monthly to daily values by assuming mid-month planting and harvesting dates.

*Analysis*

While the Qiu et al. (2003) and MIRCA2000 crop distributions for China had substantial overlap, there were enough differences that two data analysis steps were necessary to harmonize the DNDC model output with the WBMplus output and the MIRCA2000 crop maps. First, county-based DNDC yield data were gridded to the resolution of the WBMplus model (0.5 degree) assigning a crop yield value to each grid cell equal to the area-weighted yields of all counties that overlapped with the grid cell. Second, for each grid cell, the individual crop yields were aggregated to produce one average annual irrigated crop yield value and one average rainfed crop yield value, both in units of $kgCha^{-1}yr^{-1}$ (note that multiple crop yields in a single year were aggregated to a single annual average crop yield). All crop production results are therefore based on the composition of irrigated and rainfed crops in DNDC's crop maps, as opposed to MIRCA2000 crop maps. This aggregation allowed for a comparison of land under irrigated versus rainfed cropping.

16

Both models were run for 20-year periods, 1981 - 2000, to capture impacts of variability in temperature and precipitation on irrigated and rainfed crop yields and irrigation water demand. The crop areas, irrigated areas, and growing season were static, representing year 2000 agriculture (Portmann et al., 2008; Qiu et al., 2003), and all crop yields were de-trended to year-2000 values with respect to time-varying DNDC model inputs (fertilizer and manure). All time series results therefore show variability in water demand and crop yields as a function of climate, and are not meant to be representative of historical trends in crop yields. From 1981 – 2000, the NASA-MERRA climate product reports an average of 640 mm of rainfall per year over agricultural land in China. The driest year for Chinese agriculture was 1992, in which precipitation averaged 596 mm over agricultural land; the wettest year was 1990 with an average of 695 mm of precipitation over agricultural land. The range in inter-annual variation in individual counties or grid cells was greater than the national aggregate of ~10%.

Surface water and mined groundwater use for irrigation in each grid cell were assumed to be spread among all irrigated crops proportional to demand. Crop yields directly due to the application of mined groundwater were determined for each grid cell by applying the difference between irrigated and rainfed crop yields to the irrigated area supplied by the mined groundwater:

$$Y_{MinedGW_j} = f_{MinedGW_j} * A_{I_j} * (Y_{I_j} - Y_{RF_j})$$ (1.6)

Where $f_{MinedGWj}$ is mined groundwater as a fraction of total irrigation water demand for grid cell $j$, $A_{Ij}$ is irrigated area (ha) in grid cell $j$, and $Y_{I,j}$ and $Y_{RF,j}$ are irrigated and rainfed crop yields (kgCha$^{-1}$ yr$^{-1}$) in grid cell $j$. Gridded results were aggregated to 31 provincial totals.

Crop yield per unit area due to mined groundwater is a function of both mined groundwater demand and the difference between irrigated and rainfed crop yields. Both of these

variables are climate-dependent, but also vary in different ways based on other factors such as

local surface water storage, soil properties, crop composition, and management practices. The

combined effects of these variables determine how vulnerable each province's crop production is

to groundwater depletion. To capture the combined effects of these two variables, we defined

crop groundwater productivity, $CGP$ (kgC ha$^{-1}$ mm$^{-1}$) to describe the yield gains per unit of

mined groundwater for each province $p$, based on province-level data:

$$CGP_p = \frac{Y_{MinedGW_p}}{MinedGW_p} \qquad (1.7)$$

where $CGP_p$ is the crop groundwater productivity for province $p$ (kgC ha$^{-1}$ mm$^{-1}$), $Y_{minedGWp}$ is the

20-year average crop yield due to mined groundwater in province $p$ (kgC ha$^{-1}$ yr$^{-1}$), and

$MinedGW_p$ is the 20-year average amount of mined groundwater demand in province $p$ (mm yr$^{-1}$).

$CGP$ provides a method for directly comparing provinces' combined reliance on both mined

groundwater use and yield impact for crop production to each other.


## Results

### *Irrigation water demand and supply*

Over the 20 years simulated by WBMplus, Chinese agriculture required an average of

330 km$^3$ yr$^{-1}$ irrigation water withdrawals in order to fulfill gross irrigation demands (Table 1.2).

Irrigation water demand standard deviation over the 20-year simulation was 33 km$^3$year$^{-1}$, or

10% of demand. This deviation is significantly higher than the deviation in annual average

rainfall over cropped land, which is only 5% of the mean. Irrigation water demand varies

significantly across China, but generally follows patterns of irrigated area (Figure 1.2).

WBMplus tracks the sources of irrigation water, grouping them into three categories: 1)

rivers and reservoirs, 2) renewable groundwater, and 3) mined/fossil groundwater mining (Figure

1.3).  Over the 20-year simulation period, the nationally aggregated amount of irrigation water

available from the first two sources are very steady.  Rivers and reservoirs supply 115 km$^3$ yr$^{-1}$,

with a standard deviation of only 7 km$^3$ yr$^{-1}$ and a range of 105 - 134 km$^3$ yr$^{-1}$.  Similarly, the

average groundwater recharge supplied to irrigation is 107 km$^3$ yr$^{-1}$, with a standard deviation of

7 km$^3$ yr$^{-1}$ and a range of 96 to 125 km$^3$ yr$^{-1}$.  Due to the extremely steady supply of river and

reservoir water as well as groundwater recharge, most of the variability in irrigation water

demand leads to widely ranging fossil groundwater mining demands.  Average mined

groundwater demand is 125 km$^3$ yr$^{-1}$, only slightly over one third of the total average irrigation

water demand.  However, at its maximum demand of 209 km$^3$ yr$^{-1}$, mined irrigation water makes

up 49% of that year's total irrigation water demand, and at its minimum demand of 58 km$^3$ yr$^{-1}$ it

makes up only 20% of total irrigation water demand.  Demand for mined groundwater varies

significantly across China, both in absolute volume and in its relative fraction of total irrigation

water demand (Figure 1.4 and Table 1.3).  These two measures of the importance of mined

groundwater do not always vary together.

Anhui, the province with the largest irrigated area in MIRCA2000, also has the greatest

demand for mined groundwater both in terms of absolute volume (32 km$^3$ yr$^{-1}$) and fraction of

the province's total irrigation water demand (58%).  In contrast, Fujian province has a relatively

low mined groundwater demand (3 km$^3$ yr$^{-1}$), yet this demand makes up 64% of its total

irrigation water demand.  Yet a different pattern appears in Shandong province, which has the

greatest total irrigation water demand (33 km$^3$ yr$^{-1}$), but only 20% of this demand is from mined

groundwater.

*Crop yields*

We estimate an average total national crop production of 495 mega-tonnes (Mt)C yr$^{-1}$ (Table 1.2) if all irrigation requirements are fully met on irrigated cropland (as defined by MIRCA2000 irrigated crop maps). Irrigated land produces 315 Mt C, and rainfed land produces 180 Mt C. Grid-cell crop yields for irrigated and rainfed crops are area-weighted averages across all crops and rotations occurring in the grid cell (Figure 1.5). These yields are direct results from the DNDC model, and reflect the spatial variability of crop maps, soil quality, climate, and management practices, in particular single- vs. multi-cropping. The difference in yields follows a general north-south pattern (Figure 1.5), with the largest differences in the relatively dry north, and the smallest differences in the wetter south.

*Crop production from mined groundwater*

When all irrigation water demands are fully met, mined groundwater directly contributes to a large portion of China's national crop production (Figure 1.6). We estimate an average of 102 Mt C over the 20 year period modeled, with a wide range varying from 79 to 130 Mt C per year due to weather variability (Table 1.2). The maximum contribution is seen in the driest weather year, and makes up 27% of the total national production; the minimum occurs in the wettest weather year and is 15% of the total national production. Without mined groundwater, the average national crop production is 393 Mt C/year (Table 1.2), which is 79% of the 20-year average crop production with mined groundwater used as needed.

There are significant regional differences in crop yields produced from irrigation water (Figure 1.5 & Table 1.4). Of all the major agricultural provinces, Sichuan has the highest crop irrigation water productivity, gaining 26 kgC/ha/mm (Table 1.4). There is a decline in crop

irrigation water productivity with an increase in precipitation (Figure 1.7), which indicates that rainfed crop yields in the southern high-precipitation regions of China are close to the maximum yield possible given soil quality, crop selection, and management choices.

We use the crop groundwater productivity, $CGP$ (Eq. 1.7), to compare the vulnerability of provinces to the loss of unsustainably mined groundwater. Xinjiang Province had the maximum 20-year average mined groundwater demand, 596 mm/year, and Tianjin Province had the maximum 20-year average crop yield due to mined groundwater, 3,300 kgCha$^{-1}$. Tianjin also had the highest $CGP$, 11.6 kgCha$^{-1}$mm$^{-1}$. Beijing and Xinjiang have a $CGP$ of 0 kgC ha$^{-1}$ mm$^{-1}$ because simulated use of mined groundwater is zero. Of the major agricultural producers (red in Table 1.5), Henan, Shandong, and Sichuan have the highest $CGPs$ (Figure 1.8), 7.3, 5.3, and 3.9 kgC ha$^{-1}$ mm$^{-1}$, respectively. Demand for mined groundwater and crop irrigation water productivity in these provinces were also the most variable based on changes in climate over the 20-years of climate input modeled. The relative importance of mined groundwater versus crop yield gains due to mined groundwater vary between provinces; Sichuan Province has a high $CGP$ mainly due to its high mined groundwater demand, while Henan Province has a high $CGP$ because of its large crop yield gains due to mined groundwater. However, all provinces fall on a roughly linear trend between low $CGP$ and high $CGP$ (Figure 1.8). This trend is not unexpected, as both variables contributing to $CGP$ are affected by climate; dry regions have a high demand for mined groundwater and a large crop irrigation water productivity, while wet regions have the opposite (Figure 1.8).

**Discussion**

***Comparison to other studies***

*Irrigation water demand and supply*

Our results for China's irrigation water demand are similar to other model-based estimates; previous results range from 220 to 850 km$^3$ yr$^{-1}$, with most studies estimating an average of 350 – 400 km$^3$ yr$^{-1}$ circa 2000. These results assume that gross irrigation water demand was always fulfilled and that the infrastructure is in place to change the source of water depending on the availability. While global irrigation water demand estimates also vary, this study's estimate of 330 km$^3$ yr$^{-1}$ is roughly 10% - 20% of world-wide demand (Döll and Siebert, 2002; Wisser et al., 2008; Wada et al., 2012). FAO's AQUASTAT database reports China's irrigation water withdrawals to be 358 km$^3$ yr$^{-1}$, and Döll and Siebert (2002) report an irrigation water demand of 364 km$^3$ yr$^{-1}$ for all of East Asia. Wada et al. (2014) use the PCR-GLOBWB model, which is similar in structure to WBMplus, and report China's irrigation water demand to be 519 km$^3$ yr$^{-1}$ when the model is driven with the MERRA climate product.

Previous studies also report that China's groundwater withdrawals for irrigation are ~100 km$^3$ yr$^{-1}$ (Table 6), however the only study that distinguishes groundwater recharge from groundwater mining estimates a significantly smaller amount of mined groundwater, 17 - 27 km$^3$ yr$^{-1}$ (Wada et al., 2012), than this study (58 – 209 km$^3$ yr$^{-1}$). Wada et al. (2012) also estimate a "nonlocal water resource" demand of 11 km$^3$ yr$^{-1}$, a water supply that is represented in this study's mined groundwater accounting. The remainder of the difference is likely due to the use of different methods for estimating renewable groundwater, and for partitioning of inefficient irrigation water, which has the potential to significantly alter the groundwater recharge. Wada et

al. (2012) allow inefficient irrigation water to contribute to groundwater recharge, limited only by the hydraulic conductivity of the soils underlying irrigated areas. In this study, inefficient irrigation water first evaporates to fulfill the difference between AET and PET, then any remaining water is divided equally between groundwater recharge and the river network (Eq. 1.5).

*Crop production*

This study estimates a similar total crop production as the National Bureau of Statistics of China, which reports year 2000 crop production was 500 Mt C. This estimate is significantly lower than the year-2000 total crop production of 715 Mt C reported by FAO (Table 1.6). The difference between the FAO's annual production and this study's production is mainly due to differences in rainfed crops, as the irrigated crop production estimates are comparable at 315 MtC (this study) versus 323 Mt C (FAO). Our results were consistent with other crop productivity studies that show crop production increases by slightly less than two times due to irrigation (Siebert and Döll, 2010; Molden et al., 2007), since irrigated and rainfed harvested areas in China are roughly equal (MIRCA2000; Table 1.2).

In a study by Ye et al (2013), the CERES crop models (one for each crop simulated) are used to estimate baseline (years 1961 – 1990) crop production in China, as well as future (2030 – 2050) crop production under various climate change scenarios. Baseline crop yields match well with this study at 532 MtC, while future estimates show increasing yields but make no considerations of future water use or availability. Siebert and Döll (2010) use a gridded hydrology model that is similar in structure to WBMplus, and input the MERRA climate product to simulate global crop water demand. They then use an empirical method that relates the

simulated AET/PET ratio to the relative ratio between irrigated and rainfed crop yields to calculate crop production due to consumptive irrigation water (Siebert and Döll, 2010). While they do not report values for China, their results show that total crop production across all of East Asia is 545 MtC, and 78.2% (or 457 Mt) is from irrigated agriculture.

### *Groundwater depletion and crop production*

In an analysis with multiple global climate models, gridded crop models, and hydrological models of renewable surface water resources, Elliot et al. (2014) simulated future crop production (up to 2090) and found significant decreases due to a combination of climate change and reduced water supply. However, Elliot et al. (2014) included only surface water supplies in their analysis. We explicitly separated the fossil mined groundwater irrigation from renewable groundwater irrigation and surface water irrigation so that we can estimate irrigation water coming from a non-renewable source. Therefore, our estimates of crop yields due to irrigation from mined groundwater show that without this water source, China's total annual crop production would decrease by 15% - 27%, bringing it down to 352 – 448 MtC (Table 1.2). Annual grain production in China was ~ 350 MtC in the early 1980s, and ~450 MtC in the early to mid 1990s (*NBS,* 2008). Our results assume that groundwater mining occurred at the level necessary to fulfill irrigation water demand after renewable water sources were used. While it is likely that not all irrigated areas are always provided with 100% of their irrigation water demand, there is abundant evidence, both observational and modeling studies, that groundwater depletion is occurring in China, particularly in the North China Plain (*Wada et al.,* 2012b; Aeschbach-Hertig and Gleeson, 2012; Syed et al., 2008; Tang et al., 2013). Results from the Gravity Recovery and Climate Experiment (GRACE) satellite show contemporary groundwater depletion

rates of 8.4 – 14 mm yr$^{-1}$ in the North China Plain, and further analysis by Tang et al (2013) shows these rates may underestimate depletion by ~7.6 mm yr$^{-1}$.

The North China Plain includes much of Henan, Hebei, Shandong, northern Jiangsu, and northern Anhui provinces, all of which are significant agricultural producers (Table 1.5). These five provinces collectively produce 152 Mt Cyr$^{-1}$ (31% of total national production). Anhui and Jiangsu rely on mined groundwater for 42% and 49% of their total irrigation water supplies, and Henan, Hebei, and Shandong rely even more heavily on mined groundwater, requiring on average 73%, 77%, and 80% of their annual irrigation water to come from groundwater mining. Without the use of mined groundwater, crop production in the North China Plain would drop to 101 MtC yr$^{-1}$, a 10% loss in national production.

Groundwater depletion is also occurring in the northern-most and western-most parts of China (Aeschbach-Hertig and Gleeson, 2012; Syed et al., 2008), affecting Heilongjiang, Jilin, Nei Mongol, and Xinjiang Provinces, which are all relatively small agricultural producers (Table 1.5). Of these smaller agricultural producers, Jilin has the lowest relative reliance on mined groundwater, at 70%, and Xinjiang has the highest at 80%. Northern provinces also have extremely large differences between irrigated and rainfed crop yields (Figure 1.5), so that even a small loss of irrigation water causes a substantial decrease in their crop production.

We used the crop groundwater productivity, *CGP* (see Eq. 1.7), to assess which provinces are vulnerable to the loss of unsustainably mined groundwater. Vulnerability can be due to either a high reliance on mined groundwater, significant crop yields dependent upon mined groundwater, or a combination of both these factors. Provinces that have a high *CGP* and are large agricultural producers fall along a band of precipitation with precipitation roughly between 1000 and 1500 mmyr$^{-1}$ (Figure 1.8 inset). South of this band, provinces have low

reliance on mined groundwater and small increases to crop yields due to use of mined groundwater. North of this band, provinces have high *CGP*s, but they are not large agricultural producers. Therefore we expect that if groundwater levels continue to drop, the provinces in this precipitation band will have decreased crop yields unless they can secure alternative water supplies or significantly increase crop water use efficiency (e.g., Liu et al., 2010). China's southern provinces, many of which are significant agricultural producers, are less at risk of aquifer depletion than northern provinces (Döll et al., 2012), and their CGPs are low.

### *Limitations and uncertainties*

Uncertainties, sensitivity analysis, and model validation for WBMplus are discussed in Wisser et al (2010), and for DNDC in Li et al (1992a), Li et al (1994), and Wang et al (2008). Additional uncertainty introduced in this study arises from combining the two models' results, which are based on different crop maps. While the national total irrigated and rainfed areas are similar, MIRCA2000 (WBMplus input) and DNDC's crop maps disagree on which type of crops are grown on 63% of irrigated areas, and 31% of rainfed areas. DNDC also represents crop rotations at a more detailed level than MIRCA2000 (Qiu et al., 2003). This uncertainty does not significantly impact the results of this study because both sets of model results have been validated or tested against other studies, and are both predicting results consistent with observational data and other modeling studies. Our conclusions and interpretations only address the total water demand and crop production of all irrigated and rainfed crops, and we do not make any interpretations about individual crop types; this type of analyses would require crop maps with better agreement.

Paddy rice is a major consumer of irrigation water and so representation of paddy water management and water balance has significant impact on WBMplus results for China. There was a significant change in rice paddy water management in China during 1980-2000, with most farmers adopting a mid-season draining or drying management scheme (Li et al., 2002). This practice was adopted to save labor and energy, and may also conserve water (we are not aware of any studies that have quantified water savings impacts). Paddy mid-season draining/drying has not been implemented in WBMplus, so paddy irrigation water use may be overestimated in our analysis.This over-estimation may be offset by an under-estimation of the water amounts required to initially flood rice paddies. WBMplus's rice paddy flooding implementation applies 50mm of water above soil field capacity, which would be insufficient to inundate many field soils. Finally, macro-scale modeling that relates flooded paddy percolation losses to soil texture may overestimate losses in paddies that have been in managed use for centuries (e.g., the Sichuan Basin), where paddy percolation losses may be lower than would estimated from regional soil properties.

**Conclusions**

This study is the first to combine a hydrologic model with a process-based crop growth model to simulate national-scale agricultural yield and irrigation water use. This methodology allows for the direct attribution of crop yields to irrigation water from rivers and reservoirs, groundwater, and fossil mined groundwater (Figure 1.2), as well as computation of the spatially varying crop water productivity from total irrigation water, and the crop groundwater productivity. We find that mined groundwater fulfills 20% - 49% of China's national irrigation water demand, which directly leads to 15% - 27% of national crop production. Crop water

productivity and mined groundwater demand vary spatially across China, with the combination leading to a smaller percentage of crop yields dependent upon mined groundwater than the percent of total irrigation water demand fulfilled by mined groundwater. I.e., while mined groundwater fulfills 20% - 49% of irrigation water demand, crops irrigated with mined groundwater only account for 15% - 27% of national crop production, which indicates that the regions with high crop yields are not all directly dependent upon mined groundwater. This study calculated a crop groundwater productivity to determine which provinces would be most vulnerable to the loss of access to mined groundwater. We found that provinces across central China are most vulnerable due to their combination of significant agricultural production, high demand for mined groundwater, and high crop yield gains from the use of mined groundwater.

China, like all major world agricultural producers, has increased irrigation to the point of over-exploitation of water resources in order to achieve greater levels of food production. While irrigation is typically employed to reduce agriculture's vulnerability to weather and climate variability, in the case of water over-exploitation irrigated agriculture may be vulnerable to changes in the water supply caused by both climate variability and diminishing water resources. We find that these two factors – climate variability and mined groundwater demand – do not operate independently, but rather magnify one another by increasing the demand for irrigation water in a hot and dry year while simultaneously reducing the water available for irrigation use. Due to this magnification, the amount of food production dependent upon irrigation from unsustainable water supplies varies significantly from year to year.

Understanding the sources of irrigation water supply and their relative importance to crop production across China will help provide context for water resource management in China, especially with regards to groundwater storage and the South-North Water Transfer Project.

Global depletion of groundwater aquifers, including the North China Plain, will require future groundwater use to achieve greater levels of sustainability than are seen today. Simultaneously, crop production will need to increase to feed the growing world population. Increases in crop water productivity are required to achieve these two goals. A major finding of this research is that crop irrigation water productivity varies significantly across China and that inter annual weather variability results in widely ranging responses in crop irrigation water productivity, as well as in relative reliance on mined groundwater. The wide range in crop water productivity can help identify the regions in China where climate, soil conditions, and management practices work together to achieve high levels of "crop per drop". Increasing cropping efficiency will likely be a key component of feeding the world's growing population in a sustainable manner (Brauman et al., 2013). Similarly, the distinction between crop yield losses due primarily to yield differences versus reliance on mined groundwater illustrates that different strategies (e.g., increasing rainfed yields versus reducing reliance on mined groundwater) will be best suited to different regions in order to achieve the future requisite increases in crop productivity within the constraints of future water availability.

**Tables**

**Table 1.1:** Data sets input to WBM and DNDC.

| Data Type | Variables | WBM Source | DNDC Source |
|---|---|---|---|
| Climate Drivers | Temperature, precipitation | MERRA | MERRA, (Reinecker et al. 2011) |
| Soil Properties | Field capacity, wilting Point, non-crop rooting depth, soil drainage class | FAO Soil Map of the World | Third National Soil Survey (Shi et al. 2004) |
| Crop Distribution | Irrigated crops and areas, and rainfed crops and areas | MIRCA 2000 | Qiu et al. 2003 |
| Crop Calendar | Plant date, harvest date | MIRCA 2000 | Cui et al 1984 |
| Crop water use parameters | Rooting depth, crop coefficient, growth stages | Siebert and Doll (2010) | Li 2007b |

**Table 1.2:** China national 20-year mean and range of irrigated and rainfed areas, annual irrigation water demand, crop yields, and the difference between sustainable and unsustainable yields.

| Variable | Mean | Range |
| --- | --- | --- |
| Irrigated area (ha) | 79,100,000 | |
| Rainfed area (ha) | 76,600,000 | |
| | | |
| Irrigation water demand (km3/yr) | 330 | 270 - 420 |
| Mined groundwater demand (km3/yr) | 125 | 58 - 209 |
| | | |
| Fully irrigated yield (MMTC/yr) | 315 | 308 - 321 |
| Rainfed yield (MMTC/yr) | 180 | 164 - 207 |
| Sum: Irrigated + Rainfed yields | 495 | 474 - 527 |
| | | |
| Sustainable irrigated yield (MMTC/yr) | 127 | 114 - 142 |
| Sustainable rainfed yield (MMTC/yr) | 265 | 238 - 305 |
| Sum: Irrigated + Rainfed yields | 393 | 352 - 448 |
| | | |
| Unsustainable minus Sustainable Yields (Mt C/yr) | 102 (21%) | 79 – 130  (15% - 27%) |

**Table 1.3:** Province summary of irrigation water demand, and mined groundwater demand.

| Province | Harvested Irrigated Area (Mha) | Irrigation Demand (km³/yr) | Mined Groundwater Demand (km³/yr) | Mined Groundwater Demand (fraction total) |
|---|---|---|---|---|
| Anhui | 6.8 | 32 | 13 | 0.42 |
| Jiangsu | 5.2 | 26 | 13 | 0.49 |
| Hunan | 4.9 | 19 | 6 | 0.33 |
| Jiangxi | 3.6 | 16 | 5 | 0.33 |
| Zhejiang | 2.7 | 12 | 4 | 0.35 |
| Hubei | 3.4 | 13 | 5 | 0.39 |
| Shandong | 6.6 | 33 | 26 | 0.80 |
| Henan | 7.1 | 25 | 28 | 0.73 |
| Guangdong | 2.2 | 8 | 3 | 0.33 |
| Guangxi | 2.5 | 8 | 3 | 0.33 |
| Hebei | 4.6 | 24 | 18 | 0.77 |
| Fujian | 1.3 | 5 | 2 | 0.33 |
| Guizhou | 1.3 | 5 | 2 | 0.33 |
| Yunnan | 2.2 | 5 | 2 | 0.33 |
| Sichuan | 2.6 | 15 | 10 | 0.69 |
| Chongqing | 1.0 | 6 | 4 | 0.65 |
| Jilin | 2.2 | 6 | 4 | 0.70 |
| Liaoning | 2.2 | 6 | 4 | 0.69 |
| Xinjiang Uygur | 3.9 | 29 | 23 | 0.80 |
| Heilongjiang | 2.5 | 6 | 4 | 0.70 |
| Shaanxi | 2.1 | 7 | 5 | 0.75 |
| Hainan | 0.18 | 1 | 0.3 | 0.33 |
| Shanghai | 0.14 | 1 | 0.4 | 0.43 |
| Shanxi | 1.45 | 3 | 2 | 0.83 |
| Nei Mongol | 3.3 | 13 | 10 | 0.81 |
| Tianjin | 0.54 | 2 | 2 | 0.77 |
| Gansu | 1.3 | 4 | 3 | 0.80 |
| Ningxia Hui | 0.51 | 3 | 2 | 0.82 |
| Qinghai | 0.49 | 1 | 0.8 | 0.75 |
| Beijing | 0.15 | 0 | 0 | |
| Xizang | 0.16 | 0 | 0 | |

**Table 1.4:** Crop yield gains per mm of irrigation water for each of the provinces that contribute to the top 50% of cumulative national yields.

| Province | Irrigation Water (mm/yr) | Irr Yield – RF Yield (kgC/ha) | Yield Gain per mm Irrigation Water (kgC/ha/mm) |
|---|---|---|---|
| Sichuan | 31 | 810 | 26 |
| Henan | 135 | 2164 | 16 |
| Shandong | 253 | 1854 | 7 |
| Anhui | 219 | 1416 | 6 |
| Hunan | 92 | 535 | 6 |
| Jiangsu | 267 | 1413 | 5 |
| Guangdong | 49 | 182 | 4 |
| Jiangxi | 93 | 279 | 3 |

**Table 1.5:** China province 20-year average sustainable and unsustainable crop yields, with 20-year minimum and maximum yields in parentheses.

| Province | Total Crop Yield [Mt C] | Crop Yield without mined groundwater [Mt C] | Cumulative percent of national total crop yield | Cumulative % of total crop yield without mined groundwater |
|---|---|---|---|---|
| Hunan | **39** (37, 41) | **38** (36, 40) | 8 | 8 |
| Henan | **37** (33, 42) | **23** (14, 34) | 16 | 13 |
| Anhui | **36** (33, 39) | **31** (23, 37) | 24 | 20 |
| Shandong | **29** (25, 32) | **14** (5, 25) | 30 | 23 |
| Jiangsu | **28** (27, 30) | **24** (19, 31) | 36 | 28 |
| Jiangxi | **27** (26, 29) | **27** (25, 29) | 42 | 33 |
| Sichuan | **26** (23, 28) | **22** (19, 25) | 47 | 38 |
| Guangdong | **25** (24, 26) | **25** (24, 26) | 53 | 43 |
| Guangxi | **24** (23, 25) | **23** (22, 24) | 58 | 48 |
| Hubei | **23** (21, 26) | **21** (17, 24) | 63 | 53 |
| Guizhou | **22** (20, 23) | **21** (20, 22) | 67 | 57 |
| Hebei | **22** (20, 24) | **9** (3, 16) | 72 | 59 |
| Zhejiang | **16** (16, 17) | **16** (15, 17) | 75 | 63 |
| Chongqing | **15** (13, 17) | **13** (10, 15) | 79 | 65 |
| Yunnan | **14** (14, 15) | **13** (13, 14) | 81 | 68 |
| Fujian | **14** (12, 14) | **13** (12, 14) | 84 | 71 |
| Shaanxi | **13** (10, 13) | **7** (4, 10) | 87 | 72 |
| Jilin | **10** (8, 12) | **7** (5, 10) | 89 | 74 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Liaoning | **9** | (7, 10) | **6** | (5, 8) | 91 | 75 |
| Nei Mongol | **9** | (9, 10) | **2** | (1, 4) | 93 | 76 |
| Xingjiang | **9** | (8,9) | **2** | (2, 4) | 95 | 75 |
| Heilongjiang | **7** | (6, 9) | **5** | (4, 7) | 96 | 77 |
| Shanxi | **5** | (4, 6) | **3** | (1, 4) | 97 | 77 |
| Gansu | **4** | (4, 5) | **1** | (1, 3) | 98 | 77 |
| Tianjin | **3** | (3, 3) | **1** | (0, 2) | 99 | 78 |
| Ningxia Hui | **2** | (2, 2) | **0** | (0, 1) | 99 | 78 |
| Beijing | **1** | (0, 1) | **0** | (0, 0) | 99 | 78 |
| Hainan | **1** | (1, 2) | **1** | (1, 2) | 100 | 78 |
| Quinghai | **1** | (1, 1) | **0** | (0, 1) | 100 | 78 |
| Shanghai | **1** | (1, 1) | **1** | (0, 1) | 100 | 78 |
| Xizang | **0** | (0, 0) | **0** | (0, 0) | 100 | 78 |

**Table 1.6:** Estimates of irrigation water use, groundwater abstraction, and crop production from this study and others.

| China Irrigation Water Demand (km$^3$) | Groundwater Abstraction (km$^3$) | Year | Source |
|---|---|---|---|
| 270 – 420 | 58 – 209 | 2000 | This study |
| 403 | 97 (20 mined) | 2000 | Wada et al. 2012 |
| 537 | 95 | 2000 | Wada et al. 2014 |
| 358 | 101 | 2005 | FAO AQUASTAT |
| 220 – 850 | *-nr-* | 2000 | Wisser et al. 2008 |
| 358 | *-nr-* | 2005 | Jiang 2009 |
| 364* | *-nr-* | 1995 | Doll and Siebert 2002 |

| China annual crop production (Mt C) | China irrigated annual crop production (Mt C) | Year | Source |
|---|---|---|---|
| 474 – 527 | 315 | 2000 | This study |
| 500 | *-nr-* | 2000 | China Statistical Yearbook |
| 532 | *-nr-* | 1961 - 1990 | Ye et al. 2012 |
| 715 | 323 | 2000 | FAO AQUASTAT |
| 545* | 457* | 1998 - 2002 | Siebert and Doll 2010 |
| 474** | *-nr-* | 2009 | Fan et al. 2012 |

*-nr-* results not reported.

*Estimates for all of East Asia
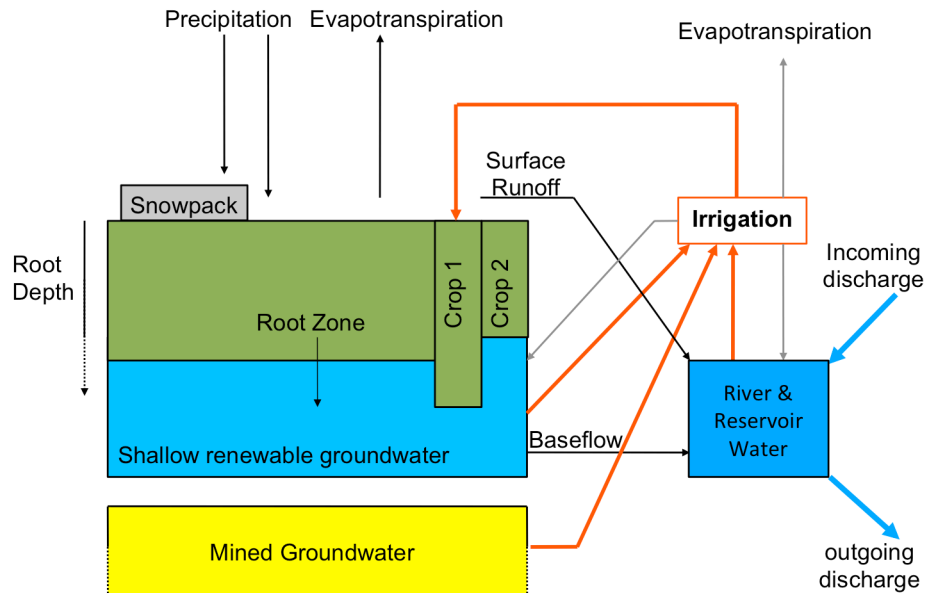
**Cereal production only

**Figure. 1.1:** Water stocks and flows in one grid cell of the Water Balance Model (WBM). Irrigation water flows are shown in orange, and inefficient irrigation water returns are shown in gray. Each grid cell can have up to 26 irrigated and rainfed crop types, each with unique root depths. The arrows into and out of the river &reservoir water stock represent water flowing into and out of the grid cell by way of the Simulated Topological River Network.
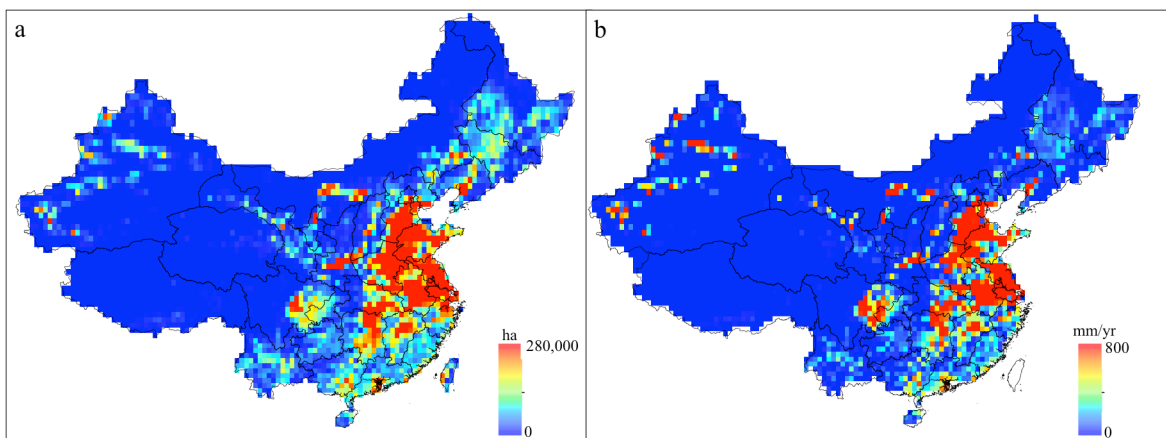


**Figure 1.2**: **a)** Irrigation area (ha) in each 0.5° grid cell from MIRCA2000, and **b)** simulated average irrigation water demand ($I_{gross}$ in Eq. 1.4, mm yr$^{-1}$).
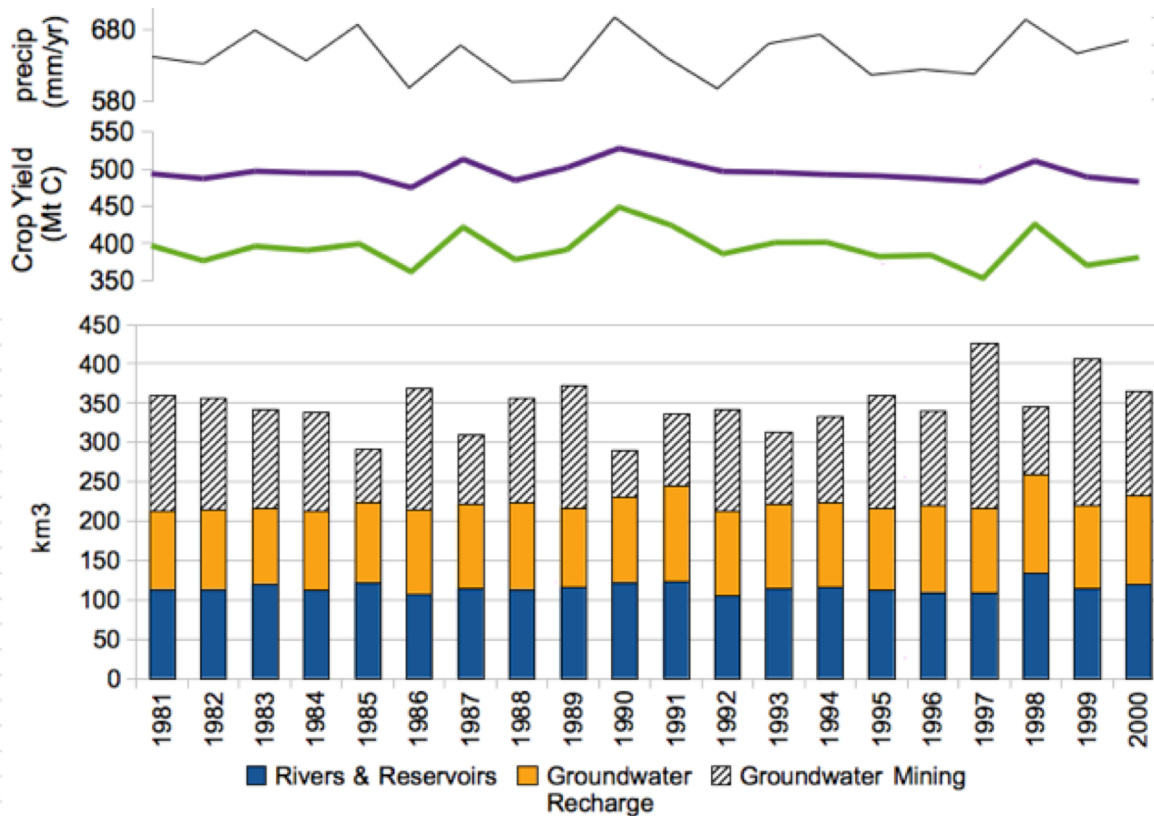
**Figure 1.3: (top)** Annual precipitation (mmyr[-1]) over all cropped areas in China from 1981 – 2000. **(middle)** Annual national crop yield (Mt Cyr[-1]) under fully irrigated conditions (purple) and under surface-water-use-only conditions, i.e., without mined groundwater (green). **(bottom)** China's annual irrigation water demand for 1981 – 2000 (interannual variation from weather variability, not changes in crop or irrigation areas). Irrigation demand is partitioned into supply from rivers and reservoirs (blue), shallow groundwater recharge (orange), and the mined groundwater (black hashed) required to fulfill all irrigated crop demands. Note that plotted yields are based on c.2000 cropping area and management, and only represent interannual variability in weather. Actual grain yields in China increased by about 2.5% - 3% per year during most of 1980 to 2000, due to improved management.
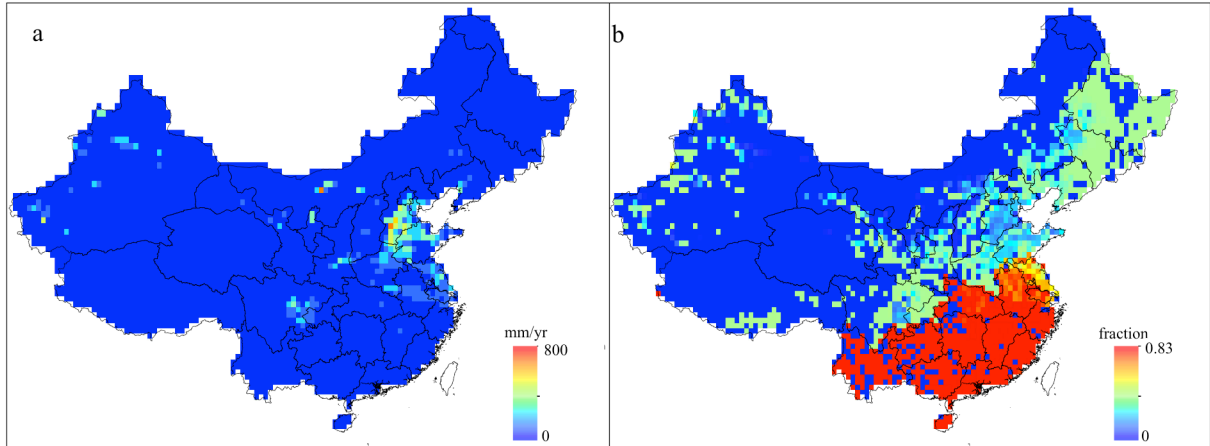
**Figure 1.4:** Mean annual mined groundwater required to meet total grid cell irrigation demand **(a)** in mm, and **(b)** as fraction of total irrigation demand.
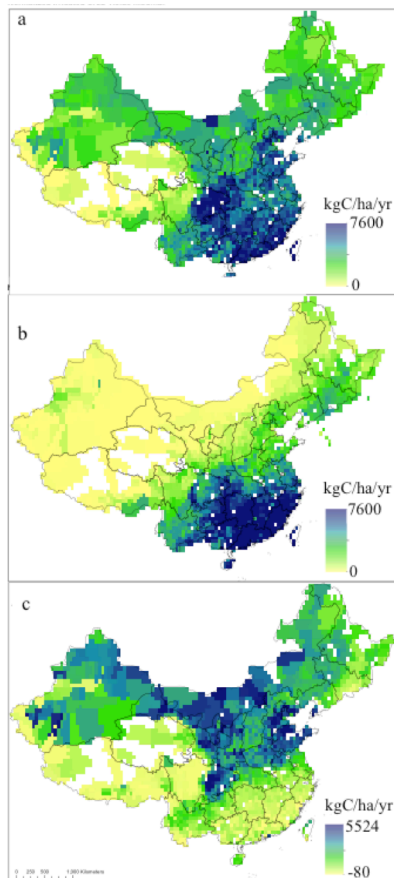


**Figure 1.5**: DNDC-simulated grid-cell mean **(a)** irrigated crop yield, **(b)** rainfed crop yield, and **(c)** difference between irrigated and rainfed yields (all kgCha$^{-1}$yr$^{-1}$). Averages are area-weighted across all crops under irrigation **(a)** or rainfed **(b)** systems, and for multiple-cropping, yields are the area-weighted total annual yield per hectare. Where the difference is negative, rainfed yields are higher than irrigated yields because different crops are grown under rainfed and irrigated conditions; in these areas the rainfed crops produce higher yields than the irrigated crops.

**Figure 1.6**: Mean annual grid-cell crop yield (kgCyr$^{-1}$) attributed to mined groundwater irrigation.

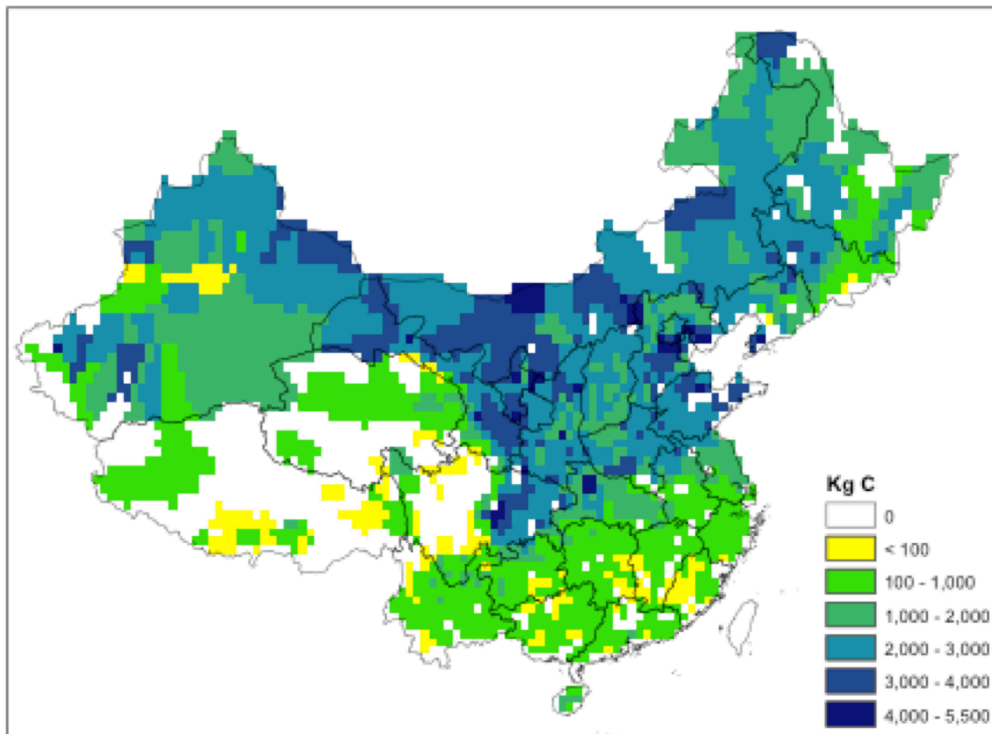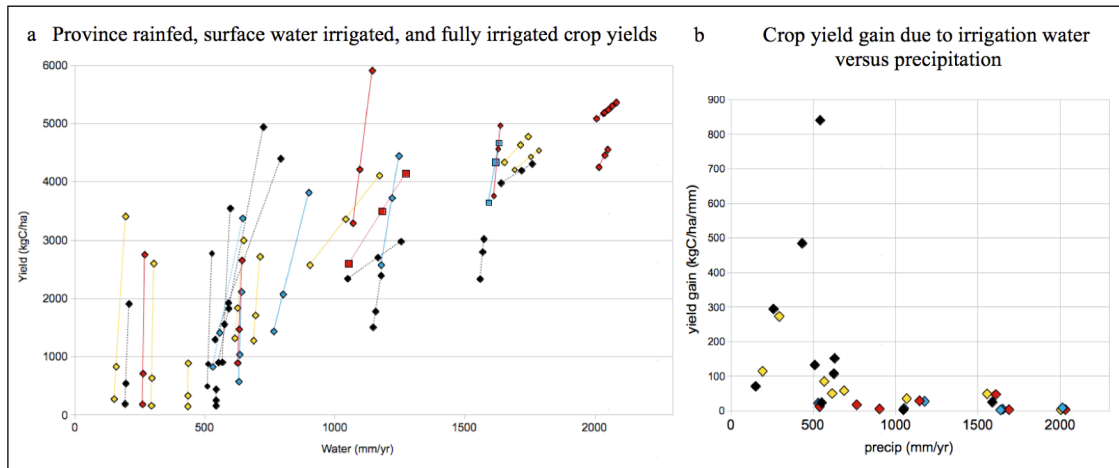**Figure 1.7**: **(a)** For each province, a line connects the rainfed crop yield (lower dot on the y-axis), the river, reservoir, and shallow groundwater irrigated crop yield (middle dot), and the fully irrigated crop yield (high dot on the y-axis); all yields in kgCha$^{-1}$yr$^{-1}$. For the rainfed yields, the water axis (x-axis) is mean annual precipitation over the cropped area in each province. For the irrigated yields, the water axis is the mean annual precipitation [mmyr$^{-1}$] over all cropped area, plus the mean annual irrigation water [mmyr$^{-1}$] applied to irrigated areas. The colors of the lines follow the color scheme for cumulative yields in Table 1.5, with the black dots and dashed black lines representing the white color-coded provinces from Table 1.5. Note that the linear relationship between rainfed, surface-water irrigated, and fully-irrigated yields illustrates the assumptions made by integrating the results from WBM and DNDC; these relationships may in reality be non-linear. **(b)** The slope of lines in part (a) [kgCha$^{-1}$ mm$^{-1}$] versus annual precipitation over cropland in each province. The slope of the lines in **(a)** represent the crop yield gain due to irrigation water.
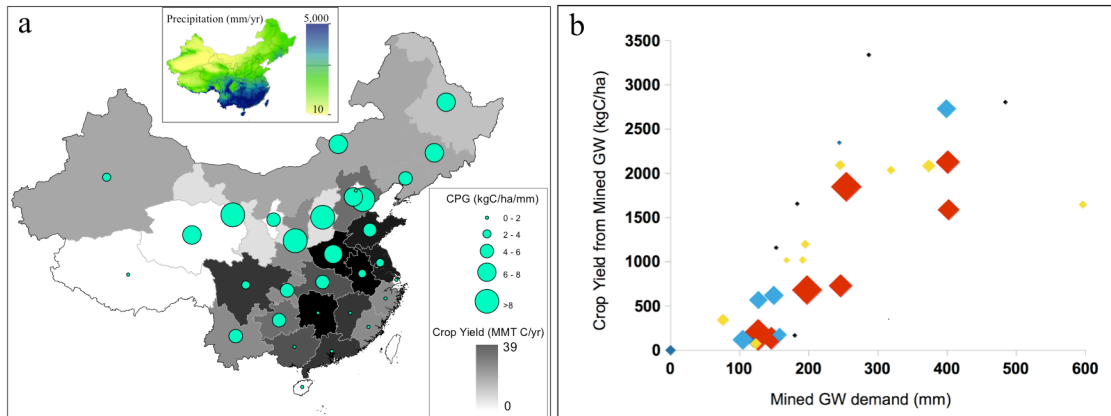
**Figure 1.8: (a)** Provinces with high crop groundwater productivity (*CPG*) and high annual crop yields fall in a region of moderate annual precipitation (inset). Each province's crop groundwater productivity (*CPG*) is indicated by the size of the filled circles, and their cumulative yield is indicated by the gray shading. **(b)** *CPG* is a function of both mined groundwater demand and crop yields due to mined groundwater; all provinces fall along a general linear trend when their crop yield gain due to mined groundwater is plotted against mined groundwater demand. Size of symbols proportional to total provincial crop yield, and coloring follows cumulative yield scheme in Table 1.5.

**Figure S 1.1:** China provinces. Color indicates the cumulative crop yield categories shown in Table 5: red = top 50% of cumulative national yields; blue = top 75% of cumulative yields; yellow = top 95% of cumulative yields.

CHAPTER II:


INVISIBLE WATER, VISIBLE IMPACT: GROUNDWATER USE AND INDIAN

AGRICULTURE UNDER CLIMATE CHANGE


**Abstract**

India is one of the world's largest food producers, making the sustainability of its
agricultural system of global significance. Groundwater irrigation underpins India's agriculture,
currently boosting crop production by enough to feed 170 million people. Groundwater
overexploitation has led to drastic declines in groundwater levels, threatening to push this vital
resource out of reach for millions of small-scale farmers who are the backbone of India's food
security. Historically, losing access to groundwater has decreased agricultural production and
increased poverty. We take a multidisciplinary approach to assess climate change challenges
facing India's agricultural system, and to assess the effectiveness of large-scale water
infrastructure projects designed to meet these challenges. We find that even in areas that
experience climate change induced precipitation increases, expansion of irrigated agriculture will
require increasing amounts of unsustainable groundwater. The large proposed National River
Linking Project has limited capacity to alleviate groundwater stress. Thus, without intervention,
poverty and food insecurity in rural India is likely to worsen.

**Introduction**

Agriculture is a significant part of India's social and political economy. While most of India's agricultural production processes are small scale in nature, in total they account for 20% of India's GDP and are India's largest employers. Moreover, the agriculture sector is the primary food supplier for India's 1.2 billion people. India is also one of the world's largest agricultural producers, and exports close to $39 billion in raw agricultural products and over 4.4 million tons of milled rice annually (Government of India, 2014; FAO, 2015). Much of Indian agriculture is heavily dependent on irrigation. Beginning in the 1960s, with the onset of the Green Revolution, India saw a significant increase in groundwater irrigation (Shah, 2010). This increase was primarily driven by the emergence of *atomistic* or personal irrigation systems and the use of subsidized power to pump groundwater from individual tube wells (Shah, 2010). Through this process, approximately 90 million rural households have come to directly depend on groundwater irrigation (Government of India, 2014). Between 1970 and 2004, while crop area remained fairly stable, irrigated area saw a rapid increase with groundwater extractions accounting for 70-80% of the value of agricultural production (World Bank, 1998, Fig. S 2.1). This underscores the important role that groundwater irrigation has played in supporting upward trends in yields and productivity.

Increased use of groundwater irrigation has led to widespread over-abstraction of groundwater resources (Rodell, Velicogna, and Famiglietti, 2009), which is unsustainable in the long term. Since 1980, groundwater levels have dropped from 8 meters below ground level (mbgl) to 16 mbgl in northwestern India, and from 1 to 8 mbgl in the rest of the country (Rodell, Velicogna, and Famiglietti, 2009; Aeschbach-Hertig and Gleeson, 2012; Sekhri, 2012). Northwestern India lost 109 km$^3$ of groundwater between 2002 and 2008 (Rodell, Velicogna,

and Famiglietti, 2009), which is an order of magnitude larger than the groundwater depletion experienced by California's Central Valley during the same period (Famiglitti et al., 2011), and twice the volume of India's largest surface water reservoir (Rodell, Velicogna, and Famiglietti, 2009). Previous research has demonstrated that groundwater declines can lead to increases in poverty and threaten food production – especially for rural households (Sekhri, 2014; Seckler et al, 1998). This directly affects small-scale farmers, who typically own < 2 ha of farmland, control the majority of the landholdings in India and produce 41% of India's food grains (Singh, Kumar, and Woodhead, 2002). These farmers use groundwater to irrigate half their land and are likely to be the hardest hit by continued declines in groundwater, thus adding to their already high vulnerability. Consequently, the sustainable use of groundwater into the future remains a serious concern for India, especially given the substantial number of rural households dependent on it for their sustenance. Previous studies in the fields of hydrology (Biemans et al., 2012; Elliot et al., 2014) and economics (Fishman, 2012) have individually assessed the future of groundwater-based agriculture in India; however, each of these fields misses key elements. Previous hydrology work has not accounted for the dynamic behavioral response of farmers to changing climate, while econometric studies have failed to account for biophysical constraints on water supplies. Researchers and policymakers need a modeling system that can account for both human irrigation decisions and the physical water supply in order to understand how climate changes may affect unsustainable groundwater (UGW) use -defined as any groundwater abstraction in excess of recharge - groundwater levels, and agricultural production.

In this study, we assess how climate change will drive human irrigation decisions, and in turn how both those human decisions and the physical changes in climate will alter both crop water requirements and available water resources. This approach requires integrating an

econometric model with a process-based hydrology model (Fig. 2.1), and allows us to quantify not only how climate change will alter each system individually, but how it will affect the entire coupled system. This multidisciplinary approach is required for study of the widely varying spatial differences in water resources and crop production across a country that is both experiencing a significant water crisis and is home to a third of the world's extreme poor, who primarily rely on agriculture for their livelihoods (World Bank, 2015). The combination of both models is critical for India in order to better plan for the future, as well as assess the role that adaptation responses and policy measures play going forward. One such policy initiative proposed by the Government of India is to move 178 billion $m^3yr^{-1}$ of water across river basin boundaries (Chellaney, 2011). This National River Linking Project (NRLP) has been proposed as a solution to groundwater stress by increasing irrigated agriculture through surface irrigation and artificial groundwater recharge. Better understanding of future irrigation water demand and availability, with emphasis on UGW dependence, is critical to assess such policies and formulate effective strategies to adapt to climate change.

## Methods

We make use of detailed crop-wise agriculture and weather data spanning 36 years from 1970-2005 for all the districts in the main agricultural states of India, and a panel data regression to estimate the relationship between inter-annual variation in monsoon climate variables and district-level irrigation decisions. This relationship is used to generate projections of irrigated area in response to climate change. The hydrology model then simulates irrigation water demand and supply from surface and renewable- and non-renewable groundwater based on these projections and future climate inputs (see Supplemental Information). Our coupled-model

approach allows us to assess India's future dependence on UGW and its impacts on groundwater-based agricultural production. We also quantify the effect of the proposed NRLP in mitigating groundwater stress.

*Econometric model: human system*

We project irrigated area for six major crops in India - the staple cereal crops rice and wheat (the focus of the Green Revolution); coarse cereals maize, sorghum and barley; and a high-value crop cotton. Barley and wheat are dry season crops, while maize, sorghum and cotton are wet season crops. Rice is grown in both seasons. Historical agricultural data was acquired from the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) and historical data of temperature and precipitation were acquired from the Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) product (See Supplemental Information for further details).

The empirical model of irrigation decisions assumes that the planting decision has already been made. Therefore, irrigation decisions reflect the second stage in a farmer's decision making process, and each crop regression only accounts for the sample of districts that grow a particular crop over our study period, 1970-2005. We estimate the following equation for each crop:

$$\log Y_{d,t} = \gamma_0 + \alpha \, logY_{d,t-1} + \beta \, log \, R_{d,t} + \gamma_1 logGDD + \gamma_2 log\overline{A_{d,t-1,t-6}} + \rho_d + \lambda_t + A_{s,t} + \epsilon_{d,t} \qquad (2.1)$$

where $Y_{d,t}$ is the irrigated area of a particular crop in district $d$ in year $t$; $Y_{d,t-1}$ is a lag of the dependent variable, which captures spillover effects from investments in irrigation infrastructure affecting all subsequent irrigation decisions; $R_{d,t}$, which follows from previous research (Fishman, 2012), represents June-September monsoon rainfall and the number of days with

precipitation >0.1mm; and $A_{d,t-1,t-6}$ is the previous 5-year average crop area, which captures the expectation to plant in the current period.

In this model, we use short-run random variation in climate in a given area to compare that area's outcomes under different weather conditions after controlling for observed and unobserved characteristics using regional fixed effects, $\rho_d$, and a time fixed effect that further neutralizes any common trends. These time fixed effects are represented by $\lambda_t$ and $A_{s,t}$; $\lambda_t$ accounts for all common contemporaneous trends, for example, national price changes, economic growth and population growth while state-specific trends, $A_{s,t}$, allow to control for differential trends by state such as state-wise technological progress and changes in electricity subsidies. In this way, a district observed during a dry year, acts as a 'control' for that same district observed during a wetter 'treatment' year. This is important because, for instance, one district might be wetter than another district, and also much wealthier. If irrigation decisions are a function of wealth, then a statistical model that just compares precipitation to irrigation outcomes between the two regions, without accounting for income differences, is likely to be biased.

We adopt linear and Tobit regression approaches when modeling irrigated area. Zero irrigated areas reflect optimal outcomes of a decision and are modeled using a Tobit approach. Standard errors are clustered at the district in the Tobit models, and corrected for spatial and serial correlation in the linear models.

*Irrigated area projections*

We combine our estimated historical irrigation response to weather (Table 2.1) with bias-corrected climate projections from 5 GCMs that have contributed to the World Climate Research

program's Coupled Model Inter-comparison Project phase 5 (CMIP5) under the RCP 8.5 scenario (Taylor, Souffer, and Meehl, 2012). The 5 GCMs used in this study are CCSM4, GFDL-CM3, GFDL-ESM2G, MIROC-ESM-CHEM, and NorESM1-M (Table S 2.1). We chose these models because they (i) come closest to characterizing India's historical monsoon (Menon et al., 2013; Sooraj, Terray and Mujumdar, 2014), and (ii) demonstrate a wide distribution of future climate changes, including both increases and decreases in monsoon rainfall and the number of rainy days within the monsoon season (Fig. S 2.3). The bias correction method and number of models used follows the Inter-Sectoral Impact Model Inter-comparison Project approach (Hempel et al., 2012). Further details are in Supplemental Information.

### *Hydrology model: physical system*

Projected changes in irrigated area and bias-corrected CMIP5 climate projections are inputs to the hydrology model, WBMplus (Wisser et al., 2010; Grogan et al., 2015). WBMplus is a gridded process-based hydrology model, used to represent the spatial and temporal water cycle in India, including crop water use. It simulates vertical water exchange between the land surface and the atmosphere, and horizontal water transport through runoff and stream networks (Wisser et al., 2010), and computes irrigation water demand, supply, and use by crop type (Grogan et al., 2015). WBMplus uses a 30-arcminute grid resolution with the simulated topological river network STN-30p (Wisser et al., 2010; Grogan et al., 2015). The climate inputs – both historical and future – are the same as used for the econometric analysis and projections (Supplemental Information Sections S 2.2 and S 2.3). WBMplus simulated the climate for each GCM separately, including a 10-year spin-up with each GCM's historical output (1996-2005). Historical agricultural data (i.e., irrigated area by crop type) is from ICRISAT (Supplemental

Information Section S 2.1), and future agricultural data is from the econometric model projections. The field capacity and wilting point of soils is a required input for calculating both crop and non-crop potential evapotranspiration (PET). We use the FAO/UNESCO Soil Map of the World (Burke and Emerick, 2015) for both, as well as for the soil drainage classes. Soil drainage class is used to estimate water seepage through flooded rice paddies. Large reservoirs are represented as river segments with the reservoir's storage capacity; data on large reservoir capacity, location, and primary purpose are from the GRanD database (Lehner et al., 2011).

### *Unmet irrigation demand: modeling the loss of groundwater access*

We quantify current (c.2000) irrigated crop production that can be attributed to UGW by simulating historical crop water requirements as described above, but restricting irrigation water supplies to sustainable sources. Unmet irrigation water demand was assumed to lead to contraction of irrigated areas (Fishman et al., 2011). We apply the contraction of irrigated area equally to all crops within a district. This is a first-order assumption; water deficits will likely cause uneven distribution of irrigated area contractions due to a range of factors. These factors may include the profitability of the crop type, the ability of the crop to survive under deficit irrigation conditions, or the crop water productivity of a crop type. Results shown here should be considered as estimates of crop production losses due to UGW restrictions; further work on this topic will help increase the accuracy of estimates shown here. In the dry season, reduction of irrigated areas is assumed to equal a reduction in crop area, as most dry season crops cannot be grown without irrigation. In the wet season, we assumed farmers would grow the same area of crops, but under rainfed conditions. Therefore, dry season crop production losses, $P_{loss,d}$ (tons), due to UGW restrictions are:

$$P_{loss,d} = \frac{I_{unmet,d}}{I_{gross,d}} * (A_1 * IY_1 + \cdots + A_N * IY_N) \tag{2.2}$$

and wet season crop production losses, $P_{loss,w}$ (tons), due to UGW restrictions are:

$$P_{loss,w} = \frac{I_{unmet,w}}{I_{gross,w}} * (A_1 * (IY_1 - RY_1) + \cdots + A_N * (IY_N - RY_N)) \tag{2.3}$$

where $I_{unmet,d(w)}$ is dry (wet) season unmet irrigation water demand, $I_{gross,d(w)}$ is the dry (wet) season gross irrigation water demand $A_N$ is the area (ha) of crop $N$, and $IY_N$ (tons/ha) is the irrigated yield of crop $N$, and $RY_N$ (tons/ha) is the rainfed yield of crop $N$, and $_{Igross,w}$ (km$^3$) is the gross irrigation demand in the wet season. Crop production data is from ICRISAT.

### *National River Linking Project*

To assess the NRLP's potential to reduce India's groundwater stress, WBMplus models all published proposed irrigation water transfers across river basins (Ghassemi and White, 2007). Two scenarios were simulated: one implements only the inter-basin water transfers through the proposed canal system; the other additionally adds reservoirs at each water recipient location. In both simulations, the daily volume of water moved through the canals is a function of river discharge at the donor location and the canal's capacity. In the second simulation, reservoir capacity is added to allow water transferred during the wet season to be stored until it is needed for irrigation in the dry season. We recognize that this simulation is highly speculative, as few details of the proposed increase in water storage have been published.

### *Method limitations*

The methods described above allow a human decision-making process – i.e., expansion or contraction of crop-specific irrigated areas – to be be fully integrated into a physical hydrology system analysis. While inclusion of the human decision-making process is an

important advance, the methods used here have several key assumptions which lead to limitations that must be defined. First, projections of irrigated areas assume that current trends in variables other than weather continue into the future: e.g., population growth, GDP growth, and technological advancement are all assumed to continue along their current trajectory. Therefore, this model system cannot account for significant extreme events or shocks to the economic or physical system. Second, projections of irrigated areas can only be made in districts that had crops planted in the historical assessment period. This limits the study by not allowing us to project the expansion of agriculture (irrigated or rainfed) into new areas. Third, the hydrologic model simulates crops that are either fully rainfed, or fully irrigated – deficit irrigation, the practice of providing only a portion of the amount of irrigation water required for optimal growth, is not simulated. Crop yields are not perfectly linear with added water and the relationship between water additions and crop yields vary between different crops (Fishman, 2012). Therefore, when modeling the crop yield loss due to losing access to UGW, our methods only provide a first-order estimate. Fourth, groundwater levels drop unevenly across the country, and so loss of access to UGW will occur at different times in different places. The methods presented here do not capture this temporal variability, nor do they capture the potential human response to the gradual loss of UGW. Lastly, the simulation of the National River Linking Project is based on imperfect knowledge of the location and storage capacity of planned reservoirs or other water storage systems (Amarasinghe, Shah and Malik, 2009). The simulation results shown here are meant to optimize water storage and release volumes and timing, and therefore may not be representative of actual reservoir construction or storage volumes.

The methods used here and there limitations define the scope of this study: we are able to (a) project how changes in climate (and continued trends in all other relevant variables) will lead

to hunan-based changes in crop-specific irriagted areas, (b) quantify how much irrigation water these projected irrigated areas will require under the same set of climate change conditions, (c) quantify how much unsustainable groundwater, UGW, would be required to fully meet these projected irrigation water requirements, (d) categorize changes in the rate at which groundwater levels will drop in response to these projections, (e) estimate the quantity of food produced as a direct result of current UGW use, and (f) estimate a potential range for UGW alleviation under full implementation of the National River Linking Project.

A detailed account of the methods and data is available in the Supplementary Information.

## Results

### *Monsoon Impact on Historical Irrigated Areas*

The first step in assessing the future of irrigation in India is to quantify how historical changes in climate have driven historical patterns of irrigation. This understanding allows us to project changes in irrigation into the future, based on a suite of potential climate change scenarios. India has a monsoonal climate with a wet (Kharif) season that receives up to 1m of rainfall and a dry (Rabi) season in which precipitation is insufficient to grow most crops and irrigation must be used. Consequently, farmers assess the supply of rain during the monsoon season and the amount that gets stored at the end of the season, in order to make decisions about increasing or decreasing irrigated areas for different crops (Fishman, 2012). Thus, there is a significant link between monsoon rainfall and irrigated areas in India (Siegfried et al., 2010; Fishman et al., 2011). The future of monsoon rainfall is extremely uncertain; some climate change studies show an increase in future precipitation (Chaturvedi et al., 2012), others predict a decrease (Ashfaq et al., 2009) (Fig. S 2.3). Increases in inter-annual and intra-seasonal variation

are also expected (Menon et al., 2013) along with rising temperatures (Chaturvedi et al., 2012; Fig. S 2.4). Such climatic change will affect irrigation water demand and supply due to farmer irrigation decisions, water supply, and physiological crop water requirements.

The econometric model estimates the effect of total precipitation, rainfall distribution, and seasonal growing degree days (GDD) on seasonal, crop-wise irrigation decisions for six major crops in India, which make up 80% of India's crop production (Table 2.1), across districts in all major agricultural states in India from 1970 to 2005. The logarithmic specification of the model enables interpretation of coefficients as elasticities, so that a 1% increase in the weather variable affects irrigated area by $\hat{\beta}$%, where $\hat{\beta}$ is the coefficient of interest. Precipitation and GDD have less than proportionate impacts on irrigation decisions (i.e., $|\hat{\beta}| < 1$).

Precipitation plays a larger role than GDD in driving changes in irrigated area (Table 2.1). The number of rainy days (i.e., the distribution of monsoon season rainfall) directly affects wet season crop irrigation, as too many days without rain during critical crop stages can reduce yields or lead to crop failure (Gadgil and Kumar, 2006). Supplemental irrigation in the wet season, largely relying on stored monsoon rainwater from previous years, can help overcome this uneven distribution of rainfall, but may not be able to offset decreases in total precipitation. Negative coefficients on the number of rainy days for maize (p<0.001), sorghum (p<0.1) and cotton (p<0.001) reflect a rise in irrigated areas for most wet season crops in response to fewer rainy days, even when total rainfall is controlled for. In contrast, the impact of the total amount of rainfall on wet season irrigated area of rice, sorghum and cotton is varied. Of the three crops, sorghum is least water intensive and most drought resistant (Brouwer and Heiblem, 1986), so a fall in monsoon rainfall can be easily compensated by supplemental irrigation to meet its irrigation needs. Rice and cotton are more water intensive, rice due to the practice of flooding

paddy fields, and cotton due to high crop water requirements for optimal growth (Brouwer and Heiblem, 1986). As a long duration crop, cotton sometimes extends into the dry season, increasing its water needs substantially. For these two crops, coefficients on the total amount of rainfall are positive and significant ($p<0.001$), implying that a reduction in total monsoon rainfall also decreases irrigated area. In the irrigation-intensive dry season, the capacity to irrigate rests on the amount of monsoon rainfall collected in surface and groundwater storage. Any decrease in precipitation during the previous monsoon season significantly ($p<0.001$) reduces the area of rice and wheat that are irrigated. Barley, another dry season crop, is of short duration, hardier than wheat, relatively drought resistant (Brouwer and Heiblem, 1986) and relies on conserved soil moisture for its water needs (Majumdar, 2013). A more even distribution of monsoon rainfall helps retain soil moisture and can significantly ($p<0.05$) decrease barley irrigation.

The impact of GDD on irrigation is limited, with higher wet season GDD significantly contracting irrigated area for only maize ($p<0.05$) and cotton ($p<0.001$). Studies suggest that with an increase in temperature and water stress on plants, farmers tend to contract agricultural activity to smaller areas during the season (Siegried et al., 2010). Higher wet season GDD can also affect irrigation in the ensuing dry season, but in the opposite manner. We find that dry rice irrigation significantly ($p<0.01$) increases in response to a rise in wet season growing degree days.

The crop-specific understanding of links between climate and irrigation presented here are necessary to generate projections of irrigated areas for each crop individually, since they are key to understanding future water requirements, as crops have varying levels of water requirements.

*Projections of Future Irrigated Areas under Climate Change*

Farmers in India tend to respond to water scarcity along the extensive margin, changing the extent of cultivated and irrigated area rather than the rate of water use per unit area (Fishman et al, 2011). We combine our elasticity estimates (Table 2.1) with predicted changes in precipitation and GDD from five different general circulation models (GCMs) to project crop-specific irrigated areas (Fig. 2.2). The GCMs used model the historical Indian monsoon well (Menon et al., 2013; Sooraj, Terray, and Mujumdar, 2014), yet they also span the range of positive to negative projected changes in future monsoon rainfall (Fig. S 2.3). These irrigated area projections implicitly assume that historical irrigation decisions in response to changes in precipitation and temperature continue into the future, and that any future adaptation to a changing climate is fully embodied in the observed ability to adapt to past changes.

In developing countries like India, the majority of farmers face credit constraints, incomplete markets, lack of information, and low levels of human capital, which limits their ability to quickly adopt new technologies or improve upon existing ones (Sui Wing and De Cain, 2014; Jack, 2011). Thus, the estimates from our panel data models reflect the effects of climate change in the short- to medium-term scenarios, where farmers might be unable to adjust or re-optimize their decisions, or can only do so very slowly (Sui Wing and De Cain, 2014). Recent research suggests that the degree to which people adapt to longer-run changes in temperature and precipitation reflects surprisingly little adaptation (Burke and Emerick, 2015). Thus, our econometric model evaluates the effects of climate on changes in irrigated area in the medium term (up to 2050), while keeping cropping decisions, growing seasons, and other variables unchanged and assuming trends in technology and population stay constant into the future.

Since the application of irrigation to cultivated cropland is a short-term adaptation response by farmers in the face of inter-annual monsoon variation, irrigated area projections are made year to year and we convert the estimated % changes in irrigated area into absolute values using the previous year's irrigated area as the base. While we acknowledge that the path of development in India will change in the future, it is nevertheless instructive to project irrigated areas to assess the possible magnitude of climate-change related UGW needs. We project irrigated area increases in both seasons, with uncertainty (+/- 15% in the wet season, +/- 50% in the dry season by 2050) due to the range of GCM-projected future climates (Fig. 2.2). These increases are due to the extent of irrigated wheat and rice continuing their historical rising trend, while irrigated extent of other crops remains the same or decreases (Fig. S 2.5)

***Impacts of Climate Change on Unsustainable Groundwater Abstraction***

UGW abstraction is unsustainable in the long term, and can exhaust groundwater resources if continued unabated (Aeschbach-Hertig and Gleeson, 2012). We use a process-based hydrology model that separately models both sustainable irrigation – that supplied by groundwater recharge and surface water rivers and reservoirs – and UGW. By integrating the agro-economic econometric model with the hydrology model, we assess the impact of climate change and the resulting changes in irrigated areas on future UGW demand (Fig. 2.3)

India's northwest region has already experienced significant groundwater level decreases due to UGW use (Rodell, Velicogna, and Famigliettie, 2009). We use our model projections of future UGW demand to infer how groundwater levels will change up to 2050. If demand increases, then groundwater levels will drop more rapidly (Fig. 2.3, dark red); continued demand will lead to continued rates of groundwater level decline (Fig. 2.3, red), while reduced but

positive demands will slow the rate of groundwater level decline (Fig. 2.3, yellow). Some districts will be able to rely solely on sustainable water supplies, allowing groundwater levels to recover (Fig. 2.3, blue). Under future climate change, most of Punjab and Haryana, northern areas of Rajasthan and Gujarat, and parts of Uttar Pradesh and Tamil Nadu will face continued and further groundwater level declines (Fig. 2.3). Additionally, the spatial extent of UGW pumping expands to pockets of Tamil Nadu, Andhra Pradesh, Uttar Pradesh and Gujarat, regions that previously did not over-abstract groundwater (Fig. 2.3, orange).

Free or flatly tariffed electricity provisions have played a critical role in enabling groundwater extraction (Badiani, Jessoe, and Plant, 2012), and might further contribute to UGW use if present day irrigation and cropping practices persist. However, despite the presence of subsidies, expensive pump technology is still needed to draw groundwater once levels drop beyond certain thresholds (Sekhri, 2011). Therefore, evidence of continued and increased future groundwater level declines reflect potential constraints on access as rising pumping costs can eventually make extraction prohibitive.

To assess how a loss of access to UGW may affect Indian agriculture, we quantify the amount of unmet irrigation water demand that will occur in its absence by restricting the use of the UGW within the hydrology model (Fig. S 2.7). Without UGW, unmet irrigation water demand will reach 170 $km^3y^{-1}$ by 2050 (Fig S 2.7), paralleling only the unmet demand in 2002, a year in which India was hit by a massive drought (Bhat, 2006).

### *Policy Implications*

Losing access to UGW directly translates to reductions in food production. We find that currently, half of dry season irrigated crop production and a quarter of the total annual irrigated

crop production is directly sustained by UGW (Table 2.2). The most affected regions primarily grow India's staple crops – wheat and rice – in the dry season (Smilovic, Gleeson, and Siebert, 2015). The fertile Indo-Gangetic Plain is one of the most intensely farmed and populated areas in the world, and includes much of Uttar Pradesh, Punjab and Haryana, which have districts that produce up to 1.8 million tons of UGW-based agricultural output each year (Fig. 2.4). The southeastern states of Tamil Nadu and Andhra Pradesh also rely heavily on UGW for crop production, with some districts producing up to 0.8 million tons per year using UGW (Fig. 2.4). Therefore, in the event that UGW becomes difficult to access, national food security will be threatened. We find that UGW in India is directly responsible for production of sufficient food calories to feed 173 million people, accounting for 14% of India's population (Table 2.2). In a country where ~194 million people go hungry every day (FAO, 2015), losing access to UGW would further aggravate food security concerns that already plague India.

Based on the range of projected future climates, UGW will either remain at historical levels ($30 - 40 \text{ km}^3\text{yr}^{-1}$), or it will increase to as much as $170 \text{ km}^3\text{yr}^{-1}$ (Fig. S 2.7). Historically, the irrigation water deficit that would occur due to restricting UGW use is $5 - 15\%$. Future irrigation water deficits may stay at this level, or increase up to 40%. If the future UGW demand remains the same as the present, than the food cost of eliminating UGW use will remain at levels summarized in Table 2.2. However, the extreme case of increasing irrigation water deficits to 40% would result in significantly larger reductions in food production. However, it is worth noting that future advances in agricultural systems may allow for higher crop yields under water stressed conditions. Additionally, increased irrigation efficiency may allow for more water to be used beneficially by crops even with reduced water extractions.

A recent government initiative has looked to the massive NRLP to overhaul water management in India. The $123 billion project intends to move 178 km$^3$ yr$^{-1}$ of water by connecting 37 rivers, building ~3000 storage dams and 12,500 km of water conveyance networks (Chellaney, 2011; Amarasinghe, Shah, and Malik, 2009). If completed, it will be the biggest infrastructure project in the world (Amarasinghe, Shah, and Malik, 2009). In addition to its stated goals of 34 GW of hydropower generation, increasing drinking water supplies, and mitigating floods in the east Amarasinghe, Shah, and Malik, 2009), it is also expected to alleviate the stress on groundwater. The NRLP is expected to increase the extent of irrigated agriculture by 35 million ha through surface irrigation and improved groundwater recharge (Amarasinghe, Shah, and Malik, 2009)

To quantify the NRLP's impact on UGW demand and surface water irrigation, we simulate a scenario in which all proposed river links are functional along with concurrent construction of large reservoirs at receiving nodes in the NRLP, and compare the UGW demand to our baseline model results. We find that with both the additional reservoirs and the inter-basin transfer network functioning, there is potential to alleviate as much as 16% of India's mid-century UGW demand (Fig. 2.5b). However, without new large reservoirs, the inter-basin transfers alone reduce only 1-4% of overall UGW demand (Fig. 2.5c). Historically, construction of large dams has been contentious in India (Bhawan, 1989). While the exact plans for dam construction under the NRLP have not yet been publicized, it is clear from these results that without a large increase in reservoir capacity, the NRLP will not alleviate groundwater stress in northwest India.

**Discussion**

In this paper, we project future unsustainable groundwater (UGW) use by accounting for changes in both demand and supply of water and their interaction under different climate change scenarios. We are able to identify regions where groundwater demand and supply may change significantly (Figs. 2.3, 2.4), which has important implications for policy decisions that affect agricultural development, poverty, food security, and adaptation.

One policy recommendation for managing water is the NRLP, but our results demonstrate canals alone have limited ability to decrease unsustainable use of groundwater nationally (Fig. 2.5c). In addition to the NRLP, several other policy and infrastructure tools have been studied, and research has indicated they have potential for reducing groundwater abstraction in different areas. Interventions researched include investments in public provision of groundwater to crowd out private construction of wells (Smilovic, Gleeson, and Siebert, 2015), decentralized rainwater harvesting schemes (Famiglietti et al., 2011), creation of groundwater markets (Foster and Sekhri, 2008), metering electricity (Mukheraji, Shah, and Giodano, 2012), and power supply rationing by separating agricultural from non-agricultural feeders (Shah and Verma, 2008). In the end, our coupled model suggests that, regardless of the mechanism, support for sustainable groundwater management will become increasingly urgent in the near future.

**Conclusions**

The results of this study provide the first multidisciplinary assessment of the extent of UGW use in India through mid-century, its importance in sustaining food production, and the potential role of large infrastructure projects in decreasing India's dependence on UGW. Our results emphasize that under a business as usual scenario, climate change induced precipitation

increases in certain areas will not necessarily alleviate groundwater stress, due to the expansion

of irrigated areas. This analysis also points to the need for a thorough analysis of farmer

decision-making responses to infrastructure projects such as the NRLP and other policy

measures that affect the availability of irrigation water supplies,  as it is likely that subsistence-

level food security concerns may drive these decisions. While we quantify here the potential for

the NRLP to alleviate groundwater stress, it is possible that expansion of irrigated areas in

response to the project (as has been promoted by NRLP-planners) will negate these potential

benefits.


**Supplemental Information**

*Agricultural data*

The historical agricultural data used in our analysis was acquired from the International

Crop Research Institute for the Semi-Arid Tropics (ICRISAT) and their Village Dynamics in

South Asia (VDSA) database, which collates data from State Directories of Agriculture, State

Bureaus of Economics and Statistics, State Planning Departments, various Agricultural Censuses,

and government reports. It provides documentation of the different sources used in compiling

data across various variables.  This is the only long-term *publicly* available recording of district

level statistics. The dataset includes district-level irrigation and crop area data for both Kharif

(wet) and Rabi (dry) season crops across all major agricultural states in India from 1966 to 2006.

Data for a large proportion of districts in India are available from 1969 to 2005. We therefore use

data from 1970-2005 in the historical econometric analysis. A district is an administrative unit

under the Indian state that is the lowest level of disaggregation for which agricultural data are

uniformly available across India. Since Indian district boundaries change over time and larger

districts have been split into smaller ones (219 new districts were formed between 1966 and 2007), we use district boundaries from the year 1966 to allow for comparisons over time. To construct time-series data, we must have a consistent district definition. Therefore, districts formed after 1966 are mapped back to their parent districts (i.e. districts from which they were formed) based on the percentage of geographical area of the parent district that was transferred to the new district.

District-level agricultural statistics provided by the government's land use surveys (LUS) report annual irrigated area for each crop, and season-wise irrigated area for sorghum and rice. The online LUS database provides these statistics only 1998 onwards. The ICRISAT database, on the other hand, uses all the land use surveys to compile season-wise irrigated area for sorghum from 1970 to 2005, but *not* rice. Apart from rice, the other crops used in the econometric analysis are largely grown in either the wet or dry season. The wet season coincides with the timing of the summer monsoon, which spans approximately June through September. The dry season spans approximately November through February. On average, most wet and dry season crops are grown across these two seasons. While rice is predominantly grown in the wet season in India, some states in the South and East (such as Andhra Pradesh, Assam, Bihar, Karnataka, Kerala, Maharashtra, Orissa, Tamil Nadu and West Bengal) grow rice in both seasons (Frolking, Yeluripati, and Dougles, 2006). For these states, we split annual irrigated area, as reported in the ICRISAT database, into seasons using district-level wet-season and dry-season irrigated area from the early- to mid-1990s (Huke and Huke, 1997), as follows:

1. We first calculated the proportion of total irrigated area for rice that fell in either the wet or dry season for each district covered in Huke and Huke (1997).

2.  Since district boundaries changed over time, we intersected the 1966 boundaries that we used in our analysis with those used in Huke and Huke (1997) and apportioned the area weighted average for each season's crop to the 1966 districts.

3.  These seasonal proportions were then multiplied by the annual total for irrigated rice area in our dataset to get each season's irrigated area.

Thus, the underlying assumption used to split data into seasons is that districts have different amounts of seasonal irrigated rice area over time, but the proportion of wet- and dry-season irrigated rice area stays constant. Only states that grow rice in both seasons are used in regressions that involve dry-season rice. Additionally, for these regressions the previous five year average of crop area is calculated using annual crop area.

### *Historical weather data*

Observed temperature and precipitation data were acquired from the gauge-based observationally-gridded daily dataset Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) (Yasutomi, Hamada, and Yatagai, 2011; Yatagai et al., 2012) compiled by the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of Japan, Meteorological Agency (MRI/JMA). Data are available at a spatial resolution of 0.25° X 0.25° for 1951-2007. Precipitation data is from the Monsoon Asia product APHRO_MA_V1101R2, and temperature data is from AphroTemp_V1204R1. This is the only long-term continent-scale daily product that contains a dense network of daily rain-gauge data for Asia including the Himalayas, South and Southeast Asia, and the mountainous areas in the Middle East (Yatagai et al., 2012). We re-scale the gridded weather data to the district level by taking an area-weighted average of grid values in

each district. There is a large degree of variation in the amount of rainfall and the frequency of rainy days during the monsoon season (Fig. S 2.2). Regions in the northwest tend to have lower amounts of both precipitation measures. Regions in the south have lower amounts of total rainfall, but a more even distribution of rain over the monsoon period

*Future Climate Projections*

Climate projections from five general circulation models (GCMs) using representative concentration pathway (RCP) 8.5 and initial condition r1i1p1 were bias-corrected and downscaled for use in this project. These models were CCSM4, GFDL-CM3, GFDL-ESM2G, MIROC-ESM-CHEM, and NorESM1-M. All five models were part of the World Climate Research Programme's Coupled Model Intercomparison Project Phase 5, or CMIP5 (Taylor, Souffer, and Meehl, 2012; Table S 2.1). We chose these models because they (i) come closest to characterizing India's historical monsoon (Menon et al., 2013; Sooraj, Terray, and Mujumdar, 2014), and (ii) demonstrate a wide distribution of future climate changes, including both increases and decreases in monsoon rainfall and the number of rainy days within the monsoon season (Fig. S 2.3). The bias correction method and number of models and initial condition used follows the approach of the Inter-Sectoral Impact Model Intercomparison Project (Warszawaski et al., 2014; Hempel et al., 2013). The bias correction method is recommended for use in hydrology models as it preserves future trends in precipitation and growing degree day variability in addition to their mean trends (Hempel et al., 2013). The climate models were downscaled to the 0.25° spatial resolution of the APHRODITE historical data (Hempel et al., 2013).

*Econometric Model Specification*

In our econometric model, we use both total monsoon rain and the number of rainy days to distinguish between cumulative impacts of rainfall and the associated impacts of its distribution over the months of June-September, which are likely to have different implications for wet and dry season crops (Fig. S 2.2), following past studies (Auffhammer, Ramanathan, and Vincent, 2012; Fishman, 2012). Since historical precipitation and temperature are correlated, omitting temperature means that the coefficient on precipitation will measure the combined effect of both temperature and precipitation (Auffhammer et al., 2013). Therefore to obtain unbiased estimates of the effects of changes in precipitation, we also include growing degree days by season, $GDD$, calculated by using daily mean temperature (Schlenker, 2006). Since irrigation can be applied at any time during the growing season in response to planting decisions, controlling for extent of crop area is necessary. This can help absorb residual variation and generate more precise estimates. However, inclusion of the contemporaneous cropping decision could create a potential source for endogeneity bias, especially if time varying unobservables that impact irrigation decisions also impact planting decisions, or if these decisions occur simultaneously as in the dry season. Additionally, contemporaneous crop area is itself an outcome of weather changes and we would be unable to estimate the true effect of weather on irrigation, due to an over-controlling problem (Dell, Jones, and Olken, 2014). To address this, we use the previous five-year averages of crop area $A_{d,t-1,t-6}$, to eliminate the bias at least contemporaneously. In all Tobit regressions we use cluster-robust SEs allowing for within-district clustering of errors and arbitrary correlation of observations across time. For the ordinary least squares (OLS) regressions, SEs are corrected for spatial and serial correlation (Hsiang, 2010).

*Dependent variable*

The dependent variable used in the econometric model is irrigated area, in 1000s of hectares, for different crops grown in the wet and dry season. It takes on properties of a nonlinear corner solution outcome since many districts report zero irrigated area in a given year, especially for wet season crops and rice grown in the dry season. Of the estimation samples used in our regression models, zeros for irrigated area range from as low as 8% to as high as 67% (wet-season rice: 8%, maize: 22.4%, sorghum: 67%, cotton: 9%, dry- season rice: 16%). A variable with this type of distribution – a variable with a large numbers of zeros with a latent mixing distribution that takes on positive values with positive probability – requires the use of a Tobit model (Wooldridge, 2010). For this type of behavioral model, the zeros reflect a natural outcome from a decision-making process conditional on changes in a set of observed independent variables. Thus, we are interested not only in the properties of $E(y|x)$, but also in $P(y = 0|x)$ which renders a linear OLS estimation inappropriate. Our results from a linear estimation model applied to wet-season crops and dry-season rice are shown in Table S 2.3. We find that the signs and significance of the coefficients are largely consistent with the results from the Tobit models reported in Tables 2.1 and S 2.2 except for the coefficient on wet season growing degree days in the dry- season rice regression. For wheat and barley, however, only 0.25% and 3% of the observations in the samples have zero irrigated areas, so we can ignore the zeros problem and perform standard OLS estimation (applying a Tobit model to a sample with a small number of censored observations is less efficient than simply running OLS). We log transform the dependent variables our econometric models as most of the variables are log-normally distributed, and allow the weather variables to affect irrigated area proportionally. Thus, for the

OLS models we take into account only positive values of the dependent variable. In the tobit models, our dependent variable is of the form $\ln(Y_{dt} + 1)$ to retain year-district observations that have zero values in the estimation sample.

*Time-varying unobservables*

Despite the use of fixed effects and state specific trends, there could be other factors affecting irrigation decisions that may be correlated with weather. In this case, panel data models could still suffer from omitted variables bias. We include an interaction between a set of five-year time dummies and the latitudes and longitudes coordinates for each of our districts to create a smooth spatial function (Banzhaf and Lavery, 2010). This function, which creates a smooth spatial surface at each of our five year increments, captures deviations from each district's long-term time trend and controls for any additional spatiotemporal confounding effects. The results from this model are shown in Table S 2.4. The signs, significance, and magnitude for the average partial effects and the coefficients are largely consistent with Table 2.1 and S 2.2, suggesting that our preferred model is able to capture all confounding factors.

*Panel fixed effects*

A potential problem with including panel level fixed effects in a non-linear model such as Tobit is the incidental parameters problem if the number of panel units goes to infinity and the number of time periods is fixed (Neyman and Scott, 1948). In theory, this can make it difficult to estimate fixed effects consistently, and can affect the consistency of parameters of interest. However Greene (2004) shows, using Monte Carlo simulations, that the bias generally believed to result from such a model is quite limited as long as the number of time periods is greater than

five (Greene, 2004). Since the time dimension in our historical analysis is as large as 35 years, with an equally large number of districts, our estimates reported in Table1 will be consistent and asymptotically efficient. An alternative specification to the fixed effects Tobit model is a more general random effects model that includes averaged values of all the time-varying variables in the model to account for time-invariant district level unobservables (Wooldridge, 2010; Chamberlain, 2004). We apply this correlated random effects model in Table S 2.5 and bootstrap standard errors allowing for within-district clustering to account for potential heteroscedasticity and correlation of observations across time within each cross-sectional unit. The direction and significance of the average partial effects (APE) remains largely similar to Table 2.1 and S 2.2; however significance on the APE for total precipitation in the regression for cotton, and that for wet season degree days in the regression for dry-season rice is lost. Also, unlike before, the regression for wet-season rice picks up significance for the APEs related to number of rainy days and wet season growing degree days

*Lagged dependent variable*

Including the lagged dependent variable in the specification is important since there is substantial serial correlation in irrigation outcomes at the district level that is not accounted for with common trends. In dynamic panel data models with unobserved effects, the treatment of the initial observations is an important theoretical and practical problem. When using short panels, including lags in OLS models biases coefficient estimates (referred to as Nickel bias). In long panels (here T=35), this bias can safely be considered second order as it declines at the rate of 1/T (Nickell, 2981). Similarly in Tobit models, using lags causes an initial condition problem caused by the presence of both the past value of the dependent variable and an unobserved

heterogeneity term in the equation, and the correlation between them. Here too, the impact of the initial conditions diminishes if the number of sample periods T is large (Honoré, Vella, and Verbeek, 2008). In Table S 2.6, we exclude the lag dependent variable and the average crop area over the last five years. The coefficients and average partial effects on the precipitation measures are largely consistent with Table 2.1 in terms of sign and significance; apart from those on rainy days for barley and sorghum which are no longer significant.

*Residual Variation in Weather*

A concern with using fixed effects is that these controls can absorb much of the variation in weather. Table S 2.7 shows the R-square and standard deviation of the residual weather variation not absorbed by fixed effects. These are calculated by running regressions of total monsoon precipitation and rainfall frequency on (1) intercept, (2) year fixed effects (3) district fixed effects (4) district-year fixed effects and (5) district-year-state specific trends, our preferred empirical approach. Including only year fixed effects preserves a significant amount of precipitation variation. When we remove district fixed effects, the remaining variation decreases substantially, suggesting that much of the precipitation variation comes from spatial differences. Including state-specific time trends does not lead to further reduction in the variation. If the variation remaining after removing the fixed effects is as large as the weather changes projected in the climate change models, then we can identify the effects of climate change on irrigated area from the data. Column 3a in Table S 2.7 shows that under NorESM1-M, annual monsoon-season precipitation is projected to increase by 27.5 mm in the RCP 8.5 scenario. The percentage of observations that have a residual greater than this projected change after controlling for our preferred fixed effects is reported in Column 3b as 83%. Across the models, we find a

considerable overlap between variation of precipitation and number of rainy days in our estimation sample, and the projected changes under different climate futures.

***Irrigated area projections: note about cropped area***

We do not project future cropped area, or constrain future irrigated areas to historical cropped areas. However, projected national net irrigated area never exceeds the historical national net cropped area. The largest historical annual net cropped area, as reported by ICRISAT, was 188.95 million hectares. Our projections of national total net irrigated area reach a maximum 110.37 million hectares. At the district level, a minimum of 8 and a maximum of 81 (out of 311) districts are projected to have greater net irrigated area than historical district-level net cropped area (range due to different GCM projected climates). All these districts lie in states with large irrigated areas: Madhya Pradesh (0 – 10 districts; range due to different GCM projected climates), Punjab (6 – 11 districts), Rajasthan (1 – 20 districts), Tamil Nadu (0 – 7 districts), Uttar Pradesh (0 – 24 districts), Andhra Pradesh (0 – 1 districts), Gujarat (0 – 6 districts), Haryana (1 – 7 districts), and Himachal Pradesh (0 – 1 districts). We recognize that in these districts, our projections may be biased upwards due to the already-large irrigated areas and high ratio of historical irrigated to cropped area.

***WBMplus details***

*Model inputs*

WBMplus is a gridded, process-based, hydrology model (Wisser et al., 2010; Grogan et al., 2015), which here uses a 30-arcminute grid resolution with the simulated topological river network STN-30p (Vörösmarty et al., 2000b; Vörösmarty et al., 2000b). The climate inputs –

both historical and future – are the same as used for the econometric analysis and projections. WBMplus simulated the climate for each GCM separately, including a 10-year spin-up with each GCM's historical output (1996-2005). Historical agricultural data (i.e., irrigated area by crop type) is from ICRISAT, and future agricultural data is from the econometric model projections. The field capacity and wilting point of soils is a required input for calculating both crop and non-crop potential evapotranspiration (PET). We use the FAO/UNESCO Soil Map of the World (FAO/UNESCO, 2003) for both, as well as for the soil drainage classes. Soil drainage class is used to estimate water seepage through flooded rice paddies. Large reservoirs are represented as river segments with the reservoir's storage capacity; data on large reservoir capacity, location, and primary purpose are from the GRanD database (Lehner et al., 2011).

*Unmet irrigation water demand*

Quantifying unmet irrigation water demand under a scenario with limited access to unsustainable groundwater required two sets of WBMplus model runs. The first run will be referred to here as WBMplus-Base. WBMplus-Base computes the net irrigation water needed by each crop to avoid water stress (Grogan et al., 2105; Allen et al., 1998). It then computes the daily gross irrigation water supply required, accounting for irrigation inefficiency:

$$I_{gross} = \frac{I_{net}}{E_{eff}} \tag{S 2.1}$$

where $I_{gross}$ (km$^3$) is gross irrigation, $E_{eff}$ (-)is irrigation efficiency, and $I_{net}$ (km$^3$) is net irrigation. In WBMplus-Base, gross irrigation water requirements, $I_{gross}$, are 100% fulfilled. To fulfill gross irrigation water requirements, irrigation water is abstracted from these sources in order: 1) groundwater recharge, $GW_{r-Base}$ (km$^3$), 2) rivers and reservoirs, $RR_{Base}$ (km$^3$), 3) unsustainable groundwater ($UGW_{Base}$) (km$^3$). Note, $UGW_{Base}$ is only used when the sustainable surface water

supplies are not sufficient to fulfill $I_{gross}$ requirements. The sum of abstracted irrigation water in WBMplus-Base, $Abst_{Base}$ (km$^3$), is equivalent to gross irrigation demands:

$$Abst_{Base} = I_{gross} = GW_{r-Base} + RR_{Base} + UGW_{Base} \qquad \text{(S 2.2)}$$

In the second model run, called WBMplus-NoUGW, all model inputs and parameters are the same as WBMplus-Base. Therefore, $I_{net}$ and $I_{gross}$ are equal in WBMplus-Base and WBMplus-NoUGW. The only difference is that for WBMplus-NoUGW, unsustainable groundwater is not available as a water source. In this model run gross irrigation water requirements cannot always be fulfilled because there are only two sources available from which to abstract water: 1) groundwater recharge, $GW_{r,NoUGW}$ (km$^3$), and 2) rivers and reservoirs, $RR_{NoUGW}$ (km$^3$):

$$Abst_{noUGW} = GW_{r,NoUGW} + RR_{NoUGW} \qquad \text{(S 2.3)}$$

where $Abst_{NoUGW}$ (km$^3$) is the water abstracted for irrigation in WBMplus-NoUGW. The difference between $Abst_{Base}$ and $Abst_{NoUGW}$ is defined as the unmet irrigation water demand, $I_{unmet}$ (km$^3$):

$$I_{unmet} = Abst_{Base} - Abst_{NoUGW} \qquad \text{(S 2.4)}$$

This allowed us to quantify the amount of irrigation water demand that would be unmet if unsustainable groundwater resources were unavailable. $I_{unmet}$ was calculated separately for the dry (rabi) season, $I_{unmet,d}$ and the wet (kharif) season, $I_{unmet,w}$. For grid cells in which $I_{unmet} > 0$, we assume there are irrigated crop production losses. Unmet irrigation water demand was assumed to lead to contraction of irrigated areas. As a first-order assumption, we applied the contraction of irrigated area equally to all crops within a grid cell. In the dry season, reduction of irrigated areas is assumed to equal a reduction in crop area, as most dry season crops cannot be grown without irrigation (Kumar et al., 2004). In the wet season, we assumed that without

sufficient irrigation water, farmers would grow the same area of crops, but under rainfed conditions.  Therefore, dry season crop production losses, $P_{loss,d}$ (tons), due to $I_{unmet,d} > 0$ are:

$$P_{loss-d} = \frac{I_{unmet,d}}{I_{gross,d}} * (A_1 * IY_1 + \cdots + A_N * IY_N) \tag{S 2.5}$$

where $A_N$ is the area (ha) of crop $N$,  and $IY_N$ (tons/ha) is the irrigated yield of crop $N$, and $I_{gross,d}$ ($km^3$) is the gross irrigation demand in the dry season.  Wet season crop production losses, $P_{loss,w}$ (tons), due to $I_{unmet} > 0$ are:

$$P_{loss,w} = \frac{I_{unmet,w}}{I_{gross,w}} * (A_1 * (IY_1 - RY_1) + \cdots + A_N * (IY_N - RY_N)) \tag{S 2.6}$$

where $A_N$ is the area (ha) of crop $N$, $IY_N$ (tons/ha) is the irrigated yield of crop $N$, and $RY_N$ (tons/ha) is the rainfed yield of crop $N$, and $I_{gross,w}$ ($km^3$) is the gross irrigation demand in the wet season.  ICRISAT provided only total crop production, without separating irrigated yields from rainfed yields.  Therefore, where both irrigated and rainfed crops were present within a district, we assumed $IY_N = 2$ x $RY_N$ (33).

It is important to note that $I_{unmet} \neq UGW_{Base}$.  Rather, $I_{unmet}$ is larger than $UGW_{base}$.  This difference is entirely due to irrigation inefficiencies, which, if the irrigation water source is UGW, will increase surface water supply. Therefore, $I_{unmet}$ is sensitive to the irrigation efficiency estimate used in WBMplus.  We use the FAO-reported 34% irrigation efficiency for all of India (AQUASTAT, 2008; Guerra et al., 1998; Rohwer, Gerten, and Lucht, 2007).  Inefficient irrigation leads to surface and sub-surface runoff and groundwater recharge from irrigated cropland.  To represent this inefficiency, WBMplus (both –Base and –NoUGW) abstracts water in response to gross irrigation water demands, allows the inefficient portion of this water (gross irrigation minus net irrigation) to evaporate to meet grid-cell evaporative demands, and then the remainder is returned to the groundwater recharge pool and to surface runoff.

$$I_{evap} = \begin{cases} PET_{grid} - AET_{grid} \ if \ \left(I_{gross} - I_{net}\right) \geq PET_{grid} - AET_{grid} \\ I_{gross} - I_{net} \qquad if \ \left(I_{gross} - I_{net}\right) < PET_{grid} - AET_{grid} \end{cases} \qquad \text{(S 2.7)}$$

$$R_{sr} = 0.5 * \left(I_{gross} - I_{net} - I_{evap}\right) \qquad \text{(S 2.8)}$$

$$R_{gr} = 0.5 * \left(I_{gross} - I_{net} - I_{evap}\right) \qquad \text{(S 2.9)}$$

where $I_{evap}$ is evaporation from the inefficient portion of irrigation water, $PET_{grid}$ is the potential evapotranspiration from the entire grid cell, $AET_{grid}$ is the actual evapotranspiration from the entire grid cell, $R_{sr}$ is inefficient irrigation water return to surface runoff, and $R_{gr}$ is inefficient irrigation water return to groundwater recharge. In order to meet gross irrigation water requirements, water is abstracted first from the sustainable sources of rivers, reservoirs, and groundwater recharge. If the sum of these sources is insufficient to meet gross irrigation water requirements, the model will abstract additional water from the unsustainable groundwater (UGW) pool (in WBMplus-Base only). The proportions of UGW and sustainable water that make up $I_{gross}$, $I_{net}$, $I_{evap}$, $R_{sr}$, and $R_{gr}$ are assumed to be equal to the proportion of water abstracted from UGW versus sustainable sources. Whenever UGW is used, some of it is returned to surface runoff and groundwater recharge ($R_{sr}$ and $R_{gr}$). This water can be re-abstracted for irrigation, but this subsequent re-abstraction is counted as water from sustainable sources. In this way, UGW increases the (apparent) sustainable surface water supplies. Therefore, in the simulation in which UGW is unavailable, no UGW enters the surface water system, and the unmet irrigation water demand reflects this loss.

A few recent studies have pointed out that irrigation from wells can have a higher efficiency than from surface water due to reduced conveyance distances and improved timing of water delivery (Biemans et al., 2013). Increased efficiency for groundwater irrigation would decrease our estimate of UGW demand and the estimate of unmet irrigation water demand.

However, since WBMplus always uses sustainable surface water supplies before turning to UGW, which is not always how farmers choose to irrigate (Shah, 2010), our estimates of UGW and unmet irrigation water demand are already conservative.

*NRLP and other inter-basin transfers*

Two types of inter-basin water transfers were modeled by WBMplus (Table S 2.8). The first type is existing transfers for which canals have already been completed: the Kurnool Cuddapah Canal System, the Periyar Project, the Parambikulam Aliyar Project, the Teluga Ganga Project, and the Indira Gandhi Canal. Data for these completed inter-basin transfers is from Ghassemi and White (2007) and Jain, Reddy and Chauge (2005). These completed transfers were implemented in all WBMplus simulations. The second type of inter-basin transfers were only included in the simulation to assess the National River Linking Project (NRLP). These transfers are not yet completed, and have only been proposed by NRLP project planners. They include all links of both the Himalayan and Peninsular components of the NRLP. Data for these inter-basin transfers is from Ghassemi and White (2007), Jain, Reddy and Chauge (2005), and Adhikari, Verhoeven and Iroch (2009).

For all inter-basin transfers (completed and proposed), five parameters are used to model the transfer. These are: the donor/from latitude and longitude, the recipient/to latitude and longitude, a minimum allowed flow, a maximum allowed flow, and a rule for flow volumes between the minimum and maximum (Table S 2.8a,b). In some cases, maximum allowed flow is based on published reported annual transfer capacities (Table S 2.8b). In addition to the reported latitudes and longitudes of the transfers, Table S 2.8a also lists STN latitudes and longitudes; these are the grid cell based locations of each transfer, which in some cases are different than the

reported location because they were adjusted to ensure they linked to the correct rivers within the STN-30p network version 6.02. Any NRLP inter-basin transfer for which latitude and longitude points could not be found are included in Table S 2.8a (starred), but could not be included in the model simulation. The completed transfers were implemented in the year that construction was completed (Table S 2.8b). Proposed NRLP transfers were turned on for the entire future simulation (2006 – 2050), as there is no set date for completing construction of these transfers. The impact of the NRLP on unsustainable groundwater demand was only assessed for the last decade of the simulation, 2040-2050, and so these results are not sensitive to any construction date prior to 2040. The simulation in which new reservoirs were implemented included an additional 10-year spin-up time to allow the reservoirs to fill.

The volume of water transferred through each canal is calculated as:

$$D = \begin{cases} 0 & if\ Q_d \leq Q_{min} \\ (Q_d - Q_{min}) * \frac{P}{100} & if\ Q_{min} > Q_d \geq Q_{max} \\ Q_{max} & if\ Q_d > Q_{max} \end{cases} \tag{S 2.10}$$

where $D$ ($m^3s^{-1}$) is the amount of water diverted through the canal, $Q_d$ ($m^3s^{-1}$) is the donor river discharge, $Q_{min}$ ($m^3s^{-1}$) is the minimum flow parameter, $Q_{max}$ ($m^3s^{-1}$) is the maximum flow parameter, and $P$ is the percent flow parameter.

The transfer volume, $D$, is corrected to $D_{corr}$ for small transfer volumes:

$$D_{corr} = 0\ if\ D < 0.01 \tag{S 2.11}$$

Evaporation from open water along the canals is removed from the transfer volume:

$$D_{corr_e} = \begin{cases} D_{corr} - E\ if\ (D_{corr} - E) > 0.001 \\ 0 \qquad if\ (D_{corr} - E) \leq 0.001 \end{cases} \tag{S 2.12}$$

where $D_{corr\_e}$ ($m^3s^{-1}$) is the transfer volume corrected for evaporation, and $E$ ($m^3s^{-1}$) is the evaporation volume:

$$E = L * W * FWE \tag{S 2.13}$$

where $L$ (m) is the length of the canal (listed in Table S 2.8b where published data is available, or calculated based on a straight line between to/from points), $FWE$ is free-water evaporation (mm/day) which can be calculated through various free-water evaporation models (Dingman, 2002) or by scaled calculated potential evapotranspiration by the Hamon method (Park, 1977); and $W$ (m) is the width of the canal:

$$W = \begin{cases} \tau * D_{corr}^{\varphi} & if \ (\tau * D_{corr}^{\varphi}) \geq 0.01 \\ 0 & if \ (\tau * D_{corr}^{\varphi}) < 0.01 \end{cases} \qquad \text{(S 2.14)}$$

where $\tau$ (8.0) and $\varphi$ (0.58) are held constant (Park, 1977).

Water is transferred on a daily time step. Several of the lengthy inter-basin transfers were split into multiple transfer segments for the purpose of the simulation (Table S 2.8a,b). This allowed for water to be released and/or stored along the canal route, from where it can be accessible for irrigation withdrawals.

We developed two model simulations to assess the potential for the NRLP to alleviate groundwater stress. The first implemented only the NRLP inter-basin transfers, with no additional water storage mechanisms for the water transferred. The plans for locations and capacity of the NRLP transfers have been published (Ghassemi and White, 2007; Gain, Reddy, and Chauge, 2005). This assessment showed that mid-century UGW demand could only be alleviated by 1% - 4% when the inter-basin transfers were implemented alone. This small percentage is primarily due to the transfer rules; since a percentage of donor-river discharge is transferred, the largest water transfers occur in the wet (Kharif) season, while the majority of UGW demand occurs in the dry (Rabi) season. While NRLP plans have included increased water storage (Amarasinghe, Shah, and Malike, 2009), there is insufficient detail on storage methods, locations, or capacities to accurately simulate an NRLP-storage scenario. We chose to develop a scenario that would allow all wet season water transfers to be stored until the dry

season, then be released to optimize supply during periods of high irrigation water demand. This scenario is hypothetical, and should only be interpreted as an upper bound on the potential of the NRLP to alleviate UGW demand. We recognize that there are other, non-irrigation NRLP goals – particularly hydropower generation (Amarasinghe, Shah, and Malike, 2009) – which may require a different water transfer and storage schedule than modeled here. Wet season water transfers were stored by implementing large reservoirs at the donor river location for each transfer (Table S 2.8). Each reservoir has a capacity equal to the 10-year average (1996-2005) simulated wet season transfer volume, and releases water in proportion to the historical irrigation water demand in the region surrounding the reservoir. Reservoir capacity was limited to the capacity of the largest existing irrigation reservoir in India (11 km$^3$, Nagarjuna Reservoir; GRanD database, Lehner et al., 2011). Notably, while the hypothetical NRLP reservoirs in the northwestern states of Punjab, Haryana, and Gujarat are each < 5km$^3$ in capacity, they sum to a total of ~15 km$^3$ of increased reservoir capacity through the region. Simulated historical (1990-2005) UGW demand across these states was ~ 20 km$^3$/year.

In practice, there are many potential problems and impediments to constructing large reservoirs; e.g., cost, the large areas of land needed (or cleared) for the project, the potentially lengthy construction and fill-up time, and the considerable displacement of people. Constructing a larger number of small reservoirs, and even investing in very small on-farm tanks and storage structures (Wisser et al., 2010) may be another way to implement the storage component of the NRLP. Further study of small reservoirs and tanks would improve upon the results shown here.

*Within-basin transfers*

WBMplus represents within-basin water transfer systems by allowing the irrigation water demand of a given grid cell can be fulfilled by surface water not only within the grid cell, but by surface waters stored in grid cells up to 150 km away. This surface water is used before the model uses any unsustainable groundwater.

Irrigation reservoirs are often the starting point of extensive canal systems, and are managed to provide irrigation water. The GrAND database (Lehner et al., 2011) indicates the primary purpose of each reservoir. In WBMplus, reservoirs with "irrigation" as their primary purpose are parameterized to release water as a function of the 30-year average downstream irrigation water demand.

*WBMplus model validation*

FAO-AQUASTAT national statistics were used for national- and large basin-scale validation. We compared AQUASTAT to WBMplus's historical output of the following key variables (Table S 2.9):

1. **Internal renewable surface water resources,** which the FAO defines as the sum of exploitable regular renewable surface water and exploitable irregular renewable surface water. **Exploitable regular renewable surface water** is defined as the annual average quantity of surface water that is available with an occurrence of 90 percent of the time. In practice, it is equivalent to the low water flow of a river. It is the resource that is offered for withdrawal or diversion with a regular flow. **Exploitable irregular renewable surface water** is defined as equivalent to the variable component of water resources (e.g. floods). It

includes the seasonal and inter-annual variations, i.e. seasonal flow or flow during wet years. It is the flow that needs to be regulated. The WBMplus result reported here for comparison is the average annual runoff across the entire country.

2. **Annual primary groundwater withdrawals**, defined as the annual gross amount of water extracted from aquifers. It includes withdrawal of renewable groundwater, water extracted from deep fossil aquifers (non-renewable water) and potential over-abstraction of renewable groundwater. The WBMplus result reported for comparison is the sum of renewable groundwater abstractions and unsustainable groundwater abstractions.

3. **Annual average runoff within major river basins**. Notably, the $R^2$ value of WBMplus versus APHRODITE average annual runoff for the 8 largest river basins is 0.99.

To validate the groundwater storage results from our model, we compare the historical model simulation of unsustainable groundwater (UGW) extraction to historical well level data (Russo, *personal communication*) from eight districts within the state of Punjab (Table S 2.10). These data were collected twice per year, once pre-monsoon and once post-monsoon, from 1973-2003. While we cannot assess two disjoint 30-year periods in the historical analysis as we did for the future projections, we can instead apply the same analysis to two disjoint 10-year periods: 1973-1982 and 1993-2002. We take the decadal average of all well data within a district to be the decadal district-level groundwater level. We use **GWL** to indicate ground water level (well level data), and **UGW** to indicate unsustainable ground water. We use the same categories of interpretation as the future analysis:

**Category 1: Decreased rate of GWL decline (yellow in Figure 2.3)**

This category is defined as:

a) *For model results:* UGW extraction in 1993-2002 is positive, but less than UGW extraction in 1973-1982.

b) *For well level data*: Ground water levels in 1993-2002 are more than 0 mbgl (meters below ground level), but the change in ground water level is less than 10% of the 1973-1982 ground water levels.

**Category 2: Same rate of GWL decline (red in Figure 2.3)**

This category is defined as:

a) *For model results:* UGW extraction in 1993-2002 is more than UGW extraction in 1973-1982, but not more than 50% greater than UGW extraction in 1973-1982.

b) *For well level data*: Ground water levels in 1993-2002 are deeper than in 1973-1982, but the change in ground water level is not more than 50% of the 1973-1982 ground water levels.

**Category 3: Increased rate of GWL decline (dark red in Figure 2.3)**

This category is defined as:

c) *For model results:* UGW extraction in 1993-2002 is more than UGW extraction in 1973-1982, and the change is > 50% of UGW extraction in 1973-1982.

d) *For well level data*: Ground water levels in 1993-2002 are deeper than in 1973-1982, and the change is > 50% of the 1973-1982 ground water levels.

**Category 4: GWL decline begins in the later period (orange in Figure 2.3)**

This category is defined as:

e) *For model results:* UGW extraction in 1973-1982 is 0 $km^3yr^{-1}$, and in 1993-2002 is > 0 $km^3yr^{-1}$.

f) *For well level data*: Ground water levels in 1973-1982 are 0 mgbl, and in 1993-2002 are > 0 mbgl.

**Category 5: GWL recovers/stays static (blue in Figure 2.3)**

This category is defined as:

g) *For model results:* UGW extraction in 1973-1982 is > 0 $km^3yr^{-1}$, and in 1993-2002 = 0 $km^3yr^{-1}$.

h) *For well level data*: Ground water levels in 1973-1982 are > 0 mgbl, and in 1993-2002 = 0 mbgl.

Validation results (Table S11) show that the hydrologic model does well at the level of interpretation described above (i.e., categories), capturing the range of increased to continued rates of groundwater level declines, as well as recovery/static groundwater levels. Notably, the model under-estimates rates of decline in two of the three districts in which the data show that rates of decline increase: in both Amritsar and Gurdaspur, data show increased rates of GWL decline while the hydrologic model shows only continued GWL declines. This disagreement is due to the conservative definition of "increased decline rates" (category 3). The model shows that in Amritsar UGW extraction is 23% higher in 1993-2002 than in 1973-1982, and in Gurdaspur UGW extraction is 14% higher in 1993-2002 than in 1973-1982. These rates do not meet our conservative definition of 50% higher extraction in the latter period, but are consistent with the well level data.

**Tables**

**Table 2.1.** Regression estimates on weather-related variables in agro-economic model

| | Wet Season | | | | Dry Season | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | **Rice** | **Maize** | **Sorghum** | **Cotton** | **Rice** | **Wheat** | **Barley** |
| No. of rain days | 0.03 | -0.327*** | -0.054+ | -0.199*** | -0.003 | -0.005 | -0.180* |
| | (0.038) | (0.087) | (0.030) | (0.000) | (0.032) | (0.071) | (0.091) |
| Rainfall JJAS | 0.070*** | -0.014 | -0.039*** | 0.045*** | 0.166*** | 0.250*** | -0.026 |
| | (0.020) | (0.019) | (0.011) | (0.000) | (0.023) | (0.034) | (0.043) |
| Wet GDD | -0.048 | -0.099* | 0.027 | -0.310*** | 0.301** | 0.089 | 0.066 |
| | (0.040) | (0.039) | (0.027) | (0.000) | (0.115) | (0.069) | (0.110) |
| Dry GDD | | | | | -0.316 | -0.159 | 0.068 |
| | | | | | (0.243) | (0.099) | (0.184) |
| Model | Tobit | Tobit | Tobit | Tobit | Tobit | OLS | OLS |
| N | 8248 | 7244 | 5178 | 3244 | 3770 | 7460 | 3882 |
| 0 Observations | 632 | 1628 | 3520 | 277 | 586 | 19 | 128 |

Average partial effects for Tobit models and coefficient estimates for ordinary least square (OLS) models represent the effect of monsoon precipitation (June-September), rainfall distribution (number of rainy days in June-September), and seasonal growing degree days (GDD) on irrigated area decisions, 1970-2005. A Tobit model is used for crops where a large fraction of the observations are clustered at 0. Standard errors in parentheses are clustered at the district level for the Tobit models and corrected for spatial and serial correlation for OLS models. The dependent variable is the logarithm of district-level crop irrigated area in 1000 ha. All regressions include district and year fixed effects, and state-specific trends to control for time-invariant district characteristics, country-wide trends, and time-varying differences between states. See S1-S4 for data sources, results for non-weather variables and robustness checks. Statistical significance is given by + $p<0.10$ * $p<0.05$ ** $p<0.01$ ***$p<0.001$

**Table 2.2** The impact of unsustainable groundwater (UGW) on irrigated agriculture and food supply.
* 2,000 kcal per day diet.

| | Irrigated agriculture production [million tons] | Production loss in absence of UGW [million tons], (% of total production) | Calorie loss in absence of UGW [billion kcal] | # People fed by UGW-dependent calories* [millions] |
|---|---|---|---|---|
| Dry Season | 75.4 | 38.7 (51%) | 121,300 | 166 |
| Wet Season | 73.4 | 2.8 (4%) | 4,750 | 7 |
| Annual total | 148.8 | 41.5 (28%) | 126,000 | 173 |
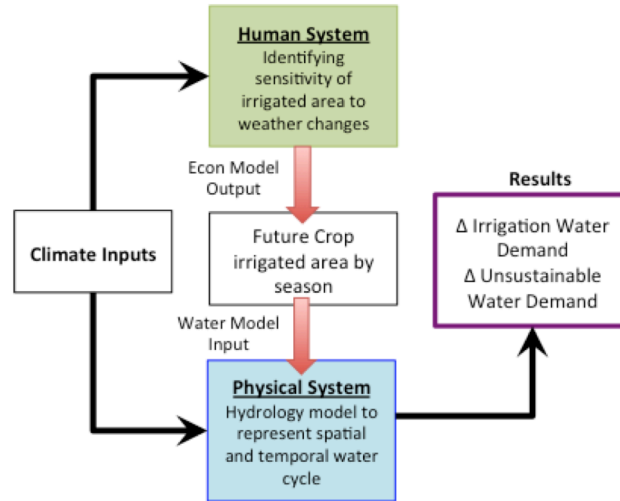
**Figure. 2.1:** Coupled human- and physical- system model schematic. Human system analysis uses an econometric model to identify historical irrigated area sensitivity to weather changes. Climate drivers from 5 GCMs are combined with historical regression estimates to project future crop-wise irrigated areas. These projections, and the climate drivers, are inputs to a physically-based hydrology model, which simulates future irrigation water demand and unsustainable groundwater demand.
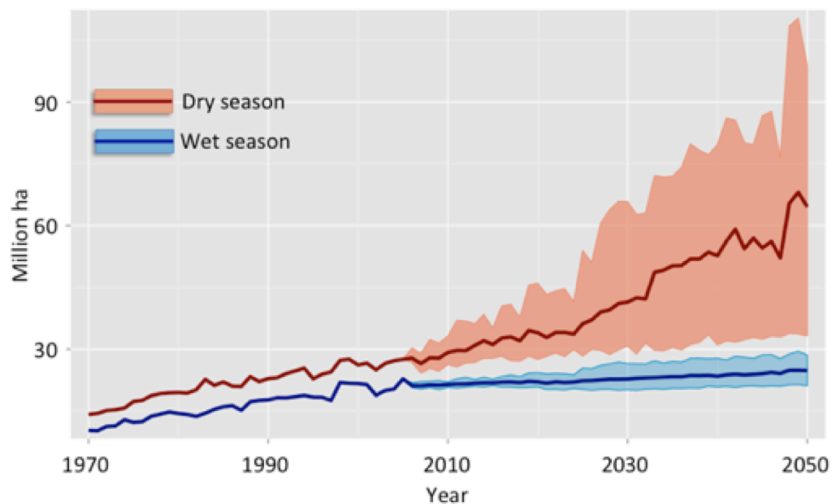


**Figure. 2.2:** Econometric model-projected aggregate dry season (red) and wet season (blue) irrigated areas. Historical period (1970-2005) data is from ICRISAT. Future period (2006-2050) solid line is the multi-model mean of projections based on 5 GCM climate futures, with a shaded range of uncertainty due to GCM differences.
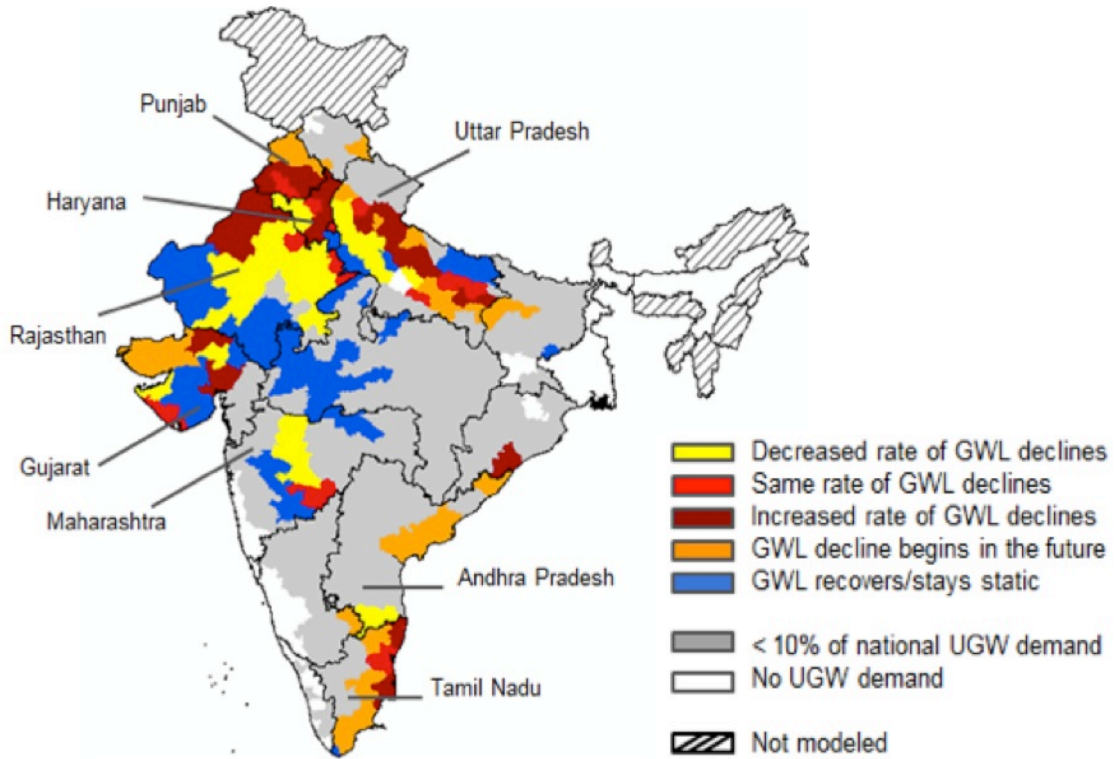
**Figure. 2.3:** Trends in district-level ground water levels (GWL) between 1979-2000 and 2029-2050, inferred from the multi-model mean of changing need for unsustainable groundwater (UGW) to meet irrigation water needs. Decreases in UGW demand will slow down GWL declines (yellow); continued demand will lead to continued GWL declines (red); increased demand will increase GWL declines (dark red); new positive demands can start GWL declines (orange); demand going to 0 can allow GWL to recover (blue). Black lines are state boundaries . Colored (non-grey) regions account for 90% of future modeled national UGW demand. Fig. S6 shows trends in GWL for 5 individual GCM climate futures.
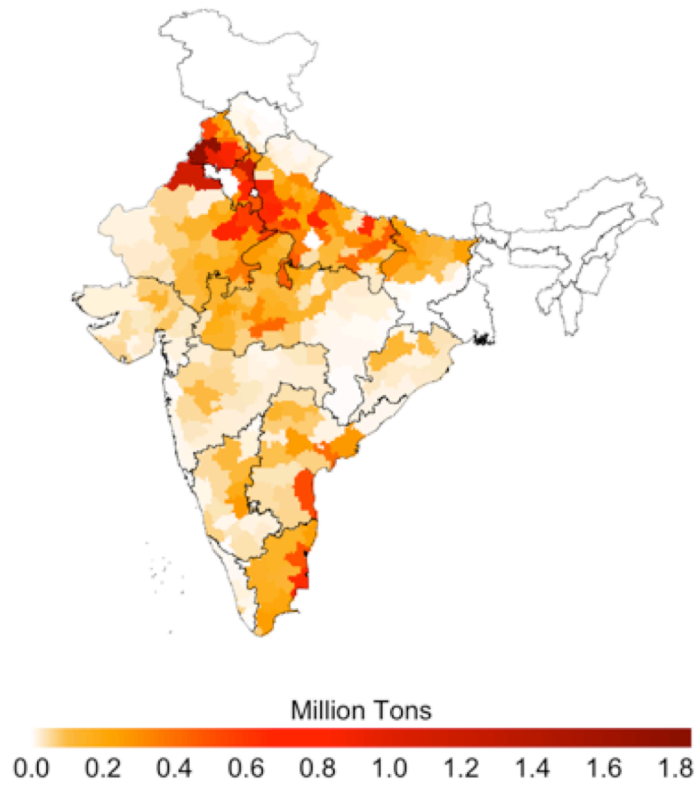
**Figure. 2.4:** District-level reduction in current (c. year 2000) annual crop production, in million metric tons, that would occur if unsustainable groundwater supplies became unavailable. Black lines are state boundaries.
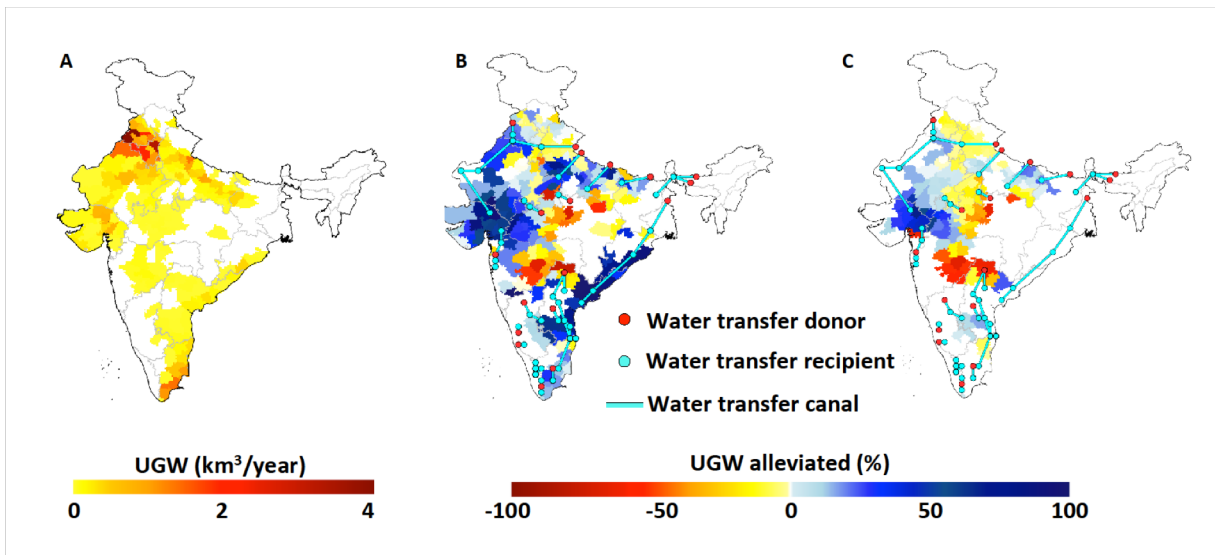
**Figure. 2.5: A)** Mid-century annual unsustainable groundwater (UGW) demand at the district level. The National River Linking Project (NRLP) is a proposed solution for alleviating this demand. **B)** and **C)** Light blue lines: NRLP water transfer canals; red dots: water donor locations; blue dots: water recipient locations. Blue dots along chained canals are both receiving and donating. **B)** The % of each district's mid-century UGW demand that could be alleviated with the implementation of NRLP canals and construction of new reservoirs along canal routes. Blue: UGW demand is alleviated; yellow and red: UGW demand is worsened. National total UGW alleviation is 16%. **C)** The % of each district's mid-century UGW demand that could be alleviated with the implementation of NRLP canals only. National UGW alleviation is 1-4% with transfers only. Gray lines are state boundaries. **B)** and **C)** share a scale bar.

**Table S 2.1:** GCMs contributing to the present analysis along with institutions that provided model output

| Modeling Center (or Group) | Institute ID | Model Name |
|---|---|---|
| National Center for Atmospheric Research | NCAR | CCSM4 |
| NOAA Geophysical Fluid Dynamics Laboratory | NOAA GFDL | GFDL-CM3 GFDL-ESM2G |
| Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies | MIROC | MIROC-ESM-CHEM |
| Norwegian Climate Centre | NCC | NorESM1-M |

**Table S 2.2** Full regression results for agro-economic model in Table 2.1

| | Wet Season | | | | Dry Season | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | **Rice** | **Maize** | **Sorghum** | **Cotton** | **Rice** | **Wheat** | **Barley** |
| No. of rain days | 0.03 | -0.327*** | -0.054+ | -0.199*** | -0.003 | -0.005 | -0.180* |
| | (0.038) | (0.087) | (0.030) | (0.000) | (0.032) | (0.071) | (0.091) |
| Rainfall JJAS | 0.070*** | -0.014 | -0.039*** | 0.045*** | 0.166*** | 0.250*** | -0.026 |
| | (0.020) | (0.019) | (0.011) | (0.000) | (0.023) | (0.034) | (0.043) |
| Kharif degree days | -0.048 | -0.099* | 0.027 | -0.310*** | 0.301** | 0.089 | 0.066 |
| | (0.040) | (0.039) | (0.027) | (0.000) | (0.115) | (0.069) | (0.110) |
| Rabi degree days | | | | | -0.316 | -0.159 | 0.068 |
| | | | | | (0.243) | (0.099) | (0.184) |
| Lag log irrigated area | 0.664*** | 0.383*** | 0.228*** | 0.572*** | 0.501*** | 0.686*** | 0.705*** |
| | (0.023) | (0.042) | (0.014) | (0.000) | (0.034) | (0.033) | (0.041) |
| Log Previous 5 yr avg crop area | 0.107*** | 0.099*** | 0.034*** | 0.128*** | 0.100** | 0.071+ | 0.146*** |
| | (0.024) | (0.019) | (0.009) | (0.000) | (0.032) | (0.036) | (0.033) |
| Model | Tobit | Tobit | Tobit | Tobit | Tobit | OLS | OLS |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 8248 | 7244 | 5178 | 3244 | 3770 | 7460 | 3882 |
| N left censored at zero | 632 | 1628 | 3520 | 277 | 586 | 19 | 128 |
| Log likelihood | -1171.7091 | -2700.00 | -1300.00 | -1800.00 | 740.535 | | |
| R-sq | | | | | | 0.990 | 0.990 |
| Pseudo-R-sq | 0.929 | 0.759 | 0.772 | 0.692 | 1.108 | | |

Notes: Table S 2.1 reports all average partial effects for Tobit models and coefficient estimates for OLS models from 7 separate crop regressions reported in Table 1. District-level data are from ICRISAT and APHRODITE for years 1970-2005. A Tobit model is used for crops where a large fraction of the observations are clustered at 0. Standard errors reported in parentheses are clustered at the district level for the Tobit models and corrected for spatial and serial correlation for OLS models. All variables are in natural logarithms. Statistical significance is given by + p<0.10 * p<0.05 ** p <0.01 ***p < 0.001.

**Table S 2.3:** Regression results using linear estimation for wet season crops and dry season rice

| | Wet Season | | | | Dry Season |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | **Rice** | **Maize** | **Sorghum** | **Cotton** | **Rice** |
| No. of rain days | 0.021 | -0.527* | -0.303 | -0.126 | 0.049 |
| | (0.097) | (0.223) | (0.274) | (0.133) | (0.088) |
| Rainfall JJAS | 0.080* | -0.151+ | -0.317** | 0.114+ | 0.275*** |
| | (0.041) | (0.079) | (0.106) | (0.059) | (0.043) |
| Kharif degree days | -0.133 | -0.003 | 0.798 | -0.338** | -0.024 |
| | (0.115) | (0.274) | (0.636) | (0.112) | (0.357) |
| Rabi degree days | | | | | -0.565 |
| | | | | | (0.424) |
| Lag log irrigated area | 0.687*** | 0.451*** | 0.566*** | 0.716*** | 0.665*** |
| | (0.025) | (0.038) | (0.029) | (0.024) | (0.037) |
| Log Previous 5 yr avg crop area | 0.226*** | 0.341*** | 0.135* | 0.118*** | 0.208** |
| | (0.035) | (0.037) | (0.064) | (0.025) | (0.069) |
| Model | OLS | OLS | OLS | OLS | OLS |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| District fixed effects | Yes | Yes | Yes | Yes | Yes |
| State specific trends | Yes | Yes | Yes | Yes | Yes |
| N | 7521 | 5298 | 1848 | 2901 | 3171 |
| R-sq | 0.998 | 0.991 | 0.988 | 0.997 | 0.999 |

Notes: Table S 2.2 reports all coefficient estimates and SEs from 5 separate regressions for crops grown in the wet season and for rice grown in the dry season. District-level data are from ICRISAT and APHRODITE for years 1970-2005. Standard errors reported in parentheses are corrected for spatial and serial correlation. All variables are in natural logarithms. Statistical significance is given by + p<0.10 * p<0.05 ** p <0.01 ***p < 0.001.

**Table S 2.4**: Regression results after accounting for additional time-varying unobservables

| | Wet Season | | | | Dry Season | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | **Rice** | **Maize** | **Sorghum** | **Cotton** | **Rice** | **Wheat** | **Barley** |
| No. of rain days | -0.015 | -0.323*** | -0.074* | -0.172*** | -0.013 | -0.051 | -0.225* |
| | (0.039) | (0.083) | (0.031) | (0.000) | (0.031) | (0.070) | (0.094) |
| Rainfall JJAS | 0.086*** | -0.014 | -0.030** | 0.048*** | 0.169*** | 0.253*** | -0.017 |
| | (0.020) | (0.020) | (0.011) | (0.000) | (0.023) | (0.034) | (0.044) |
| Kharif degree days | -0.065 | -0.092* | 0.023 | -0.275*** | 0.313** | 0.085 | 0.065 |
| | (0.040) | (0.039) | (0.025) | (0.000) | (0.121) | (0.072) | (0.113) |
| Rabi degree days | | | | | -0.281 | -0.126 | 0.078 |
| | | | | | (0.257) | (0.105) | (0.188) |
| Lag log irrigated area | 0.664*** | 0.385*** | 0.228*** | 0.571*** | 0.500*** | 0.686*** | 0.707*** |
| | (0.023) | (0.041) | (0.014) | (0.001) | (0.035) | (0.032) | (0.040) |
| Log Previous 5 yr avg crop area | 0.105*** | 0.097*** | 0.035*** | 0.125*** | 0.103** | 0.070* | 0.143*** |
| | (0.024) | (0.019) | (0.009) | (0.000) | (0.032) | (0.035) | (0.033) |
| Model | Tobit | Tobit | Tobit | Tobit | Tobit | OLS | OLS |
| Spatial variables | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 8248 | 7244 | 5718 | 3244 | 3770 | 7460 | 3882 |
| 0 observations | 632 | 1628 | 3520 | 277 | 586 | | |
| R-sq | | | | | | 0.999 | 0.998 |
| Log likelihood | -1100.00 | -2700.00 | -1300.0000 | -1800.00 | 803.028 | | |
| Pseudo-R-sq | 0.932 | 0.763 | 0.778 | 0.702 | 1.110 | | |

Notes: Table S 2.3 reports all average partial effects for Tobit models and coefficient estimates for OLS models from 7 separate crop regressions. District-level data are from ICRISAT and APHRODITE for years 1970-2005. A Tobit model is used for crops where a large fraction of the observations are clustered at 0. Standard errors reported in parentheses are clustered at the district level for the Tobit models and corrected for spatial and serial correlation for OLS models. All variables are in natural logarithms. Spatial variables indicate presence of two-way and three-way interactions of five yearly time dummies with latitude and longitude. Statistical significance is given by + p<0.10 * p<0.05 ** p <0.01 ***p < 0.001.

**Table S5:** Regression results using correlated random effects Tobit

| | Wet Season | | | | Dry Season |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | **Rice** | **Maize** | **Sorghum** | **Cotton** | **Rice** |
| No. of rain days | 0.089** | -0.284** | 0.014 | -0.218** | 0.015 |
| | (0.034) | (0.095) | (0.032) | (0.075) | (0.038) |
| Rainfall JJAS | 0.051* | -0.015 | -0.067*** | 0.055 | 0.167*** |
| | (0.020) | (0.016) | (0.012) | (0.034) | (0.023) |
| Kharif degree days | -0.115** | -0.200** | 0.023 | -0.170** | 0.141 |
| | (0.042) | (0.072) | (0.027) | (0.054) | (0.103) |
| Rabi degree days | | | | | -0.319 |
| | | | | | (0.218) |
| Lag log irrigated area | 0.835*** | 0.492*** | 0.289*** | 0.695*** | 0.592*** |
| | (0.023) | (0.039) | (0.015) | (0.040) | (0.033) |
| Log Previous 5 yr avg crop area | 0.024 | 0.057*** | 0.014 | 0.085** | 0.059* |
| | (0.022) | (0.015) | (0.010) | (0.028) | (0.027) |
| Model | Tobit | Tobit | Tobit | Tobit | Tobit |
| Means ofTime varying vars | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| District fixed effects | Yes | Yes | Yes | Yes | Yes |
| State specific trends | Yes | Yes | Yes | Yes | Yes |
| N | 8248 | 7244 | 5718 | 3244 | 3770 |
| 0 observations | 632 | 1628 | 3520 | 277 | 586.000 |
| Log likelihood | -1800 | -3300.00 | -1700.00 | -2000.00 | 460.093 |

Notes: Table S 2.4 reports all average partial effects for correlated random effect Tobit models from 5 separate regressions for crops grown in the wet season and for rice grown in the dry season. District-level data are from ICRISAT and APHRODITE for years 1970-2005. A Tobit model is used for crops where a large fraction of the observations are clustered at 0. Bootstrapped Standard errors are given in parenthesis and are clustered by district, based on 100 replications. All variables are in natural logarithms. Statistical significance is given by + $p<0.10$ * $p<0.05$ ** $p<0.01$ ***$p < 0.001$.

**Table S 2.6:** Regression results without additional controls

| | Wet Season | | | | Dry Season | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | **Rice** | **Maize** | **Sorghum** | **Cotton** | **Rice** | **Wheat** | **Barley** |
| No. of rain days | -0.059 | -0.234** | 0.037 | -0.264*** | -0.073 | -0.122 | -0.099 |
| | (0.067) | (0.083) | (0.039) | (0.000) | (0.054) | (0.094) | (0.198) |
| Rainfall JJAS | 0.083*** | -0.025 | -0.057*** | 0.076*** | 0.146*** | 0.284*** | -0.090 |
| | (0.025) | (0.022) | (0.015) | (0.000) | (0.026) | (0.046) | (0.077) |
| Kharif degree days | 0.095 | 0.036 | 0.117* | -0.150*** | 0.489** | -0.017 | 0.143 |
| | (0.091) | (0.053) | (0.055) | (0.000) | (0.179) | (0.123) | (0.176) |
| Rabi degree days | | | | | 0.212 | -0.178 | -0.035 |
| | | | | | (0.479) | (0.128) | (0.221) |
| Model | Tobit | Tobit | Tobit | Tobit | Tobit | OLS | OLS |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 8613 | 7544 | 5903 | 3359 | 3989 | 7739 | 4054 |
| 0 observations | 665 | 1723 | 3649 | 289.000 | 617.000 | 19 | 128 |
| R-sq | | | | | | 0.999 | 0.994 |
| Log likelihood | -4850.002 | -4200.00 | -2000.0000 | -2800.00 | -899.256 | | |
| Pseudo-R-sq | 0.7174 | 0.640 | 0.651 | 0.536 | 0.884 | | |

Notes: Table S 2.5 reports all average partial effects for Tobit models and coefficient estimates for OLS models from 7 separate crop regressions. District-level data are from ICRISAT and APHRODITE for years 1970-2005. A Tobit model is used for crops where a large fraction of the observations are clustered at 0. Standard errors reported in parentheses are clustered at the district level for the Tobit models and corrected for spatial and serial correlation for OLS models. All variables are in natural logarithms. Statistical significance is given by + $p<0.10$ * $p<0.05$ ** $p<0.01$ ***$p < 0.001$.

**Table S 2.7:** Residual variation in weather

**Panel A**

| | Monsoon precipitation(mm) | | Frequency of rainy days | |
|---|---|---|---|---|
| | 1a | 1b | 2a | 2b |
| | R sq | SD of residual | R sq | SD of residual |
| No FE | | 428.9242 | | 18.01787 |
| Year Fe | 0.0346 | 421.4372 | 0.0998 | 17.09494 |
| District FE | 0.8238 | 180.0324 | 0.7336 | 9.29931 |
| District FE, Year FE | 0.8539 | 161.3846 | 0.8334 | 7.353272 |
| District FE, Year FE, state specific trend | 0.8553 | 160.5031 | 0.8365 | 7.285278 |

**Panel B**

| | Monsoon precipitation(mm) | | Frequency of rainy days | |
|---|---|---|---|---|
| | 3a | 3b | 4a | 4b |
| | Projected change | % observations | Projected change | % observations |
| NorESM1-M | 27.5 | 83 | -1.4 | 77 |
| MIROC-ESM_CHEM | 82.31 | 54 | 3.55 | 56 |
| CCSM4 | 50.89 | 70 | -0.13 | 99 |
| GFDL-CM3 | -15.69 | 92 | -1.02 | 88 |
| GFDL-ESM2G | -53.87 | 69 | -8.4 | 25 |

Notes: Variation of monsoon precipitation and number of rainy days absorbed by fixed effects. Panel **A** summarizes regressions of monsoon precipitation and no. of rainy days on various sets of fixed effects. Columns (a) report the R-square of the regression and Columns (b) report the standard deviation of the residuals (remaining monsoon precipitation, and no. of rainy days variation) in mm and days. Panel **B** reports the percentage of observations (Columns (b)) with absolute value of residuals greater than the projected change in precipitation and number of rainy days between 1970-79 and 2040-29 as shown in Columns (a).

**Table S 2.8a:** Project names and locations of India's National River Linking Project canals. Columns "From" and "To" indicate geographic location of canal starting and ending points; columns "From Lat.", "From Long.", "To Lat.", and "To Long." Indicated the latitude and longitude closest to the known starting and ending locations of the canals; columns beginning with "STN" are the latitude and longitude of canal beginning and ending points as they fit onto the half-degree spatial resolution of the Simulated Topological Network STN-30p used by WBM.

| Project Name | From | To | From Lat. | From Long. | To Lat. | To Long. | STN From Lat. | STN From Long. | STN To Lat. | STN To Long. |
|---|---|---|---|---|---|---|---|---|---|---|
| *Completed Inter-basin transfers* | | | | | | | | | | |
| Kurnool Cuddapah Canal System | Krishna Basin | Pennar Basin (Srisailam Dam) | 16.00 | 78.32 | 14.61 | 78.73 | 15.75 | 78.25 | 14.75 | 78.75 |
| Parambikulam Aliyar Project - West to East Combined | Multiple Reservoirs | Aliyar Reservoir | 10.25 | 76.75 | 10.75 | 77.25 | 10.25 | 76.75 | 10.75 | 77.25 |
| Parambikulam Aliyar Project - Northwest Combined | Aliyar Reservoir | Northwest | 10.25 | 76.75 | 10.75 | 76.75 | 10.25 | 76.75 | 10.75 | 76.75 |
| Teluga Ganga Project | Krishna R. in Andhra Pradesh (Srisailam Dam) | Chennai (Madras) (Poondi Reservoir) | 16.09 | 78.90 | 13.19 | 79.86 | 16.25 | 78.75 | 13.25 | 79.75 |
| Teluga Ganga Project | Krishna R. in Andhra Pradesh (Srisailam Dam) | Somasila Reservoir | 16.09 | 78.90 | 14.47 | 79.26 | 16.25 | 78.75 | 14.75 | 79.25 |
| Teluga Ganga Project | Somasila Reservoir | Kandaleru Reservoir | 14.47 | 79.26 | 14.35 | 79.60 | 14.75 | 79.25 | 14.25 | 79.75 |
| Indira Gandhi Canal Project | Harrike Barrage at Ravi & Beas R. junction | Canal junctions at Bhullar | 31.15 | 74.95 | 30.40 | 74.60 | 31.25 | 74.75 | 30.25 | 74.75 |
| Indira Gandhi Canal Project | Canal junctions at Bhullar | Canal junctions at Tibbi | 30.40 | 74.60 | 29.56 | 74.57 | 30.25 | 74.75 | 29.75 | 74.75 |
| Indira Gandhi Canal Project | Canal junctions at Tibbi | Thar Desert | 29.56 | 74.57 | 27.47 | 71.68 | 29.75 | 74.75 | 27.25 | 71.75 |
| Indira Gandhi Canal Project | Thar Desert | Ramgarh | 27.47 | 71.68 | 27.41 | 70.47 | 27.25 | 71.75 | 27.25 | 70.25 |

98

## Proposed NRLP Inter-basin transfers

| Link | Location | River | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| NRLP, Himalayan Component, Kosi-Mechi Link - 1* | Kosi | Mechi | | | | | | | | |
| NRLP, Himalayan Component, Kosi-Ghagra Link - 2 | Kosi in Nepal | Ghagra | 26.87 | 86.81 | 26.09 | 84.09 | 26.75 | 86.75 | 26.25 | 84.25 |
| NRLP, Himalayan Component, Gandak-Ganga Link - 3 | Gandak in Nepal | Ganga | 27.92 | 83.48 | 25.76 | 81.40 | 27.75 | 83.25 | 25.75 | 81.25 |
| NRLP, Himalayan Component, Ghagra-Yamuna Link - 4 | Ghagra in Nepal | Yamuna | 28.93 | 80.99 | 26.85 | 78.80 | 28.75 | 80.75 | 26.75 | 78.75 |
| NRLP, Himalayan Component, Sarda-Yamuna Link - 5 | Sarda River near Nepal border | Yamuna River | 29.01 | 80.10 | 29.19 | 77.14 | 29.25 | 80.25 | 29.25 | 77.25 |
| NRLP, Himalayan Component, Yamuna-Rajasthan Link - 6 | Yamuna | Rajasthan (use junction to existing at Tibbi | 29.19 | 77.14 | 29.56 | 74.57 | 29.25 | 77.25 | 29.75 | 74.75 |
| NRLP, Himalayan Component, Rajasthan-Sabaramati Link - 7 | Rajasthan (use existing canal at Ramarh) | Sabaramati | 27.41 | 70.47 | 22.57 | 72.44 | 27.25 | 70.25 | 23.75 | 72.75 |
| NRLP, Himalayan Component, Chunar-Stone Barrage Link - 8* | Chunar | Stone Barrage | | | | | | | | |
| NRLP, Himalayan Component, Sone Dam-Southern Tribs of the Ganga Link - 9* | Stone Dam | Southern Tribs of the Ganga Link | | | | | | | | |
| NRLP, Himalayan Component, Manas-Ganga Link - 10a | Manas and Sankosh Rivers at Bhutan border | Tista River | 26.76 | 90.94 | 26.78 | 88.57 | 26.75 | 90.75 | 26.75 | 88.75 |
| NRLP, Himalayan Component, Manas-Ganga Link - 10b | Tista River | Ganga River | 26.78 | 88.57 | 25.38 | 87.47 | 26.75 | 88.75 | 25.25 | 87.25 |

| NRLP Link | Location | River | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| NRLP, Himalayan Component, Jogighopa-Frakka Link - 11 | Jogighopa | Tista River | 26.15 | 90.26 | 26.78 | 88.57 | 26.25 | 90.25 | 26.75 | 88.75 |
| NRLP, Himalayan Component, Farakka-Sunderbans Link - 12* | Farakka | Sunderbans | | | | | | | | |
| NRLP, Himalayan Component, Ganga (Farakka)-Subernarekha Link - 13 | Ganga (Farakka) | Subernarekha | 24.80 | 87.91 | 22.22 | 86.71 | 24.75 | 88.25 | 22.25 | 86.75 |
| NRLP, Himalayan Component, Subernarekha-Mahanadi Link - 14 | Subernarekha | Mahanadi | 22.22 | 86.71 | 20.38 | 85.16 | 22.25 | 86.75 | 20.25 | 85.25 |
| NRLP, Peninsular Components, Mahanadi-Godavari Link - 1 | Mahanadi | Godavari | 20.38 | 85.16 | 17.25 | 81.66 | 20.25 | 85.25 | 17.25 | 81.75 |
| NRLP, Peninsular Components, Godavari-Krishna Link - 3 | Godavari | Tributary of Krishna River | 18.85 | 79.24 | 17.35 | 79.46 | 18.75 | 79.25 | 17.25 | 79.25 |
| NRLP, Peninsular Components, Godavari-Krishna Link - 2 | Godavari | Tributary of Krishna River | 18.85 | 79.24 | 16.80 | 78.23 | 18.75 | 79.25 | 16.75 | 78.25 |
| NRLP, Peninsular Components, Godavari-Krishna Link - 4 | Godavari | Krishna River - lower | 17.44 | 81.54 | 16.53 | 80.57 | 17.25 | 81.75 | 16.25 | 80.75 |
| NRLP, Peninsular Components, Krishna-Pennar Link - 5 - SPLIT | Almatti Dam, Krishna R. (upper) | Tungabhadra Reservoir | 16.33 | 75.89 | 15.25 | 76.25 | 16.25 | 75.75 | 15.25 | 76.25 |
| NRLP, Peninsular Components, Krishna-Pennar Link - 6* | Bhima R. (Krishna trib.) | Pennar | | | | | | | | |
| NRLP, Peninsular Components, Krishna-Pennar Link - 7*** | Krishna | Krishna (above existing link to Pennar) | 16.80 | 78.23 | 16.14 | 78.84 | 16.75 | 78.25 | 16.25 | 78.75 |
| NRLP, Peninsular Components, Pennar-Palar, Cauvery Link - 8 | Pennar (Poondi Lake to complete string) | Palar, Cauvery, Thiruchirapalli area | 13.21 | 79.88 | 10.86 | 78.69 | 13.25 | 79.75 | 10.75 | 78.75 |

100

| Link | Location 1 | Location 2 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| NRLP, Peninsular Components, Cauvery-Vaigai, Gundar Link - 9a | Cauvery | Vaigai at Madurai | 10.96 | 78.21 | 9.93 | 78.10 | 10.75 | 78.25 | 9.75 | 78.25 |
| NRLP, Peninsular Components, Cauvery-Vaigai, Gundar Link - 9b*** | Vaigai at Madurai | Gundar | 9.93 | 78.10 | 9.79 | 78.05 | 9.75 | 78.25 | 9.75 | 78.25 |
| NRLP, Peninsular Components, Ken-Betwa Link - 10 | Ken River | Betwa River | 24.64 | 79.86 | 25.37 | 78.65 | 24.75 | 79.75 | 25.25 | 78.75 |
| NRLP, Peninsular Components, Parbati-Chambal Link - 11a | Parbati River | Kalsindh River | 23.79 | 77.21 | 24.02 | 76.25 | 23.75 | 77.25 | 24.25 | 76.25 |
| NRLP, Peninsular Components, Parbati-Chambal Link - 11b | Kalsindh River | Chambal | 24.02 | 76.25 | 24.52 | 75.59 | 24.25 | 76.25 | 24.75 | 75.75 |
| NRLP, Peninsular Components, Par-Narmada Link - 12a | Par River | Tapi (Ukal Reservoir | 20.49 | 73.07 | 21.23 | 73.68 | 20.25 | 73.25 | 21.25 | 73.75 |
| NRLP, Peninsular Components, Par-Narmada Link - 12b | Tapi (Ukal Reservoir | Narmada River | 21.23 | 73.68 | 22.23 | 73.62 | 21.25 | 73.75 | 22.25 | 73.75 |
| NRLP, Peninsular Components, Damanganga-Pinjal Link - 13 | Damanganga | Pinjal | 20.18 | 73.10 | 19.81 | 73.27 | 20.25 | 73.25 | 19.75 | 73.25 |
| NRLP, Peninsular Components, Bedti-Varda Link - 14 | Bedti | Varda | 14.16 | 74.86 | 14.39 | 75.37 | 13.75 | 75.25 | 14.25 | 75.25 |
| NRLP, Peninsular Components, Netravati-Hemavati Link - 15 | Netravati River | Hemavati River | 12.82 | 75.64 | 12.82 | 76.05 | 12.75 | 75.25 | 12.75 | 75.75 |

| Link | River 1 | River 2 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| NRLP, Peninsular Components, Cauvery-Vaigai, Gundar Link - 9a | Cauvery | Vaigai at Madurai | 10.96 | 78.21 | 9.93 | 78.10 | 10.75 | 78.25 | 9.75 | 78.25 |
| NRLP, Peninsular Components, Cauvery-Vaigai, Gundar Link - 9b*** | Vaigai at Madurai | Gundar | 9.93 | 78.10 | 9.79 | 78.05 | 9.75 | 78.25 | 9.75 | 78.25 |
| NRLP, Peninsular Components, Ken-Betwa Link - 10 | Ken River | Betwa River | 24.64 | 79.86 | 25.37 | 78.65 | 24.75 | 79.75 | 25.25 | 78.75 |
| NRLP, Peninsular Components, Parbati-Chambal Link - 11a | Parbati River | Kalsindh River | 23.79 | 77.21 | 24.02 | 76.25 | 23.75 | 77.25 | 24.25 | 76.25 |
| NRLP, Peninsular Components, Parbati-Chambal Link - 11b | Kalsindh River | Chambal | 24.02 | 76.25 | 24.52 | 75.59 | 24.25 | 76.25 | 24.75 | 75.75 |
| NRLP, Peninsular Components, Par-Narmada Link - 12a | Par River | Tapi (Ukal Reservoir | 20.49 | 73.07 | 21.23 | 73.68 | 20.25 | 73.25 | 21.25 | 73.75 |
| NRLP, Peninsular Components, Par-Narmada Link - 12b | Tapi (Ukal Reservoir | Narmada River | 21.23 | 73.68 | 22.23 | 73.62 | 21.25 | 73.75 | 22.25 | 73.75 |
| NRLP, Peninsular Components, Damanganga-Pinjal Link - 13 | Damanganga | Pinjal | 20.18 | 73.10 | 19.81 | 73.27 | 20.25 | 73.25 | 19.75 | 73.25 |
| NRLP, Peninsular Components, Bedti-Varda Link - 14 | Bedti | Varda | 14.16 | 74.86 | 14.39 | 75.37 | 13.75 | 75.25 | 14.25 | 75.25 |
| NRLP, Peninsular Components, Netravati-Hemavati Link - 15 | Netravati River | Hemavati River | 12.82 | 75.64 | 12.82 | 76.05 | 12.75 | 75.25 | 12.75 | 75.75 |

**Table S 2.8b:** Project names, year constructed, capacities, and variables $P$, $Q_{min}$, $Q_{max}$, used for calculating canal flows.

| Project Name | Year** | Annual Capacity ($10^9$ m3 yr-1) | Percent flow, $P$ | MinFlow ($m^3$/s), $Q_{min}$ | Max flow ($m^3$/s), $Q^{max}$ | Canal Length (m) | New Reservoir Capacity ($km^3$) |
|---|---|---|---|---|---|---|---|
| *Completed Inter-basin transfers* | | | | | | | |
| Kurnool Cuddapah Canal System | 1863 | 2.68 | 30 | 0 | 85 | 304 | na**** |
| Parambikulam Aliyar Project - West to East Combined | 1965 | 0.85 | 15 | 0 | 26.9192 | | na**** |
| Parambikulam Aliyar Project - Northwest Combined | 1965 | 0.08 | 15 | 0 | 2.69192 | | na**** |
| Teluga Ganga Project | 2004 | 2.14 | 30 | 0 | 67.68576 | | na**** |
| Teluga Ganga Project | 2004 | 2.14 | 30 | 0 | 67.68576 | | na**** |
| Teluga Ganga Project | 2004 | 0.43 | 30 | 0 | 13.46744 | | na**** |
| Indira Gandhi Canal Project | 1986 | 9.36 | 30 | 0 | 296.6005 | | na**** |
| Indira Gandhi Canal Project | 1986 | 10.6 | 70 | 0 | 335.89373 | | na**** |
| Indira Gandhi Canal Project | 1986 | 10.6 | 70 | 0 | 335.89373 | | na**** |
| Indira Gandhi Canal Project | 1986 | 5.3 | 70 | 0 | 167.94687 | | na**** |
| *Proposed NRLP Inter-basin transfers* | | | | | | | |
| NRLP, Himalayan Component, Kosi-Mechi Link - 1* | | | | | | | na**** |
| NRLP, Himalayan Component, Kosi-Ghagra Link - 2 | | | | | | | |
| NRLP, Himalayan Component, Gandak-Ganga Link - 3 | 2006 | 47.34 | 30 | 80 | 1500 | | 0.1015 |
| NRLP, Himalayan Component, Ghagra-Yamuna Link - 4 | 2006 | 47.34 | 30 | 50 | 1500 | | 0.2569 |
| NRLP, Himalayan Component, Sarda-Yamuna Link - 5 | 2006 | 6.31 | 30 | 100 | 200 | | 0.3992 |
| NRLP, Himalayan Component, Yamuna-Rajasthan Link - 6 | 2006 | | 30 | 0 | 2000000 | | 1.8448 |
| NRLP, Himalayan Component, Rajasthan-Sabaramati Link - 7 | 2006 | | 30 | 0 | 2000000 | | 1.9004 |

| Link | Year | | | | | | |
|---|---|---|---|---|---|---|---|
| NRLP, Himalayan Component, Chunar-Stone Barrage Link - 8* | 2006 | | 70 | 0 | 2000000 | | 1.5396 |
| NRLP, Himalayan Component, Sone Dam-Southern Tribs of the Ganga Link Link - 9* | 2006 | | | | | | na**** |
| NRLP, Himalayan Component, Manas-Ganga Link - 10a | 2006 | | | | | | na**** |
| NRLP, Himalayan Component, Manas-Ganga Link - 10b | 2006 | | 30 | 0 | 2000000 | | 1.8755 |
| NRLP, Himalayan Component, Jogighopa-Frakka Link - 11 | 2006 | | 30 | 0 | 2000000 | | 11.6 |
| NRLP, Himalayan Component, Farakka-Sunderbans Link - 12* | 2006 | | 30 | 0 | 2000000 | | 11.6 |
| NRLP, Himalayan Component, Ganga (Farakka)-Subernarekha Link - 13 | | | | | | | na**** |
| NRLP, Himalayan Component, Subernarekha-Mahanadi Link - 14 | 2006 | | 30 | 0 | 2000000 | | 11.6 |
| NRLP, Peninsular Components, Mahanadi-Godavari Link - 1 | 2006 | | 30 | 0 | 2000000 | | 11.6 |
| NRLP, Peninsular Components, Godavari-Krishna Link - 3 | 2006 | 12.17 | 30 | 0 | 385.48559 | 822 | 5.8804 |
| NRLP, Peninsular Components, Godavari-Krishna Link - 2 | 2006 | 1.66 | 3 | 0 | 52.72898 | 290 | 0.2252 |
| NRLP, Peninsular Components, Godavari-Krishna Link - 4 | 2006 | 14.76 | 27 | 0 | 467.77955 | 300 | 2.0217 |
| NRLP, Peninsular Components, Krishna-Pennar Link - 5 - SPLIT | 2006 | 2.27 | 30 | 0 | 71.77352 | 174 | 1.4631 |
| NRLP, Peninsular Components, Krishna-Pennar Link - 6* | 2006 | 1.98 | 30 | 0 | 62.74241 | 552 | 1.2263 |
| NRLP, Peninsular Components, Krishna-Pennar Link - 7*** | | | | | | 204 | na**** |
| NRLP, Peninsular Components, Pennar-Palar, Cauvery Link - 8 | | | | | | 392 | na**** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| NRLP, Peninsular Components, Cauvery-Vaigai, Gundar Link - 9a | 2006 | 26.12 | 30 | 0 | 827.75623 | 529 | 0.1259 |
| NRLP, Peninsular Components, Cauvery-Vaigai, Gundar Link - 9b*** | 2006 | 2.25 | 30 | 0 | 71.36157 | 256 | 1.1746 |
| NRLP, Peninsular Components, Ken-Betwa Link - 10 | | | | | | | |
| NRLP, Peninsular Components, Parbati-Chambal Link - 11a | 2006 | 1.02 | 30 | 0 | 32.32185 | 231 | 0.2892 |
| NRLP, Peninsular Components, Parbati-Chambal Link - 11b | 2006 | 0.46 | 30 | 0 | 14.70327 | 244 | 0.116 |
| NRLP, Peninsular Components, Par-Narmada Link - 12a | 2006 | 0.49 | 30 | 0 | 15.62223 | | 0.1943 |
| NRLP, Peninsular Components, Par-Narmada Link - 12b | 2006 | 1.35 | 15 | 0 | 42.77892 | 395 | 0.8197 |
| NRLP, Peninsular Components, Damanganga-Pinjal Link - 13 | 2006 | 1.35 | 30 | 0 | 42.77892 | | 0.8814 |
| NRLP, Peninsular Components, Bedti-Varda Link - 14 | 2006 | 0.909 | 15 | 0 | 28.80447 | | 0.5549 |
| NRLP, Peninsular Components, Netravati-Hemavati Link - 15 | 2006 | 0.242 | 30 | 0 | 7.66852 | | 0.1606 |
| NRLP, Peninsular Components, Pamba-Vaippar Link - 16 | 2006 | 0.188 | 30 | 0 | 5.95736 | | 0.1197 |
| NRLP, Peninsular Components, Krishna-Pennar Link - 5 - SPLIT | 2006 | 0.634 | 30 | 0 | 20.09025 | 51 | 0.0628 |

**Year 2006 is used as the start year for all NRLP transfers. This does not represent a real historical start date, but was used to initialize the mid-century results

****WBMplus simulations included storage reservoirs only for proposed NLRP inter-basin transfers

**Table S 2.9:** Comparison of WBMplus national-level simulated irrigation results with reported water use statistics from FAO AQUASTAT's report: *Irrigation in Southern and Eastern Asia in Figures – AQUASTAT Survey – 2011. All values are km³ yr⁻¹.*

|  | AQUASTAT | WBMplus[a] |
|---|---|---|
| Year 2010 national total irrigation water withdrawals | 688 | 716 |
| Year 2010 internal renewable surface water resources | 1,404 | 1,283 |
| Year 2010 surface water withdrawals | 396 | 395 / 415[b] |
| Year 2010 primary groundwater withdrawals | 251 | 320 |
| Year 1990 national total irrigation water withdrawals | 460 | 375 |
| Year 1990 primary groundwater withdrawals | 190 | 153 |
| Year 1996 average annual runoff of major river basins: |  |  |
| Ganges | 525 | 411 |
| Godavari | 111 | 110 |
| Krishna | 78 | 79 |
| Indus | 73 | 78 |
| Mahanadi | 67 | 43 |
| Narmada | 46 | 44 |
| Cauvery | 21 | 22 |
| Pennar | 6.3 | 3.5 |

[a] WBMplus values for years post-2005 are the average results from simulations using 5 bias-corrected GCM climate drivers, and econometrically projected crop maps. WBMplus values for years pre-2005 are from simulations using APHRODITE Monsoon Asia climate drivers, and ICRISAT crop maps.

[b] WBMplus separates surface water withdrawals from groundwater recharge withdrawals. However, AQUASTAT statistics say that there is an overlap between their reported surface water withdrawals and renewable groundwater withdrawals. WBMplus reports surface water withdrawals of 395 km³ yr⁻¹ when only rivers and reservoirs are considered. Adding 10% of WBMplus's renewable groundwater withdrawals to this result brings the estimate up to 415 km³ yr⁻¹.

**Table S 2.10: Description of Punjab well-level data, collected from 1973 – 2003.**

| District | No. of wells | No. of wells that go dry within sampling period |
|---|---|---|
| Amritsar | 231 | 86 |
| Bathinda | 84 | 1 |
| Gurdaspur | 174 | 24 |
| Firozpur | 136 | 3 |
| Hoshiarpur | 116 | 12 |
| Jalandhar | 177 | 62 |
| Patiala | 160 | 2 |
| Sangrur | 140 | 5 |

**Table S 2.11: Model validation: comparison of historical district-level groundwater level data and hydrologic model simulation of changes in groundwater levels.**

| District | Model-based category | Data-based category | Agree | Comment |
|---|---|---|---|---|
| Amritsar | 2: same rate | 3: increase rate | NO | Model is conservative |
| Bathinda | 5: recover/ static | 5: recover/ static | **YES** | |
| Gurdaspur | 2: same rate | 3: increase rate | NO | Model is conservative |
| Firozpur | 5: recover/ static | 5: recover/ static | **YES** | |
| Hoshiarpur | 3: increase rate | 3: increase rate | **YES** | |
| Jalandhar | 2: same rate | 2: same rate | **YES** | |
| Patiala | 2: same rate | 2: same rate | **YES** | |
| Sangrur | 2: same rate | 2: same rate | **YES** | |
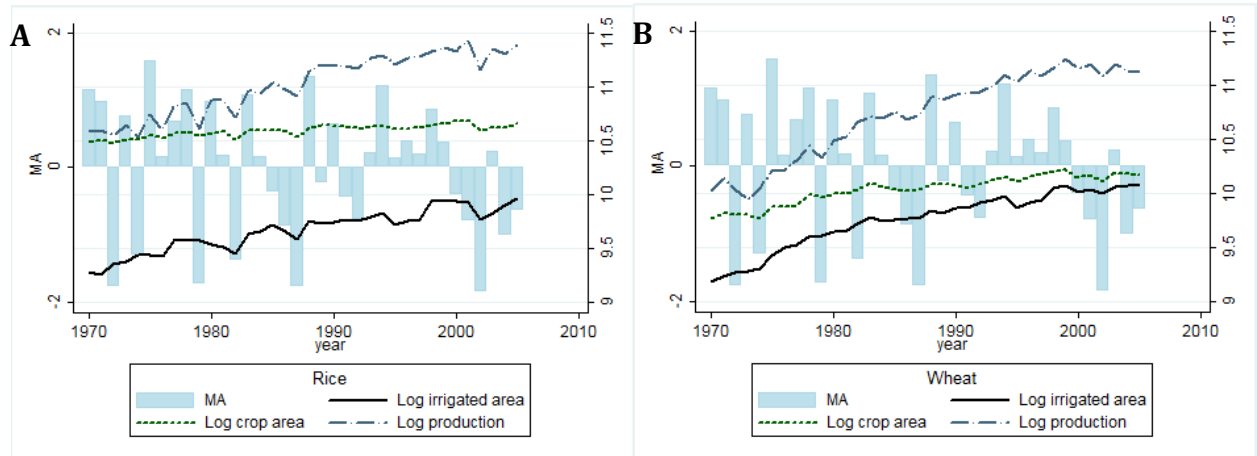
**Fig. S 2.1**. Each panel reports values aggregated in each period over the district sample used in the analysis of (**A**) rice and(**B**) wheat, with values of log irrigated area, log crop area and log production reported on the right axis. Blue bars are normalized monsoon anomalies (MA), whose values are reported on the left axis. Crop data are from ICRISAT, and monsoon anomalies are from the APHRO_MA_V1101R2 precipitation product.
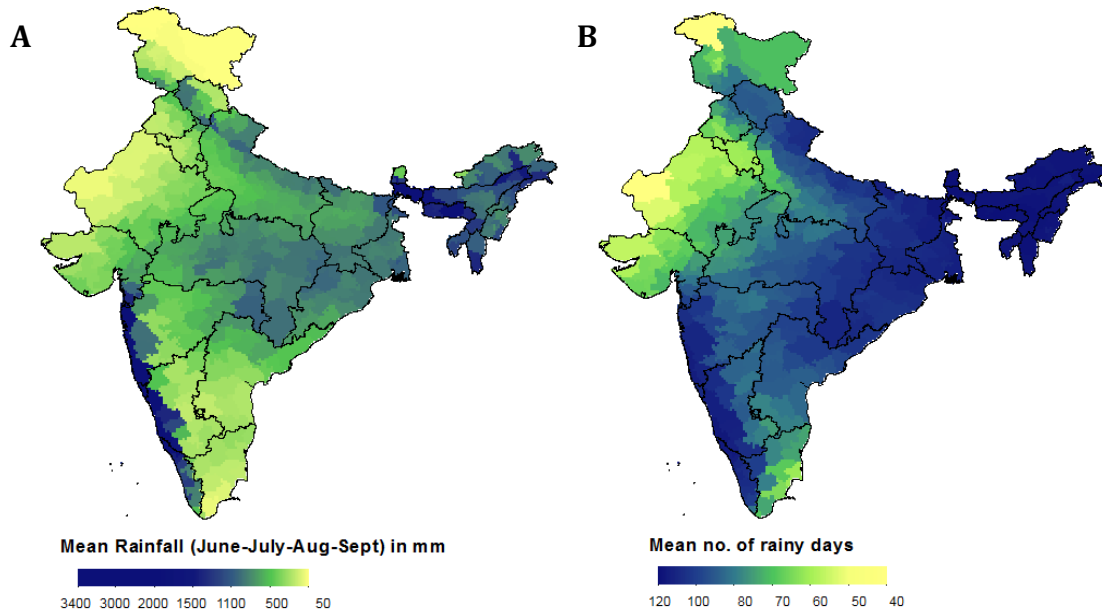


**Fig. S 2.2.** Average (1970-2005) (**A**) June to September monsoon rainfall (mm), and (**B**) number of monsoon rainy days (days with precipitation>0.1mm from June to September). Gridded APHRO_MA_V1101R2 data were aggregated to district values; state borders are in black.
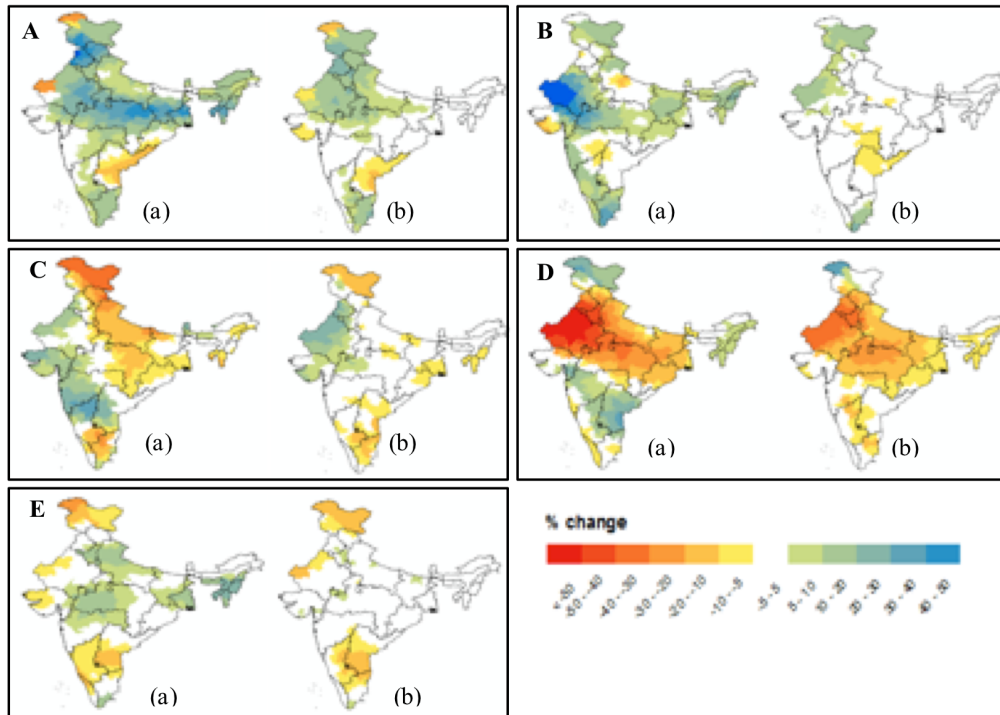
**Fig. S 2.3**. **(a)** Percent change in decadal average monsoon precipitation and **(b)** Percent change in decadal average number of rainy days in the monsoon season from the 1970s to the 2040s from 5 CMIP5 GCMs: **A)** MIROC-ESM-CHEM, **B)** CCSM4, **C)** GFDL-CM3, **D)** GFDL-ESM2G, and **E)** NorESM1-M. Differences are taken between the bias-corrected model historical runs and bias-corrected model future runs for RCP 8.5.
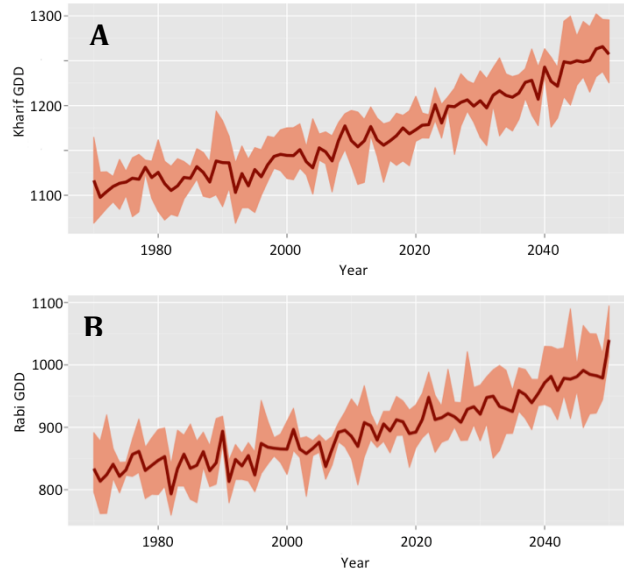
**Fig. S 2.4**. Seasonal growing degree days in the (**A**) wet (Kharif) and (**B**) dry (Rabi) seasons. The solid lines represent the multi-model mean of five different GCM climate futures, and the shade bands the five-model range.
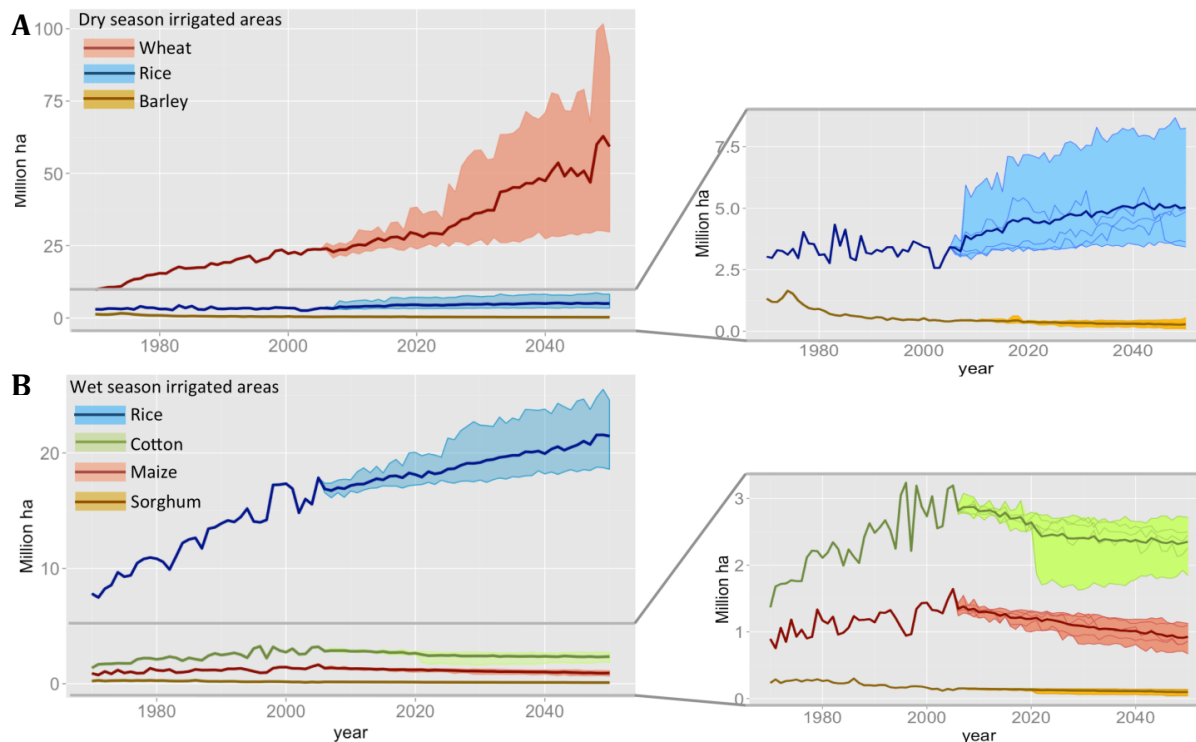


**Fig. S 2.5** Econometric model generated irrigated area projections for (**A**) dry season and (**B**) wet season crops in million hectares. The historical period (1970-2005) reflects data from ICRISAT. For the future period (2006-2050), the solid line reflects the multi-model mean of econometric projections based on five different GCM climate futures, with the range of uncertainty due to differences in GCM projections in the shaded region. Note that the y-axis scales are different
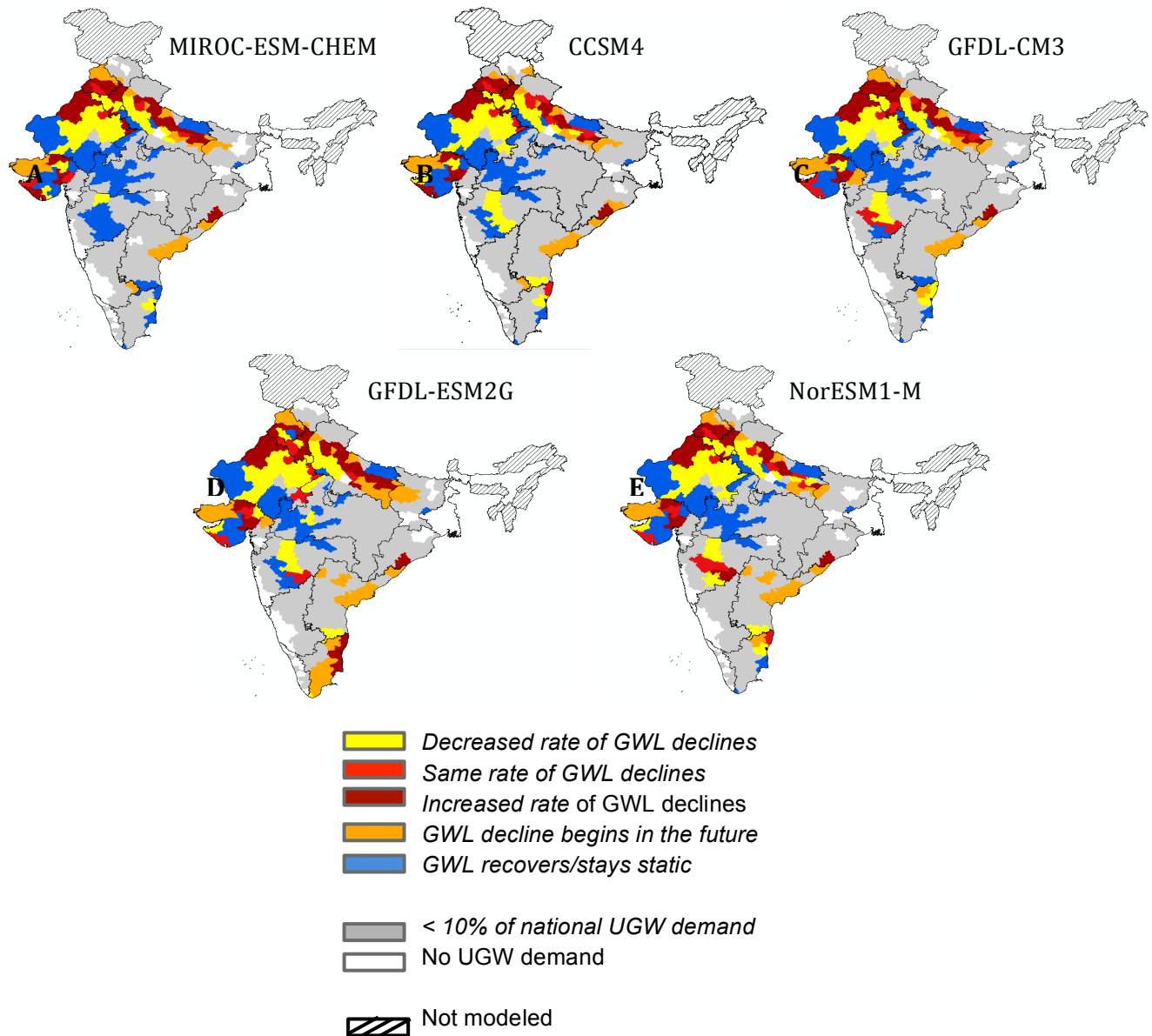
**Fig. S 2.6**. Trends in groundwater levels (GWL) between 1979-2000 and 2029-2050, inferred from need for unsustainable groundwater (UGW) to meet irrigation water demand, from 5 GCMs: **A**) MIROC-ESM-CHEM, **B**) CCSM4, **C**) GFDL-CM3, **D**) GFDL-ESM2G, and **E**) NorESM1-M. Decreases in UGW demand will slow down GWL declines (yellow); continued demand will lead to continued GWL declines (red); increase in UGW demand will increase groundwater level (GWL) declines (dark red); relying on unsustainable sources for the first time can start GWL declines (orange); future reliance on sustainable sources can allow GWL to recover (blue). Map has district-level summaries, with state boundaries drawn in black. Colored (non-grey) regions account for 90% of modeled mean national UGW demand (2029-2050).
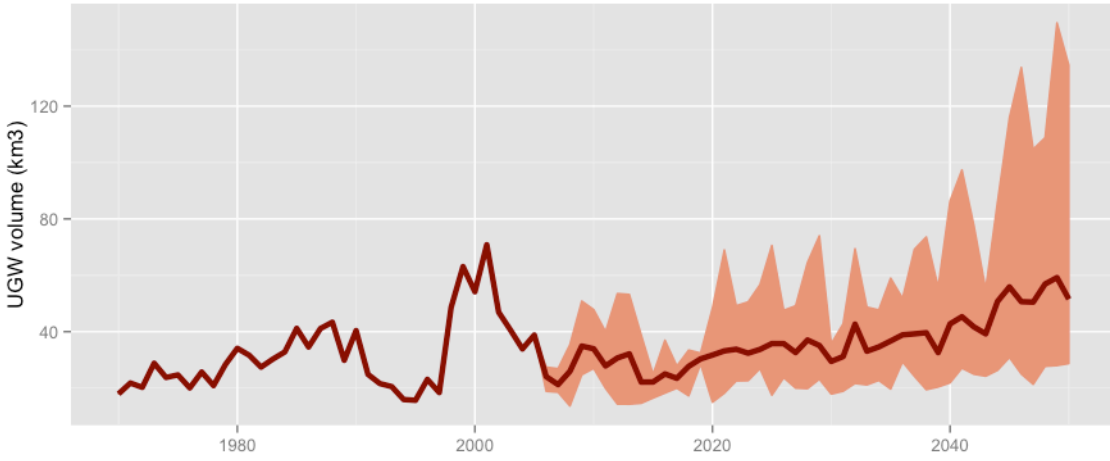
**Fig. S 2.7**. Historical and projected UGW demand. The shaded ribbon shows the range in future projections of UGW based on 5 different climate model simulations.

CHAPTER III:


THE USE AND REUSE OF UNSUSTAINABLE GROUNDWATER: A GLOBAL

BUDGET

**Abstract**

Groundwater supplies nearly half of all water used in irrigated agriculture (Aeschbach-

Hertig and Gleeson, 2012), making it critical to global food production.  Many of the

world's major groundwater aquifers are rapidly depleting due to unsustainable groundwater

pumping, much of it for agriculture (Aeschbach-Hertig and Gleeson, 2012; Rodell et al,

2009; Gleeson et al, 2012; Wada et al 2012), while demand for food production – and

therefore demand for irrigation water – is increasing (FAO, 2015).  Unsustainable

groundwater is defined broadly as groundwater extraction in excess of recharge (Aeschbach-

Hertig and Gleeson, 2012; Rodell et al, 2009; Gleeson et al, 2012; Wada et al 2012).  While

it is clear that groundwater users will be impacted by reductions in groundwater availability,

there is a major gap in our understanding of potential impacts downstream of groundwater

pumping locations.  Here, we quantify the amount of unsustainable groundwater extracted

and reused downstream of pumping sites, estimate the amount of unsustainable groundwater

needed to sustain agriculture even under high irrigation efficiency scenarios, and present the

first global budget of unsustainable groundwater reuse through agricultural systems.

Previous global studies of agricultural water reuse assessed only surface water supplies (see

review by Simmons et al., 2015): we find that groundwater reuse is responsible for a quarter to a third of total irrigation water supplies in many Asian river basins, and ~10% total irrigation water suppliesglobally. Some studies have called for increasing irrigation efficiency as a solution to water shortages (Gleick, 2001; Wada, 2014). We find that, because inefficiencies allow irrigation water to be reused, increasing irrigation efficiency in many major agricultural river basins reduces but does not eliminate the demand for unsustainable groundwater pumping; even with 100% irrigation efficiency, global demand for unsustainable groundwater is 186 $km^3yr^{-1}$. As inefficiencies also allow groundwater to enter surface water systems, in many basins an increase in irrigation efficiency leads to decreased downstream river flows.

## Introduction

Classical irrigation efficiency is defined as the ratio beneficial crop water use (evapotranspiration and water required to flood rice fields) to gross irrigation water extracted from water sources. This measure of irrigation efficiency is a misleading indicator of water use efficiency in the case of groundwater (Foster and Perry, 2010; Hafeez et al., 2007), because a fraction of the unused portion of extracted groundwater can return to both surface water and groundwater pools. Classical irrigation efficiency is estimated at ~37% - 50%,(FAO, 2011; Postel, 1993), allowing significant amounts of extracted groundwater to become runoff and recharge. Groundwater provides nearly half of all global irrigation water (Aeschbach-Hertig and Gleeson, 2012), and recent global-scale assessments have found that unsustainable groundwater provides ~20% of total irrigation water supplies (Gleeson et al., 2012; Wada et al., 2012). Here, we define unsustainable groundwater (UGW) as the

average annual groundwater extracted in excess of average annual recharge. While this definition does not account for potentially complex surface water-groundwater interactions (Aeschbach-Hertig and Gleeson, 2012), it serves as a large-scale indicator of groundwater depletion (Gleeson et al., 2012; Wada et al., 2010, 2012; Wisser et al., 2010). Previous studies of irrigation efficiencies and return flows have not included UGW in their analyses, though it has been shown that groundwater extractions can alter surface water storage volumes and river flows (Döll et al., 2012; de Graaf et al., 2014). Both satellite-based and model-based estimates of global groundwater depletion show that aquifers in important agricultural regions are losing mass (Aeschbach-Hertig and Gleeson, 2012; Rodell et al., 2009; Gleeson et al., 2012; Wada et al., 2010, 2012), and cannot continue providing current levels of groundwater supplies indefinitely. To more fully understand the implications of decreasing groundwater supplies, it is important to take UGW reuse into account in order to assess both the reliance of agriculture on UGW, and the potential for increased irrigation efficiency to reduce this reliance.

While classical irrigation efficiency is sufficient for assessing on-field water use efficiency, it is now recognized as insufficient at larger scales (Simmons et al., 2015; Vörösmarty et al., 2005; Seckler et al., 2003), leading several studies to highlight the need to understand return flows from irrigated areas (Simmons et al, 2015; Vörösmarty et al., 2005; Seckler et al., 2003; Tornqvist and Jarsio, 2011), and develop of over a dozen different metrics to quantify irrigation water reuse at field- to basin-scales (Simmons et al, 2015). Alternatives include the basin closure concept, which assesses water use efficiency at a whole-basin scale by comparing basin inflows to outflows (Molden and Boss, 2005), the water reuse index, which quantifies surface water reuse along an entire river transect

(Vörösmarty et al., 2005), and the net efficiency concept, which aims to assess what proportion of agricultural runoff is suitable for reuse (Seckler et al., 2003). These metrics cannot separate UGW– an unsustainable water source – from analyses of sustainable surface water supplies.

Here, we introduce both the unsustainable groundwater reuse index **R**, as well as the first estimate of minimum unsustainable groundwater dependence. The unsustainable groundwater reuse index quantifies how many times extracted UGW is reused within a river basin due to irrigation inefficiencies. The minimum unsustainable groundwater dependence is the quantity of unsustainable groundwater extraction required to meet irrigation water requirements under a 100% irrigation efficiency scenario. This metric quantifies the lower bound of the current agriculture system's reliance on unsustainable water sources.

We calculate gross irrigation water requirements, UGW extraction, and the amount of UGW that enters river systems and groundwater recharge using the global gridded Water Balance Model (Wisser et al., 2010) (WBM, see Materials and Methods). UGW that enters streams by way of runoff and baseflow is tracked downstream through WBM's simulated river network, and can be extracted from the (well-mixed) rivers and large reservoirs, as well as groundwater recharge pools, to meet simulated irrigation water requirements. Use and re-use of UGW is tracked through model storages and flows, including soil moisture, reservoir storage, groundwater storage, baseflow, and river discharge. All values reported here are 30-year mean annual values and one standard deviation, based on contemporary distribution of crops and weather variability from 4 different climate input datasets (see Materials and Methods).

## Results

*A global groundwater budget*

We find that global UGW pumped for irrigation is 378 ($\pm$49) $km^3yr^{-1}$, or ~12% of gross irrigation (3,244 ($\pm$240) $km^3yr^{-1}$). By tracking UGW reuse, we find that 238 ($\pm$35) $km^3yr^{-1}$ contributes directly to crop evapotranspiration (ET) (Fig. 1). This crop ET volume is due to 145 ($\pm$18) $km^3yr^{-1}$ of direct use of the efficient portion of irrigation, and an additional 91 ($\pm$20) $km^3yr^{-1}$ of crop ET from UGW_r. In total, 63% ($\pm$9%) of the initial pumped volume of UGW becomes crop ET. Only 97 ($\pm$8) $km^3yr^{-1}$, or 26% ($\pm$2%) of the initial pumped volume, leaves river systems by discharge to the ocean and internal basins (Fig. 1). UGW discharge directly to the ocean is 90 ($\pm$7) $km^3yr^{-1}$, or 0.25 ($\pm$0.02) mm of Sea Level Equivalents, which is within the range of previous estimates (0.075 – 0.8 mm) (Konikow, 2011). An additional 43 ($\pm$13) $km^3yr^{-1}$ is lost to non-beneficial evaporation from fields and conveyance structures.

*The groundwater reuse index*

Basins that have high levels of UGW reuse are more dependent upon this unsustainable source of water than estimates of groundwater depletion imply. We define a UGW reuse factor, ***R***, which is calculated as [***R = ((UGW + UGW_r) / UGW) - 1***], as a measure of the number of times a unit of water extracted from UGW is reused in a river basin. In basins with R > 1.0, all extracted UGW is reused more than once, effectively acting as twice as much water (or more) within the basin irrigation system. In basins with a reuse factor of 0.5 half of extracted UGw is reused.

The highest values of both UGW (Fig. 3.2a) and *R* (Fig. .32b) occur in central, south, and east Asia. By overlapping maps of irrigated areas (MIRCA 2000) and population density data (UN Dept. of Economic and Social Affairs, 2013) with the basin reuse factor map, we find that basins with *R* > 0.5 contain 89 million hectares of irrigated land (41% of global irrigated area), and are home to 1.7 billion people. Basins with *R* > 1.0 contain 33 million hectares of irrigated land and are home to 1.3 billion people.

*Unsustainable groundwater contributes to surface water flows*

We find that UGW returning from irrigated areas contributes 10% – 50% to basin-level average annual discharge (Fig. 3a). Given this dependence on an unsustainable water source for river flows, how will increasing irrigation efficiency alter river flows, and how will reductions in access to UGW alter river flows? To answer these questions, we developed three model simulations, referred to here as Business As Usual (BAU), Global Irrigation Efficiency 70% (GIE70), and No UGW (NoUGW). The BAU simulation uses WBM to track UGW under current (c. year 2000) conditions, as described above. In GIE70, irrigation efficiency is increased to 70% in all model grid cells with irrigation. In noUGW, irrigation efficiency remained at current FAO-based levels, but no UGW was extracted, and total irrigation demand was not met. Average annual river discharge decreased in both experiments for river segments across Asia, the Middle East, and the US southwest (Fig. 3). GIE70 resulted in 100% decrease (i.e., complete drying of the river segment) in multiple headwater grid cells across all three of these regions (Fig. S 3.1). UGW can contribute a significant amount of irrigation water through river system reuse. For example, the Ganges and Indus river systems supply 285 and 262 $km^3yr^{-1}$ of irrigation water, respectively.

However, 17% and 15% of these surface water supplies are only present due to the inefficiencies of upstream UGW irrigation, and half of the river and reservoir water used in both the Sabarmati and Brazos basins is attributable to UGW runoff from upstream irrigation (Table 3.1).

*Effects of increasing irrigation efficiency*

Much of the debate over water reuse metrics, measurements, and concepts stem from proposals to increase irrigation efficiency (Simmons et al, 2015; Gleick, 2001; Wada 2012). Several studies have shown that irrigation efficiency improvements can lead to unintended increased water use due to the perceived increase in water availability (Contor and Taylor, 2013; Ward and Pulido-Velazquez, 2008), but none have assessed the downstream impact of reducing upstream groundwater pumping.

Increased irrigation efficiency will impact irrigation requirements, extraction of UGW, and river flow; these changes will depend upon the balance between irrigation water requirements, and the reuse index R. With a higher irrigation efficiency, total irrigation water requirements - and therefore UGW - may be decreased, leading to decreased return flow and consequently lower river discharge. Lower total irrigation water demands could also reduce irrigation water extraction from rivers, leading to increased river discharge.  To explore the effect of increaseing irrigation efficiencies, we simulated a series of incremental increases in the national minimum irrigation efficiency: 34% (current minimum national irrigation efficiency: FAO, 2011), 40%, 50%, …, 100%.  With each increase, the irrigation efficiencies of all countries that are below the new minimum threshold are raised to that threshold.  These simulations represent hypothetical changes; we do not propose pathways

or timelines for implementing these changes, but rather use these simulations to quantify the potential lower bound for global UGW demand. At 100% efficiency, UGW_r goes to zero, and gross irrigation water demand is equal to net irrigation water demand. The 100% irrigation efficiency scenario identifies the minimum global UGW dependence, i.e., the minimum potential volume of UGW that is required to meet current agricultural water requriements.

We find that the mimimum global UGW dependence is 180 (±28) $km^3yr^{-1}$ (Fig. 3.4), a reduction of ~52% from current UGW demand. To reduce UGW demand further will require additional or alternative changes irrigated agriculture (e.g., switching to less water consuming crops, or varieties with increased water use efficiency; Wada 2012). For each increased irrigation efficiency scenario, we also quantify the total amount of UGW used for agriculture (UGW + UGW_r), which is equivalent to the water deficit that would occur under each irrigation efficiency scenario if UGW resources were unavailable (Fig. 3.4). At 100% irrigation efficiency, UGW_r is zero. However, at current efficiency levels, total UGW (UGW + UGW_r) is 748 (±107) $km^3yr^{-1}$, approximately one quarter of global gross irrigation water requirements. This shows that the loss of UGW resources – either due to complete depletion of aquifers, or economical reasons – would lead to a 25% shortage of irrigation water supplies globally. Under a 100% efficiency scenario, UGW makes up about 15% of global gross irrigation water requirements.


### Discussion

Use of unsustainable groundwater for irrigation has been the focus of several recent global-scale analyses (Wada et al., 2012; Aeschbach-Hertig and Gleeson, 2012; Gleeson et

al., 2012), all showing aquifer depletion and the spatial patterns of regional reliance on groundwater for irrigated agriculture. However, there are significant differences in how this water is used and reused once it has been extracted. Here, we find that the unsustainable groundwater reuse index, **R**, is a tool that can be used in conjunction with estimates of unsustainable groundwater extraction to reveal the reliance of irrigated agriculture not only on unsustainable groundwater, but also on reuse of unsustainable groundwater. Basins with high reuse factors have an outsized reliance on this unsustainable resource, which estimates of groundwater extraction alone cannot accurately assess. Additionally, basins with high reuse factors have a diminished ability to reduce their reliance on unsustainable groundwater by improving irrigation efficiency, as higher efficiencies lead to reductions in reuse. By finding the potential minimum amount of unsustainable groundwater required by irrigated agriculture, we are able to quantify the volume by which other (non irrigation efficiency) water saving measures must reduce the reliance on UGW in order to achieve sustainable groundwater use.

## Materials and Methods

*Irrigation*

Crop maps of both irrigated and rainfed land, along with crop type and season length are from the MIRCA2000 database (Portmann, Siebert, and Doell, 2010). National statistics on the ratio of surface water (from rivers and reservoirs) to groundwater (groundwater recharge and UGW) supplies used for irrigation at a country level (FAOSTAT, 2015) are used to determine source of irrigation water withdrawals in WBM. WBM's grid-cell level irrigation water extractions occur in three stages: 1) surface water and renewable

groundwater (i.e., groundwater ) are extracted in the FAOSTAT (FAOSTAT, 2015) based

ratio of surface- to groundwater if possible; 2) if (1) is not possible, irrigation water demand

is fulfilled using any remaining surface water or groundwater; and 3) if irrigation water

demand is still not met after (2), then UGW is extracted to fulfill the remaining demand. All

other methods for irrigation water demand and application are based on Allen (1998), as

described in Wisser et al (2010) and Grogan et al (2015). See Supplemental Information for

validation of UGW extractions and total groundwater extractions.

The inefficient portion of all irrigation water extractions (the difference between

gross and net irrigation water volumes) is split into three portions. First, it can evaporate to

meet local evaporative demand:

$$E = (PET_i - AET_i) * IrrAreaFrac_i \tag{3.1}$$

where $PET_i$ is the potential evapotranspiration volume grid cell $i$, $AET_i$ is the actual

evapotranspiration for grid cell $i$ (calculated after applying irrigation water to soils within

the irrigated area), and $IrrAreaFrac_i$ is the fraction of grid cell $i$ that is irrigated. After

evaporation occurs, the remaining inefficient portion of irrigation water extractions is

divided equally between surface runoff and groundwater recharge. See Supplemental

Materials for a sensitivity analysis of this parameterization of return flows.


***Tracking UGW***

WBM tracks UGW, as well as non-UGW sources (precipitation and snow melt) of

water through all model stocks and flows (Fig. S 3.2). At each daily time step, the

proportion of each stock is updated based on inflows and outflows of water, and that water's

proportional composition of water sources. We assume all stocks are well-mixed; if a stock

is X% UGW, then outflows from that stock are also X% UGW.  In this way, a unit of water retains its identity as UGW even as it passes through surface water flows.  We define UGW extracted from surface water flows and groundwater recharge pools as UGW_r.

### Supplemental Information

*Uncertainty and sensitivity analysis*

Five sources of uncertainty were assessed: 1) inter-annual weather variability, 2) climate input data sets, 3) rice paddy percolation rates, 4) irrigation return flow distribution between surface runoff and groundwater recharge, and 5) WBM's irrigation water search distance parameter.  Uncertainties due to the first four sources are reflected in all reported values in this study.

To address (1) and (2), we simulated 30 years (1980-2009) of weather based on 4 different climate input datasets (Dee et al, 2011 *ERA interim*; Rienecker et al, 2011 *MERRA*; Saha et al, 2011 *NCEP;* UDEL (2.0) Willmott and Matsuura, 2001).  These 30 years do not represent a time series of irrigation water demand because irrigation land areas were not altered.  Rather, the 30-year time period allows quantification of inter-annual variability in all hydrological variables of interest, circa year 2000.  Inter-annual variability in UGW is 36 – 51 $km^3yr^{-1}$, based on the climate data set used.

Rice paddy percolation rates are based on soil drainage class (FAO/UNESCO, 2003), and range from 2 mm $day^{-1}$ to 8 mm $day^{-1}$, as described in Wisser et al (2010). Irrigation water applied to rice paddies to maintain flooding is considered part of net irrigation, and therefore paddy percolation is not counted towards inefficient water "losses". However, since this water percolates through soils at the rates described above and enters

groundwater recharge as opposed to crop evapotranspiration, it has the potential to significantly effect UGW flows to groundwater recharge and baseflow.  To assess the impact of rice percolation rates on irrigation water percolation, UGW demand, and UGW_r, we simulated alternative rice paddy percolation rates, reducing rates by 10%, 20%, 30%, 40% and 50%.  We find that global total paddy percolation of irrigation water ranges from 322 $km^3yr^{-1}$ to 535 $km^3yr^{-1}$.  This volume is significant, as 535 $km^3yr^{-1}$ is 18% of gross irrigation water demand.  However, reducing percolation rates by 50% globally only resulted in a 5% decrease in global UGW demand.

Irrigation return flow distribution between surface runoff and groundwater recharge is assessed by altering the distribution within WBM.  Actual distribution of irrigation return flows will be dependent upon the process by which irrigation water is lost (e.g., canal leakage versus flood irrigation), soil drainage properties, previous levels of soil saturation, and irrigation methods.  These all vary in space and time.  However, we find that the distribution has little impact (1% – 8%) in total UGW_r, although it alters the relative proportion of UGW_r through surface water versus groundwater.

Uncertainty (5) WBM's irrigation water search distance parameter, is a significant source of uncertainty to estimates of global UGW.  Within WBM, irrigation water demand and available water resources (from surface water, reservoirs, and groundwater recharge) are assessed in each 30° grid cell on a daily basis.  Grid cells which contain large rivers therefore have access to significantly more sustainable surface water than other grid cells.  This makes the amount of sustainable surface water available – and therefore the demand for UGW – dependent upon the spatial resolution of the model.  Additionally, in reality many agricultural regions contain extensive canal networks that transport water away from

river mainstems.  To address both the spatial resolution issue and the reality of small canal systems, a "search distance" is set within the model; this search distance identifies the radius of a circle around the centroid of each grid cell.  The grid cell at the center is able access sustainable surface water (from rivers, reservoirs, and groundwater recharge) from all other grid cells that fall within the circle.  How large the radius of such a circle should be to best represent the spatial characteristics of a given regions is likely a function of topography and infrastructure development.  Future work in modeling macro-scale irrigation systems should identify spatial variations in appropriate values for the search distance.

In the results shown here, the search distance is set globally to 75 km, resulting in an annual average UGW demand of 378 ($\pm$49) $km^3yr^{-1}$ (averaged across 30 years of climate input data for each of 4 climate data sets).  Eliminating the search distance function (effectively setting it to 0 km) results in an higher UGW demand of 663 ($\pm$55) $km^3yr^{-1}$; using a larger search distance of 150 km results in a lower UGW demand, 232 ($\pm$23) $km^3yr^{-1}$.  This large range of values brackets previous estimates of UGW from Wada et al. (2012), who report 256 ($\pm$58) $km^3yr^{-1}$ of groundwater depletion, and the 527 $km^3yr^{-1}$ UGW calculated from the results provided in Gleeson et al. (2012).  The spatial distribution of UGW demand sensitivity to the search distance parameter matches the spatial distribution of UGW demand: regions with high UGW demand are highly sensitive to the search distance parameter (Fig. S 3.3).  While this large sensitivity makes the UGW estimates of WBM and similar models more uncertain, it also suggests that within-basin canals infrastructure can play a significant role in reducing UGW demand.  This subject would benefit from further analysis.

*Validation*

We compare our country-level results to FAO reports of year-2000 (AQUASTAT) irrigation water withdrawal, fresh surface water withdrawal, and fresh groundwater withdrawal (Fig. S 3.4). There are 61 countries for which FAO repots at least one of these values for the year 2000 (Table S 3.1); we used all of these countries for comparison. For countries with no irrigation water withdrawal data available, we used FAO reported total agricultural water withdrawal data. FAO does not report unsustainable groundwater use separately from sustainable groundwater use; therefore, we compare the sum of WBM-simulated UGW and WBM-simulated irrigation withdrawals from groundwater recharge to FAO fresh groundwater withdrawal data. WBM-simulated irrigation withdrawals from rivers and reservoirs are compared to FAO's fresh surface water withdrawals for irrigation. Considering countries for which there are comparable values for each category, the $R^2$ value of WBM-simulations compared to FAO data is 0.92 for irrigation water withdrawals, 0.90 for groundwater withdrawals, and 0.75 for surface water withdrawals.

We compare WBM global and basin-level UGW results to the groundwater recharge and extraction data provided in the supplemental data from Gleeson et al. (2012). This supplemental information is provided as 30-minute gridded data; by subtracting the groundwater extraction grid from the recharge grid, we estimate UGW in each grid cell. Aggregation to the basin level is appropriate to account for potential spatial differences between the model parameters and input data used in Gleeson et al. (2012) and this study. Globally, we estimate Gleeson et al (2012) UGW volume is 527 $km^3yr^{-1}$. Thi sis 149 $km^3yr^{-1}$ higher than this study's average global UGW volume (378 ±49 $km^3yr^{-1}$), but only 12 $km^3yr^{-1}$ higher than the maximum global UGW – resulting from 30 years of climate data

from each of 4 different climate drivers – of 515 $km^3yr^{-1}$. Three basins account for more than 50% of the difference between the global mean WBM and Gleeson et al. (2012) results: compared to Gleeson et al (2012), WBM underestimates UGW in the Indus basin by 44 $km^3yr^{-1}$, in the Mississippi basin by 25 $km^3yr^{-1}$, and in the Ganges basin by 15 $km^3yr^{-1}$ (Fig S 3.5). While UGW in some basins are overestimated by comparison, none are overestimated by more than 2 $km^3yr^{-1}$.

**Table 3.1.** The 20 basins which supply the largest volume of reused UGW through river and reservoir water to irrigated crops are listed. **RR Irr. Water** is the volume of river and reservoir water used for irrigation. **UGW in RR Irr. Water** is the volume (percent) of UGW in the RR Irr. Water due to inefficiencies and reuse. **RR % of Gross Irr. Water Demand** is the % of gross irrigation water demand supplied by RR Irr. Water. **UGW** is the unsustainable groundwater pumping required to meet gross irrigation water requirements. **Basin UGW Discharge Fraction** is the basin-wide fraction of average annual discharge that is due to UGW.

| River Basin | RR Irr. Water [km$^3$yr$^{-1}$] | UGW in RR Irr. Water [km$^3$yr$^{-1}$] (%) | RR % of Gross Irr. Water Demand [%] | UGW [km$^3$yr$^{-1}$] | Basin UGW Discharge Fraction [-] |
|---|---|---|---|---|---|
| Ganges | 285 | 47 (17%) | 62% | 59 | 0.07 |
| Indus | 262 | 39 (15%) | 67% | 43 | 0.14 |
| Hai Ho | 26 | 5 (19%) | 55% | 9 | 0.13 |
| Huai | 44 | 4 (9%) | 68% | 5 | 0.09 |
| Godavari | 46 | 3 (7%) | 72% | 4 | 0.02 |
| Huang He | 40 | 2 (5%) | 73% | 4 | 0.02 |
| Amu-Darya | 20 | 2 (10%) | 67% | 6 | 0.02 |
| Narmada | 11 | 2 (19%) | 55% | 4 | 0.03 |
| Sabarmati | 4 | 2 (50%) | 36% | 4 | 0.21 |
| Mahi | 7 | 2 (29%) | 47% | 4 | 0.11 |
| San Joaquin | 4 | 2 (50%) | 71% | 2 | 0.02 |
| Krishna | 51 | 2 (4%) | 65% | 6 | 0.01 |
| Shatt el Arab | 21 | 2 (10%) | 60% | 2 | 0.01 |
| Tarim | 13 | 2 (15%) | 57% | 3 | 0.07 |
| Liao | 9 | 1 (11%) | 64% | 5 | 0.01 |
| Mississippi | 27 | 1 (4%) | 74% | 1 | 0.01 |
| Cauweri | 23 | 1 (4%) | 29% | 3 | 0.12 |
| Brazos | 2 | 1 (50%) | 51% | 4 | 0.02 |
| Colorado (Arizona) | 5 | 1 (20%) | 71% | 2 | 0.02 |
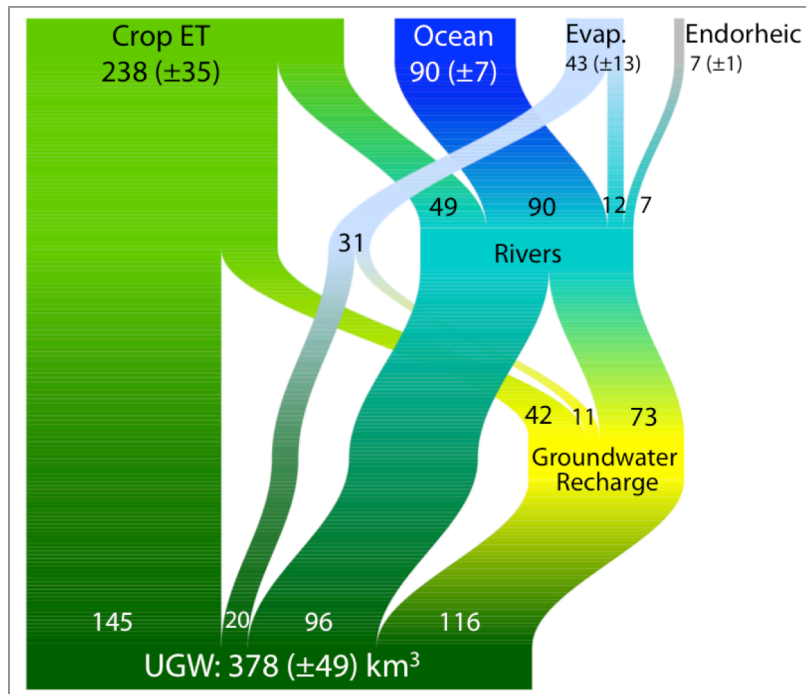
**Fig. 3.1: A global budget of unsustainable groundwater (UGW).** UGW used for irrigation is tracked, starting from initial pumping volumes (at bottom) and through efficient irrigation use to crop ET, inefficient losses to rivers and groundwater recharge, and reuse from river and groundwater recharge. The total amount of UGW contribution to crop evapotranspiration is 238 ($\pm$35) km$^3$yr$^{-1}$. 145 ($\pm$18) km$^3$yr$^{-1}$ is from direct use of UGW; the remainder is from UGW reuse through rivers and groundwater recharge.
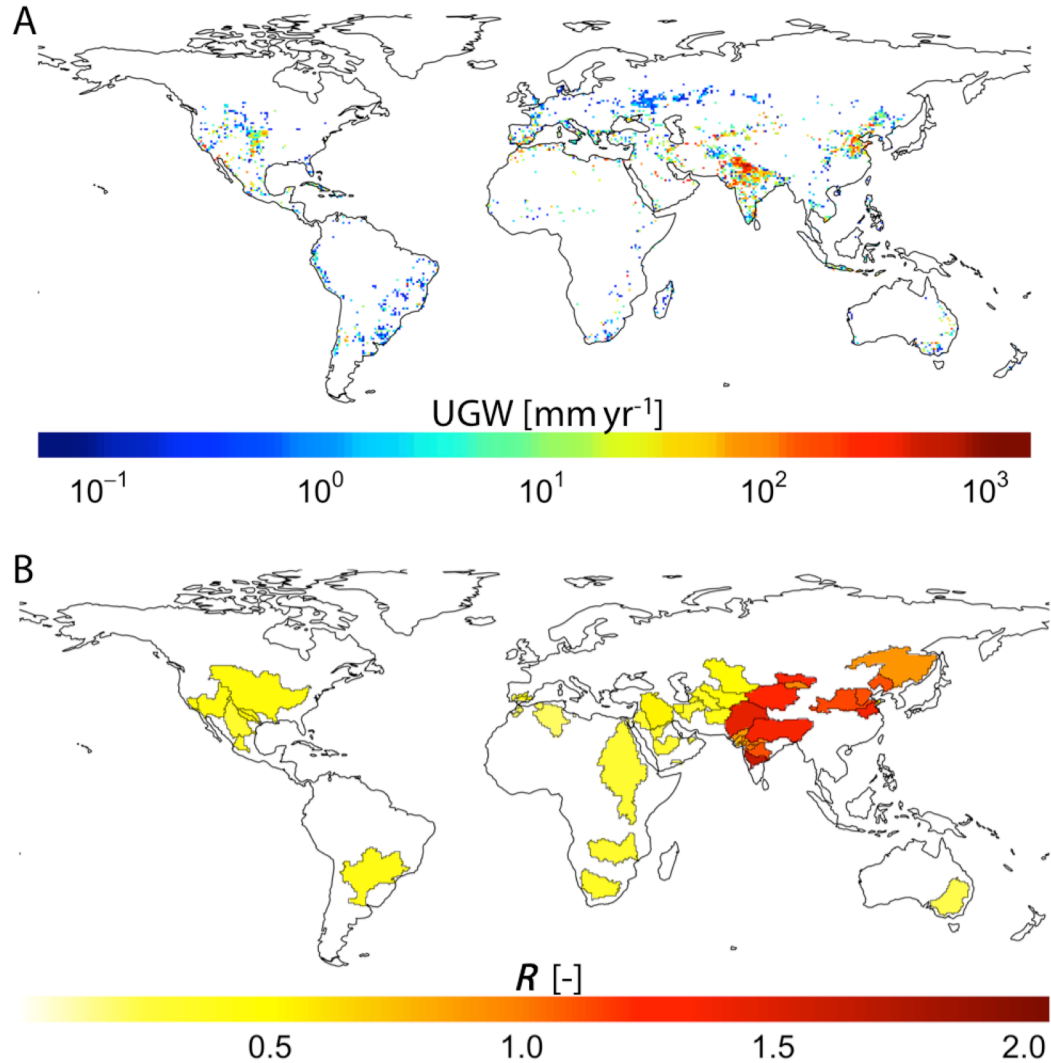
**Fig. 3.2:  Unsustainable groundwater pumped and reused. (A)** Average annual UGW demand [$km^3yr^{-1}$] at 0.5° spatial resolution. **(B)** Basin-level UGW reuse factor, **$R$,** aggregated from 0.5° grid cells, for all major river basins in which UGW demand is ≥1 $km^3yr^{-1}$. **$R$** is the number of times UGW is reused due to agricultural runoff and percolation.  **$R$** > 1.0 indicates that all UGW extracted is reused more than once.  Basins with a high **$R$** have a lower potential to reduce UGW dependency through increased irrigation efficiency.
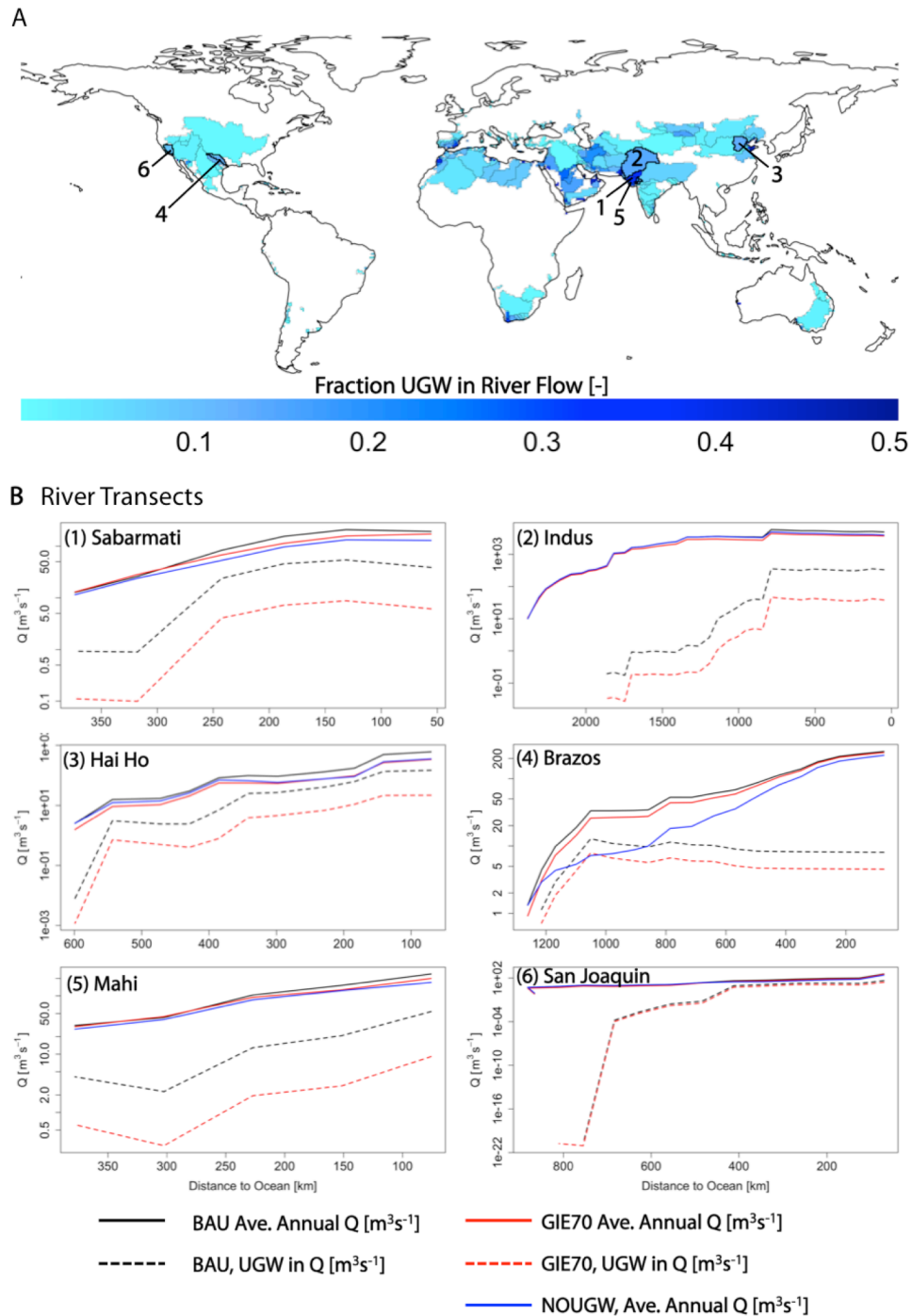
**Fig. 3.3: Unsustainable groundwater contribution to river flow. (A)** Basin-level average UGW % in river discharge. Basin-level values are the aggregate of grid-cell values. Basins and grid cells with average annual discharge <10 m³s⁻¹ or UGW < 0.1 % are not shown. **(B)** River transects of average annual discharge (black line, Q [m³s⁻¹]) following the mainstem of the six basins with the greatest volume of UGW used for irrigation by way of reuse through river systems (Table 3.1). Average annual Q is shown under both GIE70 (red dashed line) and NoUGW (blue line) scenarios. UGW within river discharge is shown for both the BAU simulation (black dashed line) and the GIE70 scenario (red dashed line).
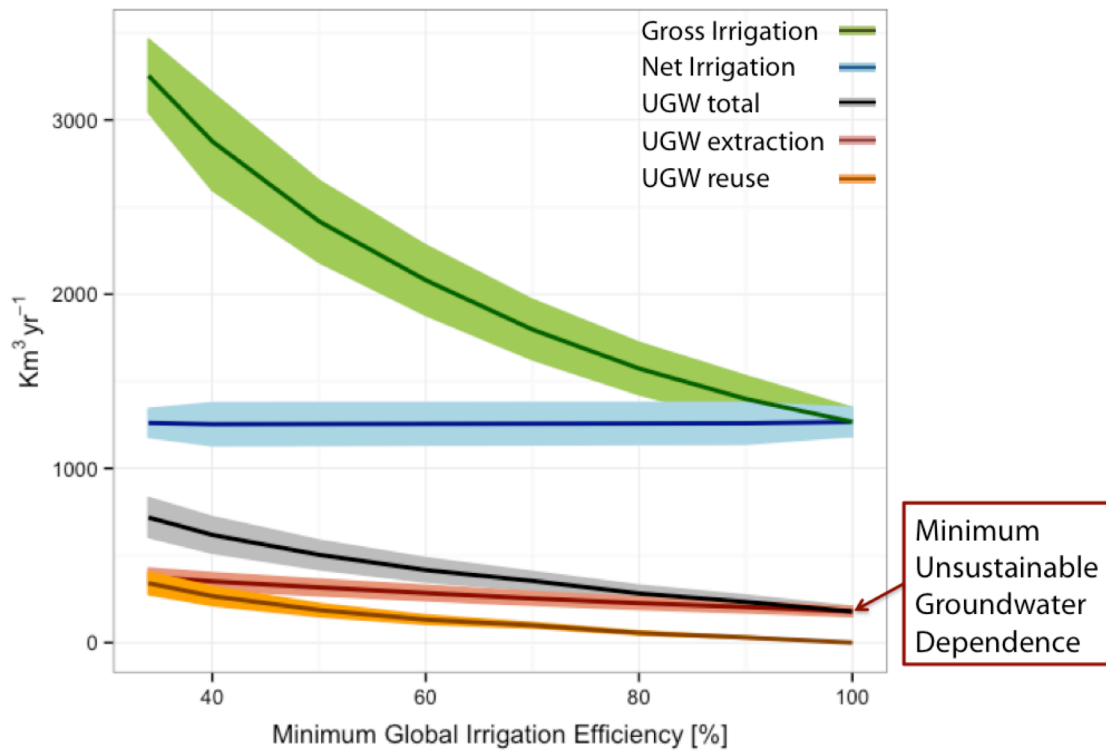
**Fig. 3.4:** Changes in global gross irrigation water requirements (green line), UGW (red line), UGW_r (orange), total UGW (black) as the minimum global irrigation efficiency increases. At 100% irrigation efficiency, 300 ($\pm$24) km$^3$yr$^{-1}$ UGW is still required to meet contemporary irrigation water requirements. Bands around each solid line show 1 standard deviation due to 30 years of inter-annual climate variability.

# Supplemental Information Tables

**Table S 3.1** Country-level comparison of WBM results to FAO data. Countries are listed in order of ascending irrigation water withdrawals. **IWW** is irrigation water withdrawals (FAO) and WBM gross irrigation water withdrawals. **SW** is fresh surface water withdrawals, and **GW total** is fresh groundwater withdrawals.

| | | IWW | | SW | | GW total | |
|---|---|---|---|---|---|---|---|
| Index | country | WBM | FAO | WBM | FAO | WBM | FAO |
| 1 | Lebanon | 1 | 1 | 0.3 | 0.4 | 1 | 1 |
| 2 | Guyana | 1 | 1 | 1 | | 0.1 | |
| 3 | Nigeria | 1 | 7 | 1 | | 0.3 | |
| 4 | United Republic of Tanzania | 1 | 5 | 1 | | 1 | |
| 5 | Armenia | 1 | 1 | 1 | 2 | 0.4 | 1 |
| 6 | Israel | 1 | 1 | 0.2 | | 1 | |
| 7 | Oman | 1 | 1 | 0.1 | 0.0 | 1 | 1 |
| 8 | Tunisia | 2 | 3 | 0.4 | 1 | 1 | 2 |
| 9 | Dominican Republic | 2 | 6 | 0.4 | | 1 | 0.4 |
| 10 | Kyrgyzstan | 2 | 7 | 1 | 7 | 1 | 0.3 |
| 11 | Colombia | 2 | 6 | 2 | 11 | 0.1 | 1 |
| 12 | Canada | 2 | 5 | 1 | 40 | 1 | 2 |
| 13 | Malaysia | 3 | 3 | 2 | 5 | 0.5 | 0.2 |
| 14 | Mali | 3 | 5 | 3 | 5 | 0.3 | 0.1 |
| 15 | Libya | 3 | 4 | 0.1 | 0.0 | 3 | 4 |
| 16 | Azerbaijan | 3 | 10 | 3 | 11 | 0.4 | 1 |
| 17 | Uruguay | 3 | 3 | 2 | 4 | 1 | 0.1 |
| 18 | Malawi | 3 | 1 | 1 | | 2 | |
| 19 | Chile | 3 | 29 | 1 | 33 | 2 | 3 |
| 20 | Tajikistan | 3 | 10 | 3 | 9 | 0 | 2 |
| 21 | United Arab Emirates | 3 | 3 | 0 | 0 | 3 | 3 |
| 22 | Cuba | 4 | 5 | 1 | 4 | 3 | 3 |
| 23 | Ukraine | 4 | 1 | 3 | | 1 | |
| 24 | Portugal | 5 | 6 | 3 | | 2 | |
| 25 | France | 5 | 3 | 2 | 25 | 2 | 6 |
| 26 | Peru | 5 | 12 | 3 | 12 | 2 | 2 |
| 27 | Kazakhstan | 6 | 14 | 4 | 19 | 1 | 1 |
| 28 | Yemen | 6 | 3 | 0 | 1 | 6 | 2 |
| 29 | Cambodia | 6 | 2 | 6 | | 1 | |
| 30 | Greece | 7 | 8 | 2 | 6 | 4 | 4 |
| 31 | Ecuador | 8 | 8 | 5 | | 2 | |
| 32 | Morocco | 8 | 9 | 2 | 8 | 6 | 2 |
| 33 | Algeria | 8 | 4 | 1 | | 7 | |
| 34 | Italy | 10 | 13 | 6 | | 4 | |

| 35 | Sudan | 10 | 26 | 8 | | 2 | |
|----|-------|-----|-----|-----|-----|-----|-----|
| 36 | Afghanistan | 10 | 20 | 5 | 17 | 5 | 3 |
| 37 | Turkmenistan | 10 | 26 | 7 | 28 | 3 | 0 |
| 38 | Nepal | 10 | 9 | 7 | | 3 | |
| 39 | South Africa | 11 | 8 | 3 | | 7 | 1 |
| 40 | Argentina | 11 | 28 | 6 | 26 | 5 | 11 |
| 41 | Madagascar | 11 | 16 | 2 | | 9 | 0 |
| 42 | Saudi Arabia | 11 | 21 | 1 | 1 | 11 | 22 |
| 43 | Turkey | 12 | 34 | 6 | 32 | 6 | 12 |
| 44 | Australia | 13 | 13 | 4 | | 10 | 5 |
| 45 | Sri Lanka | 14 | 11 | 13 | | 1 | |
| 46 | Myanmar | 18 | 30 | 17 | 30 | 1 | 3 |
| 47 | Iraq | 23 | 52 | 18 | | 4 | |
| 48 | Spain | 24 | 20 | 10 | 26 | 14 | 6 |
| 49 | Mexico | 28 | 62 | 12 | 49 | 16 | 30 |
| 50 | Brazil | 28 | 45 | 18 | | 10 | |
| 51 | Japan | 31 | 55 | 25 | 73 | 6 | 16 |
| 52 | Philippines | 31 | 67 | 28 | 78 | 4 | 3 |
| 53 | Uzbekistan | 35 | 50 | 23 | 44 | 12 | 5 |
| 54 | Egypt | 35 | 59 | 27 | | 9 | 7 |
| 55 | Thailand | 82 | 52 | 69 | 47 | 13 | 10 |
| 56 | Indonesia | 109 | 93 | 74 | 96 | 35 | 18 |
| 57 | Bangladesh | 116 | 32 | 84 | 7 | 31 | 28 |
| 58 | United States of America | 116 | 192 | 66 | 367 | 50 | 108 |
| 59 | Pakistan | 303 | 172 | 202 | 122 | 101 | 62 |
| 60 | China | 799 | 358 | 596 | 453 | 203 | 101 |
| 61 | India | 1017 | 688 | 615 | 397 | 402 | 251 |

## A Fraction of Q as UGW
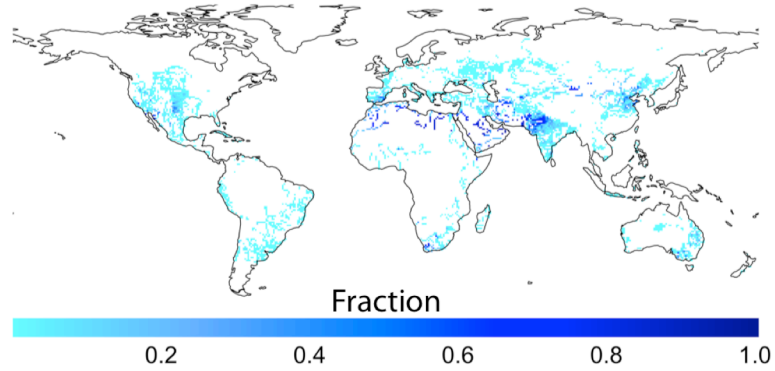


## B Irrigation efficiency = 70%
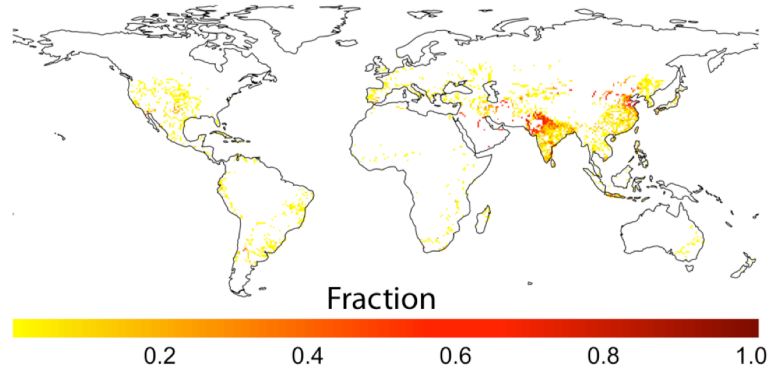


## C No Unsustainable groundwater use
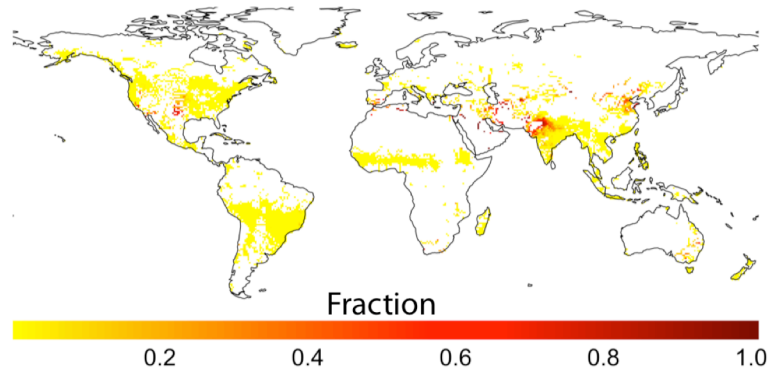


**Fig. S 3.1. UGW impacts surface water flows. (A)** Grid-cell level fraction of river discharge composed of UGW. **(B)** Grid-cell level decrease (fraction) in average annual river discharge in the GIE70 simulation. **(C)** Grid-cell level decrease (fraction) in average annual river discharge in the NoUGW simulation.
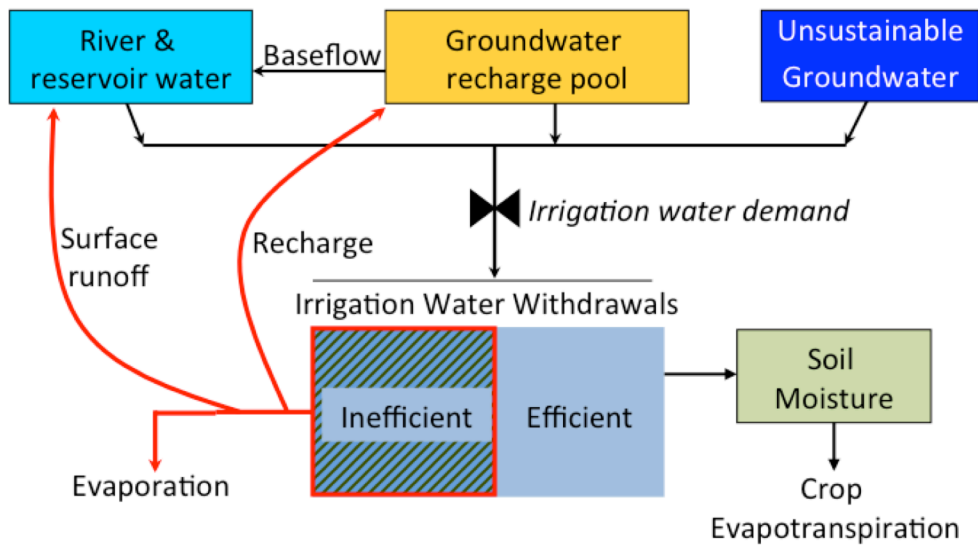
**Fig. S 3.2:** Stocks and flows of water in WBM, including unsustainable groundwater (UGW). Unless otherwise specified (as in GIE70) UGW is assumed to be fully available to meet irrigation demand. UGW is only connected to other water stocks and flows via irrigation withdrawals. UGW is tracked through all these stocks and flows.
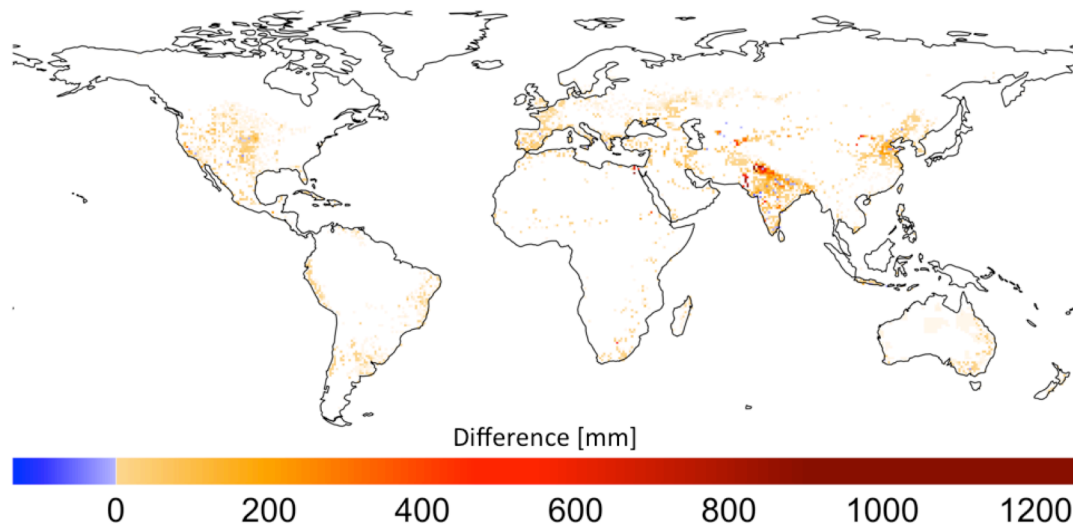


**Fig. S 3.3.** Average annual difference in UGW [mm] between a model simulation with a search distance of 0 km and a with a search distance of 150 km. Blue values indicate UGW is lower with a 0 km search distance (minimum value: -132 mm), and red values indicate UGW is higher with a 0 km search distance than a 150 km search distance.
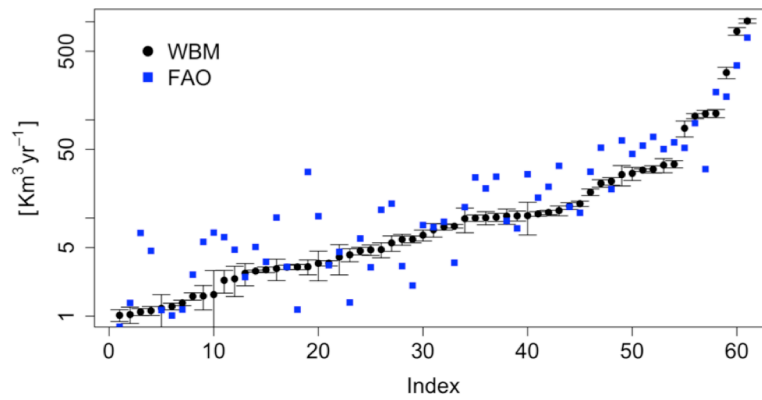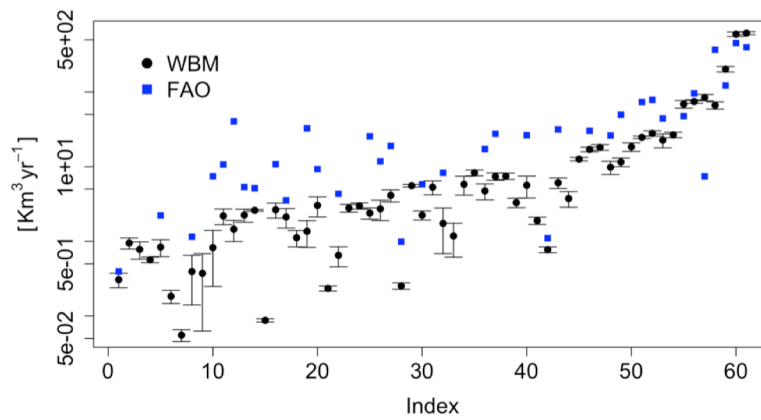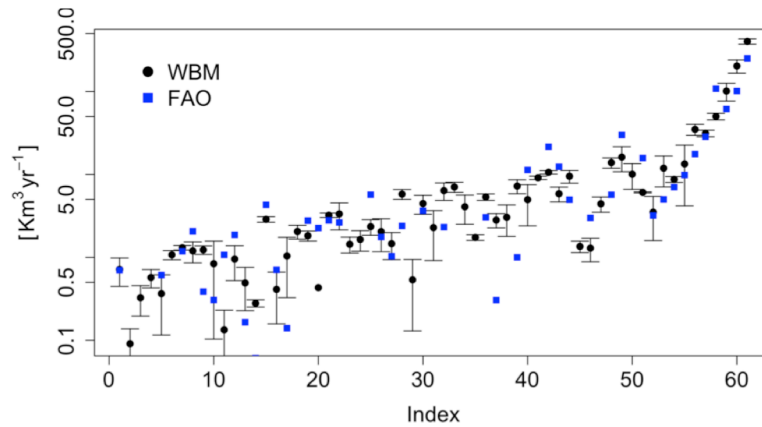
**Fig. S 3.4. Country-level comparison to FAO.** In all panels, WBM simulation results are black dots with one standard deviation shown as error bars; FAO year-2000 reported data are shown in blue squares. Note the log axis on the y scale in all panels. See Table S 3.1 for index of country names. **(A)** Country-level irrigation water withdrawals; **(B)** Country-level surface water irrigation water withdrawals; **(C)** Country-level total groundwater irrigation water withdrawals.
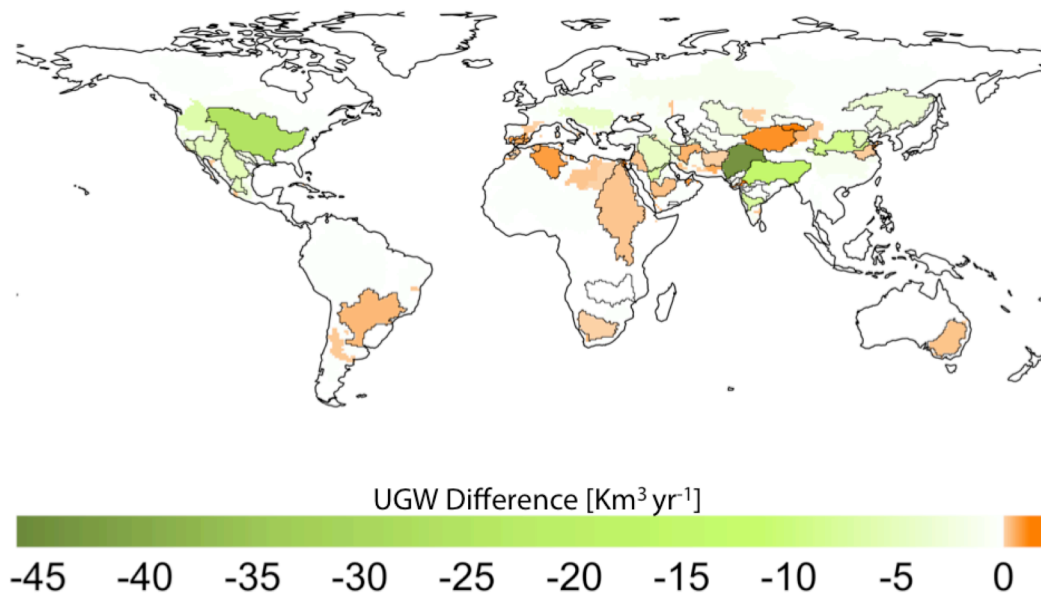
**Fig. S 3.5. Basin-level comparison of UGW.** Green indicates that the mean (of 30 years of climate inputs from each of 4 different climate data products) WBM-simulated annual UGW is lower than the basin-aggregate difference between groundwater extraction and recharge reported in Gleeson et al (2012). Orange (maximum value: 1.46 $km^3yr^{-1}$) indicates WBM-simulated results are higher.

CONCLUSIONS

Agriculture's reliance on unsustainable groundwater has been recognized by numerous studies and reports (e.g., Giordano and Villholth, 2007; Siebert and Doll, 2010; Aeschbach-Hertig and Gleeson, 2012; Wada et al., 2012), with widespread agreement that changes to groundwater management will be necessary in the future if we are to both produce more food and achieve sustainable levels of groundwater extraction (FAO, 2003; Giordano and Villholth, 2007; Shah, 2009; Foley et al, 2011). In order to achieve these goals, we must first understand what happens to groundwater after it is extracted; simply accounting for the volumes currently removed from aquifers is not sufficient due to the spatial agricultural networks created by river systems, as well as changes to the water cycle due to both climate change and human-built infrastructure. This dissertation contributes to understanding the impacts of over-extracting groundwater now and into the future by answering the following questions: How much food does unsustainable groundwater grow, and what are the regional differences in agriculture's reliance on this resource? How will this change into the future, considering both climate change and human responses to climate change? What happens to the non-consumed portion of groundwater extractions? How can constructed water infrastructure alter agriculture's demand for unsustainable groundwater?

Chapter I quantifies the amount of crop production directly attributable to unsustainable groundwater irrigation in China. We found that without unsustainable groundwater, grain production would be reduced by 15-27% (range due to 20 years of climate variability), representing a reversion to 1980s levels of agricultural production. The North China Plain alone would suffer a loss of 101 million tons of carbon (crop production)

per year, a 10% loss in national production. While 15-27% is a large portion of national crop production, it is worth noting that unsustainable groundwater currently supplies 20-49% of all irrigation water in China, indicating a non-linear relationship between reliance on unsustainable groundwater and crop production due to use of this groundwater resource. Even in the case that unsustainable groundwater supplies 49% of all irrigation water supplies, crop production suffers only a 27% loss. This non-linearity is due to regional differences in both the volumes of groundwater needed for irrigation and the amount by which irrigation boosts crop productivity.

One potential strategy for reducing agriculture's reliance on unsustainable groundwater in China would be to plan future agricultural expansion and intensification in regions with both low relative reliance on unsustainable groundwater and small gains in crop production due to using irrigation water in excess of available surface water supplies. An alternative approach would be to construct canals to move surface water into regions that are currently relying on unsustainable water sources and that also have large yield gains due to irrigation. The Chinese government has already committed to the latter approach, and are constructing the North-South Water Transfer which will move water from the Yangtze to the Yellow River. Based on the conclusions of Chapter I, providing more water to the Yellow River – which passes through the important agricultural regions of Henan, Shandong, and Shaanxi – is likely to alleviate demand for unsustainable groundwater. Further work is required to fully assess the hydrologic and agricultural impacts of the North-South Water Transfer, as this transfer will not only increase the Yellow River flow, but also decrease the Yangzte river flow and could therefore have unintentional downstream consequences. Additionally, the combined hydrologic-crop production model framework

provided here could be used to compare the impact of the North-South Water Transfer to different methods of water management and agricultural development.

Chapter II assesses the impact of humans, climate change, and water infrastructure on the future of unsustainable groundwater demand in India. The most important aspect of this work is the interdisciplinary collaboration, highlighting the role of (aggregate) farmer-level human decisions on the trajectory of water resources into the future. While the physical sciences have thoroughly shown that there are many possible trajectories for climate change, combined econometric-hydrologic model results from Chapter II emphasize that for each climate trajectory, there is a matching human response. It will be the combined climate-human trajectory that ultimately determines irrigation water demand, and the future reliance of India's agricultural system on unsustainable groundwater.

Chapter II also simulates a future in which a massive water infrastructure project has been completed in India. India's National River Linking Project (NRLP) involves both a series of canals, and a large number of "storages". While the location of the canals, along with some details of transfer capacities, have been publicized, the "storages" component of the project remains extremely vague in NRLP-related publications (Amarasinghe et al, 2008). Results in Chapter II show that implementing NRLP canals without concurrently increasing reservoir storage capacities along the canal routes has little ability to alleviate India's unsustainable groundwater demand (1-4% alleviation nationally). Increased storage capacity can boost the potential of the NRLP to alleviate unsustainable groundwater demand to 16% nationally. Clearly, it would be helpful to all researchers interested in India's NRLP to know both the location and storage capacities of the planned NRLP storages; whether or not these storages are even agreed upon by the government implementers of the project are

unclear.  Historically, construction of new large reservoirs has been contentious in India (and in other countries) due to their impact on ecosystems, downstream river flows, and human displacement (Dhawan, 1989).  Recognition for the need to increase water storage capacity in India has lead to proposals for small tanks and ponds (e.g., Wisser et al, 2010), as well as increased artificial recharge (Sakthivadivel, 2007; Amarasinghe, 2016).  Further work is needed to quantifying the potential for each of these possible infrastructure solutions to alleviate groundwater stress, both individually and as concurrent efforts.

Projections of changes in the rates of regional groundwater level declines in India are included in Chapter II.  These projections are based on the fundamental fact that extracting groundwater in volumes that far exceed recharge (which is the case across much of India, including extractions two orders of magnitude greater than recharge in northwestern India) leads to declines in groundwater levels.  The current volume of water contained within India's aquifers is unknown as of this date, and therefore these projections cannot be used to estimate when aquifer reserves will run out.  However, complete aquifer depletion is unlikely to be the first cause of unsustainable groundwater "loss" – rather, limitations to human access will most likely occur first, as groundwater levels drop below both economic and technological thresholds in well drilling and pumping.  These thresholds have inspired the authors of Chapter II's associated publication to propose new research to build upon the results in Chapter II.  This proposed research includes: a) developing an improved groundwater table module for the hydrologic model WBM that will more accurately represent changes in the groundwater table due to pumping, and b) completing the feedback loop between the hydrologic model and economics model used in Chapter 2,

allowing not only the hydrologic system to respond to human-based changes, but also the human-economic system to respond to changes in surface and groundwater supplies.

Chapter III presents the first global budget of groundwater use and reuse, illustrating the importance of groundwater to agriculture and surface water systems after it has been extracted from aquifers. We show that despite a ratio of global net irrigation to gross irrigation of ~ 38% (indicating that 62% of water extracted for irrigation is "wasted" due to inefficiencies), 63% of the total volume of extracted unsustainable groundwater becomes crop evapotranspiration. Approximately one third of this crop evapotranspiration is due to reuse of unsustainable groundwater. Because of reuse through agricultural systems, increasing irrigation efficiency – decreasing the "wasted" portion of irrigation water extractions – can reduce the demand for unsustainable groundwater, but it cannot eliminate it. This result shows that while improved irrigation efficiency may be one way to reduce global agriculture's reliance on unsustainable groundwater, it should not be the only strategy employed by water resource managers, particularly in regions where river flows are highly dependent upon groundwater entering surface water systems. Furthermore, any assessment of future surface water resources – whether for irrigation, domestic or industrial water supply, or hydropower production – should take into consideration potential reductions in surface water flows due to decreased groundwater extractions in the case that irrigation efficiencies improve or access to groundwater resources is reduced.

The United Nations has declared access to sufficient, safe, and accessible water to be a human right (United Nations General Assembly, 2010). Assuring this right is realized on a global level will require protection against the extremes of climate, as precipitation patterns and temperature both contribute to droughts, and climate projections point to the

increased possibility of long-term mega-droughts (Cook, Ault, and Smerdon, 2015). Groundwater acts as a natural reservoir, built over millions of years, and is often of good quality for drinking and agricultural use. While some groundwater aquifers are drawn down only in times of drought others have become a baseline resource for agriculture, drinking, and economic development (Giordano and Willholth, 2007). Recognizing that many of the world's large aquifers are being depleted (Aeschbach-Hertig and Gleeson, 2012), water managers and national- to international planning organizations are seeking ways to transition away from unsustainable groundwater use (United Nations General Assembly, 2010, 2015; FAO, 2011). The results presented in this dissertation help quantify agriculture's current reliance on groundwater, as well as assess possible strategies for reducing this reliance in the future. More work must be done to identify local, regional, and global strategies for achieving groundwater sustainability; the cross-disciplinary (hydrology-crop biogeochemistry, and hydrology-econometrics) frameworks presented here can be used to assess a wide range of strategies of water management and irrigation system development. These frameworks can also be built upon, further developing the areas of macro-scale groundwater modeling, hydrology-economics system modeling, and infrastructure optimization, all of which are identified as important areas of further study by the results presented here.

# APPENDIX A
# Water Balance Model Documentation

## A1  WBM purpose and scope

The Water Balance Model (WBM) was first developed by the Water Systems Analysis Group at the University of New Hampshire, Durham, (C. J. Vörösmarty, 1990s). Since then, it has branched into a family of models including WBMplus (e.g., Wisser et al., 2008, 2010) and FrAMES (e.g. Wollheim et al., 2008a,b). This appendix describes the version of WBMplus that was used for all studies within this dissertation.

WBM is a variable-scale, gridded model that simulates both the vertical exchange of water between the ground and the atmosphere, and the horizontal transport of water through runoff and stream networks. In addition to the natural water cycle, WBM can simulate human use and management of water, such as irrigation withdrawals and water transfers through canals. The core of the model is a water accounting system, tracking all water entering and leaving each grid cell. There are optional components of the model that can be turned on or off, including: large reservoirs, small reservoirs, dam operation rules, irrigated and rainfed croplands, inter-basin water transfers, and glacier melt.

## A2  Spatial and temporal resolution

WBM can be run at various spatial resolutions. It can also simulate grid cells over the entire land surface of the world, or be restricted to specific geographic regions. WBM calculates all water flows and stocks at a daily time step.

## A3  Model inputs

WBM takes data as gridded static layers (maps), gridded time series, or geographically referenced databases. Non-gridded data formats such as shapefiles can be pre-processed and turned into a gridded format. Gridded files with spatial resolutions different from the river network resolution can be aggregated (if a finer resolution) or resampled (if larger resolution) to match the river network spatial resolution.

Table A1: **WBM input data.** Inputs with the note "optional" after them will be given a default value in the model if the associated process is implemented but no input is provided.

(a) Required inputs

| Process | Inputs |
| --- | --- |
| Soil moisture, runoff, & evapotranspiration | Soil wilting point map |
| | Soil field capacity map |
| | Rooting depth map |
| | Soil available water capacity map (optional: can be substituted for above three map inputs) |
| | Precipitation time series |
| Evapotranspiration | Temperature time series |
| River flow | River network |

(b) Optional inputs

| Process | Inputs |
| --- | --- |
| Canopy interception of precipitation | Leaf area index time series |
| Reservoir storage and release | Large reservoir database |
| Small reservoir storage | Small reservoir capacity map |
| Inter-basin water transfers | Inter-basin transfer database |
| Storm runoff over impervious areas | Impervious area map |
| Glacier melt contribution to runoff | Glacier area map |
| | Glacier melt time series |
| | Temperature and precipitation time series |
| Water temperature | Leaf area index time series |
| | Cloud fraction time series |
| | Wind direction & speed time series |
| | Humidity time series |
| | Groundwater temperature time series |
| | Canopy height time series |
| Surface processes over different land cover/land use types | Land cover/land use map |
| | Leaf area index map associated with land cover/land use map |
| Irrigation water demand & use | Irrigation efficiency map |
| | Irrigated cropland time series |
| | Crop coefficient ($Kc$) time series |
| | Crop available water time series |
| | Rice paddy percolation rate time series |
| | Ratio of surface water to groundwater extractions map (optional) |

146

## A4  Potential Evapotranspiration

Potential evapotranspiration, PET, is the maximum amount of water that can be lost from soil through combined evaporation and transpiration, assuming no shortage of soil water. It provides an upper bound on non-irrigated actual evapotranspiration, and is used as a baseline reference for calculating irrigated evapotranspiration. There are many methods available for calculating PET. Vörösmarty, Federer, and Scloss (1998) compared 11 different methods of calculating daily potential evapotranspiration for use in a 30 min global gridded water balance model. These methods fall into two general categories: the first category is reference-surface PET, defined as the evapotranspiration that would occur from a hypothetical "reference crop" under given climate conditions; the second category is surface-dependent PET, defined as the evapotranspiration that would occur under a given (variable) land cover or surface.

WBMplus uses the Hamon method (Hamon,1963) to calucalte PET: $PET_H$ [mm]. This is the least data-intensive method, and it was found to estimate global average PET as well as other, more data-intensive methods. Additionally, Vörösmarty (1998) found that amongst the reference-surface PET methods, the Hamon method produced both the lowest mean annual error and the smallest bias when compared to obseravation data. Previous versions of WBMplus, as well as the FrAMES member of the WBM model family (Wollheim et al., 2008a,b) have alternative PET function options.

$$PET_H = 330.2 \, \Lambda \, \rho_{sat} \tag{A.1}$$

where: $T$ = daily mean temperature

$\Lambda$ = day length, expressed as a fraction of a 12-hr period,

$$\rho_{sat} = 2.167 P_{sat}/(T + 273.15), \text{ and} \tag{A.2}$$

$$P_{sat} = \begin{cases} 0.61078 e^{(17.26939\,T/(T+237.3))} & 0 \leq T \\ 0.61078 e^{(21.87456\,T/(T+265.5))} & T < 0 \end{cases} \tag{A.3}$$

$T$ is in [° C], $\rho_{sat}$ is in units of [g m$^{-3}$], and $P_{sat}$ is in units of [kg m$^{-1}$s$^{-2}$].

## A5  Snow

While precipitation as rainfall adds water directly to the soil moisture balance, snow may form a snow pack, then melt at a later time. The WBMplus keeps track of snowfall, $P_s$ [mm], snow pack, $S_p$ [mm], and snow melt, $M_s$ [mm]. When mean daily air temperature is below the snowfall threshold, $T_s$ [° C] (default value $-1$°C) all precipitation is treated as snow; otherwise, all precipitation is treated as rain. When mean daily air temperature is above the snowmelt threshold $T_m$ [° C] (default value

1°C) a portion of the snow is melted. Each time step's snow pack is updated, taking into account precipitation as snow, and snowmelt.

In the following equations, the subscript "$i$" is used to indicate the current timestep, "$i$-1" for the previous time step. Snow melt is never allowed to exceed the size of the snow pack.

$$S_{p_i} = \delta S_p / \delta t + S_{p_{i-1}} \tag{A.4}$$

$$\text{where:} \quad \delta S_p / \delta t = P_s - M_s, \tag{A.5}$$

$$P_s = \begin{cases} P & T < T_s \\ 0 & T_s \leq T \end{cases} \quad \text{and,} \tag{A.6}$$

$$M_s = \begin{cases} 2.63 + 2.55 \cdot T + 0.0912 \cdot T \cdot P & T_m < T \\ 0 & T \leq T_m. \end{cases} \tag{A.7}$$

where $P$ is total precipitation [mm] (snow or rainfall).

## A6  Canopy interception

Vegetation canopies intercept precipitation, preventing it from becoming soil moisture. WBMplus calculates interception by the canopy, $I_c$ [mm] as:

$$I_c = s \cdot LAI \tag{A.8}$$

where $s$ [mm] is a storage capacity constant (default value 0.25 mm (Dickinson, 1984)), and $LAI$ is the leaf area index [-]. If no leaf area index input dataset is provided, $I_c = 0$.

## A7  Soil Moisture

Soil moisture balance, $W_s$ [mm], is calculated with an accounting system that tracks a grid cell's water inputs, water outputs, and soil moisture pool holding capacity. The soil moisture pool depth is determined by the rooting depth. Inputs come in the form of precipitation as rain, $P_r$ [mm], and as snow melt, $M_s$. Water intercepted by the canopy, $I_c$, reduces how much precipiatation reaches the soil. Output is via potential evapotranspiration, $PET$ (here, $PET = PET_H$), modified by a soil drying function, $g(W_s)$. The drying function is employed when water inputs are less than potential evapotranspiration. The amount of water that can be drawn out of the soil moisture pool depends on the current soil moisture, and the available water capacity (see Figure A1). The drying function $g(W_s)$ also depends on an empirical constant, $\alpha$ [-], (default value 5.0) in order to best match the drying curve of Pierce

(1958). The coefficient $\alpha$ can be adjusted to calibrate the model based on regional climate.

Available water capacity, $W_{cap}$ [mm], indicates the portion of the soil moisture storage pool within the grid cell that is available to receive water inputs. Available water capacity is determined by taking the difference between the field capacity, $F_{cap}$ [-], and the wilting point, $W_{pt}$ [-], each expressed as fractions of the total depth. This difference is then scaled by the total rooting depth, $R_d$ [mm], to determine the depth in mm which the grid cell has available to receive water. Field capacity, wilting point, and rooting depth are all input from global datasets based on soil and vegetation type. Alternatively, available water capacity $W_{cap}$ can be input directly into the model instead of calculated.

$$W_{s_i} = \frac{\delta W_s}{\delta t} + W_{s_{i-1}} \tag{A.9}$$

where:

$$\frac{\delta W_s}{\delta t} = \begin{cases} 0 & \text{if } W_{cap} = 0 \\ g(W_s) \cdot (P_a - PET) & \text{if } P_a < PET \\ P_a - PET & \text{if } PET \leq P_a \text{ and } (P_a - PET) < (W_{cap} - W_s) \\ W_{cap} - W_s & \text{if } PET \leq P_a \text{ and } (W_{cap} - W_s) \leq (P_a - PET), \end{cases} \tag{A.10}$$

where:

$$g(W_s) = \frac{1 - e^{(-\alpha W_s/W_{cap})}}{1 - e^{-\alpha}}, \tag{A.11}$$

$$W_{cap} = (F_{cap} - W_{pt})R_d, \tag{A.12}$$

$$P_a = P + M_s - I_c. \tag{A.13}$$

## A8  Evapotranspiration

Actual evapotranspiration cannot be equal to potential evapotranspiration, $PET$, when there is a soil mositure deficit. Therefore, we calculate an estimated actual evapotranspiration, $AET$ [mm]:

$$AET = \begin{cases} P_a - (\delta W_s/\delta t) & \text{if } (P_a - (\delta W_s/\delta t)) \leq PET \\ PET & \text{if } (P_a - (\delta W_s/\delta t)) > PET. \end{cases} \tag{A.14}$$
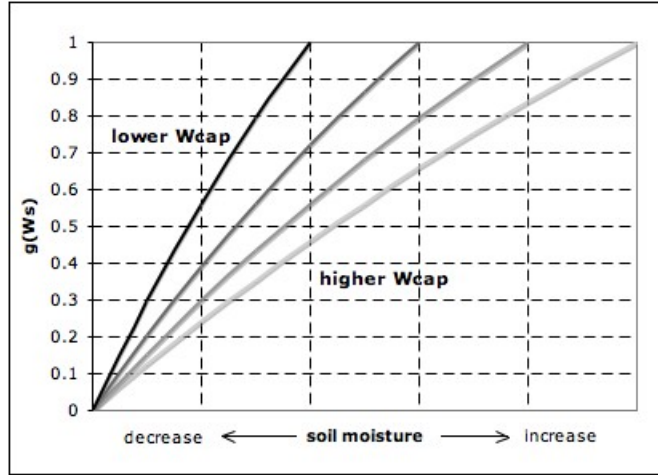
Figure A1: **Soil drying function schematic. Solid lines show how $g(W_s)$ changes with higher (gray) and lower (black) available water capacity, and with increasing and decreasing soil moisture. $g(W_s)$ is always $\leq 1$ because the amount of water removed from the soil moisture pool cannot exceed 100% of the available water capacity.**

## A9  Runoff and Baseflow

When water inputs to a grid cell exceed the amount of water that can be stored as soil moisture and be lost to evapotranspiration, then a surplus, $S$ [mm], is formed. A fraction $(1 - \gamma$ [-]) of this surplus becomes surface runoff, $R_s$ [mm]. The remaining fraction ($\gamma$ [-]) of the surplus infiltrates into the groundwater storage pool, $W_g$ [mm]. Water leaks out of the groundwater storage pool through baseflow, which is a fraction ($\beta$ [-]) of the total groundwater storage. The total change in groundwater is then the infiltration from surplus, (i.e., recharge), minus the loss to baseflow.

The combined surface runoff and baseflow exit the grid cell, and are collected in a river network that allows the water to be transported downstream, the details of which will be discussed in Section A 10. The total amount of water that exits the grid cell and enters the network (total runoff, $R_t$ [mm]) is the sum of the surface runoff and baseflow. $\gamma$ and $\beta$ are empirical values (defalut: 0.5 and 0.0167, respectively). Note that here, $\gamma$ is an infiltration fraction, setting how much of the surplus enters the groundwater pool. In Vörösmarty, 1998, $\gamma$ indicates a surface runoff fraction, setting how much of the surplus becomes surface runoff.

$$R_t = R_s + \beta W_g \tag{A.15}$$

$$\text{where: } R_s = (1 - \gamma)S, \tag{A.16}$$

$$\delta W_g/\delta t = \gamma S - \beta W_g, \tag{A.17}$$

$$W_{g_i} = W_{g_{i-1}} + \delta W_g/\delta t, \text{ and} \tag{A.18}$$

$$S = P_a - \delta S_p/\delta t - E_s - \delta W_s/\delta t \tag{A.19}$$

## A9.1 Runoff from impervious areas

Impervious areas prevent water from entering soils and increase overland runoff. If provided with an impervious area map, WBMplus calculates overland runoff in impervious areas, $R_{imp}$ [mm] as:

$$R_{imp} = C_{imp} A_{imp}(P_a + M_s) \tag{A.20}$$

where $C_{imp}$ [-] is a unitless scalar for impervious surfaces that determines the fraction of precipitation over impervious areas that is directly routed to rivers, $A_{imp}$ [m$^2$] is impervious area, $P_a$ is precipitation and snowmelt (minus canopy interception), and $M_s$ is snowmelt.

## A10 Water transport model

WBMplus transports surface runoff downstream from one grid cell to another, simulating a river network. There are two components to the water transport mode: 1) a river network map that identifies the direction of river flow within each grid cell, and 2) a water routing scheme that determines the flow velocity of river discharge. The river network is a model input; the Water Systems Analysis Group has developed such an input dataset, called the Simulated Topological Network (STN-30p), (Vörösmarty et al. 2000a; Vörösmarty et al., 2000b), described below in Section A 10.1. Other networks can be used, provided they include consistant river basin boundaries.

## A10.1 Simulated Topological Network (STN-30p)

The Simulated Topological Network (STN-30p) was first described in (Vörösmarty et al., 2000b) and (Vörösmarty et al., 2000c). To determine the maximum topographic gradient in each grid cell, (Vörösmarty et al., 2000c) first spatially aggregated (at 30 min longitude and latitude grids) the Global Gridded Elevation and Bathymetry (ETOPO5) 5-10 min DEM (Ewards, 1989), then used an ARC/INFO (ESRI, 1992)

algorithm to find the gradient. Each grid cell represents 1000 - 3000 km$^2$ of land surface, depending on latitude. Based on the maximum gradient, flow directions were assigned and given a direction of N, NE, E, SE, S, SW, W, or NW. While inflow of water to a grid cell can be from any and all directions, outflow is assigned to only one direction.

The STN-30p was checked against known stream locations, and manual reconfiguration was required to correct errors. The University of New Hampshire - Global Hydrological Archive and Analysis System (UNH-GHAAS) compared the STN-30p to atlases, regional maps, operational navigational charts, and other digital river maps (Figure A2). See (Vörösmarty et al., 2000b) for comparison details.

## A10.2  Flow routing

WBMpus estimates the flow rate and water level in each grid cell's stream using a distributed flow routing model based on the Saint-Venant partial differential equations for one-dimensional flow. Specifically, we use the Muskingum-Cunge kinematic wave model that approximates the solution to the Saint-Venant partial differential equations (Maidment, 1992). These equations require six assumptions:

1. Flow from grid $j$ to grid $j + 1$ is one-dimential,

2. The stream length through the grid cell is significantly larger than the flow depth,

3. Vertical acceleration and vertical changes in pressure are negligable,

4. Water density is constant,

5. Channel bed and banks are immobile, and

6. Channel bottom slope is small, less than 15%.

Additionally, WBM assumes a rectangular channel bed and no loss of water from the channel to groundwater

The Muskingum-Cunge solution estimates the outflow, $Q_{j+1}^{t+1}$ [m$^3$s$^{-1}$], at time $t+1$ and grid cell $j+1$, as a linear combination of three known inflows and outflows. These are: 1) the inflow of the current time step and previous grid cell, $Q_j^{t+1}$ [m$^3$s$^{-1}$], 2) outflow of the previous time step and current grid cell, $Q_{j+1}^t$ [m$^3$s$^{-1}$], and 3) inflow from the previous time step and adjacent upstream grid cell, $Q_j^t$ [m$^3$s$^{-1}$]:

$$Q_{j+1}^{t+1} = C_0 Q_j^{t+1} + C_1 Q_{j+1}^t + C_2 Q_j^t \tag{A.21}$$

The coefficients C$_0$ [-], C$_1$ [-], and C$_2$ [-], are defined such that:

$$C_0 + C_1 + C_2 = 1, \tag{A.22}$$
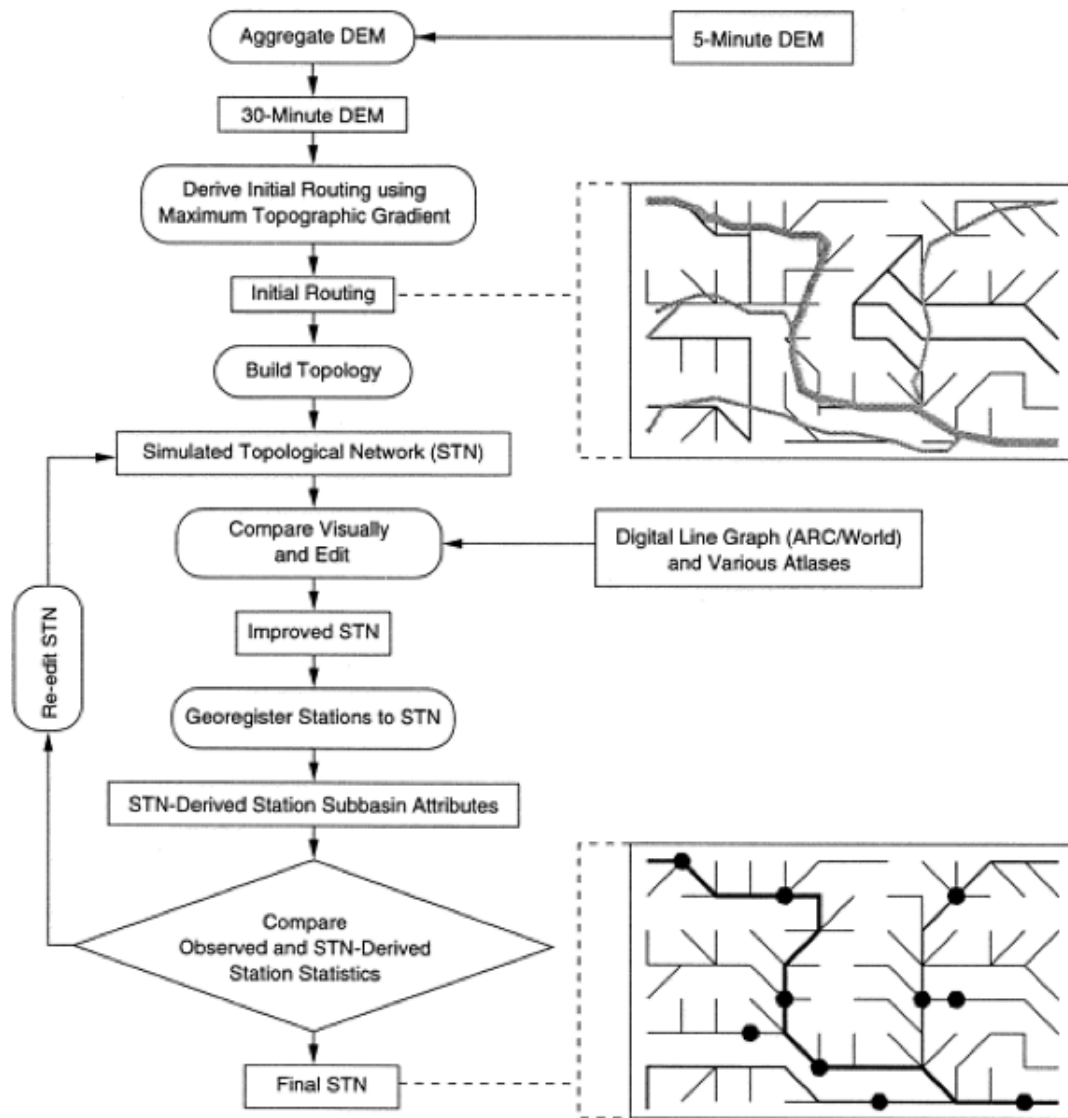
Figure A2: **Figure 1 from Vörösmarty et al. (2000b), showing the process of building and modifying data to form the STN-30p. Black dots in the lower right panel refer to river monitoring sites for which independent basin and sub-basin attributes were available.**

and if any of these three coefficients are less than 0, they are reset to 1, 0, and 0, respectively. The coefficients are unitless functions of the Courant number, $C$, and Reynolds number, $D$.

$$C_0 = \frac{-1 + C + D}{1 + C + D} \tag{A.23}$$

$$C_1 = \frac{1 + C - D}{1 + C + D} \tag{A.24}$$

$$C_2 = \frac{1 - C + D}{1 + C + D} \tag{A.25}$$

Both $C$ and $D$ depend on riverbed geometry, and are defined as:

$$C = U_w V_m \frac{\delta t}{L} \tag{A.26}$$

$$D = \frac{Y_m}{S_0 U_w L} \tag{A.27}$$

where $U_w$ [m$^3$ s$^{-1}$] is the speed of wave propogation (also referred to as the wave celerity), $V_m$ is the mean fluid velocity [m s$^{-1}$], $L$ is the river length in the grid cell [m], $\Delta t$ [s] is the time step length (daily), $Y_m$ is the mean flow depth [m], and $S_0$ is the riverbed slope [m km$^{-1}$]. These variables are defined or calculated as:

$$U_w = 1 + \frac{\frac{2}{3}\sigma}{\sigma + 1} \tag{A.28}$$

where the shape parameter $\sigma = 2$ [-],

$$V_m = \frac{Q_m}{Y_m W_m} \tag{A.29}$$

where $Q_m$ is the mean annual discharge in the river segment [m$^3$s$^{-1}$] , and $W_m$ is the corresponding mean annual channel width [m]:

$$W_m = \tau Q_m^{\phi} \tag{A.30}$$

where $\tau$ [-] and $\phi$ [-] are constants 8.0 and 0.58, respectively (Knighton, 1998).

$$L = N\sqrt{A_c} \tag{A.31}$$

$$\text{where } N = \begin{cases} 1 & \text{for flow directions N, S, E, W} \\ \frac{1}{sin(\pi/4)} & \text{for flow directions NW, SW, NE, SE} \end{cases} \tag{A.32}$$

154

where $A_c$ is the area of the grid cell [m$^2$].

$$\delta t = 86,400 = \text{seconds in one day}, \tag{A.33}$$

$$Y_m = \eta Q_m^\nu \tag{A.34}$$

where $\eta$ and $\nu$ are empirical constants of 0.25 and 0.4, respectively (Knighton, 1998), and

$$S_0 = \frac{0.1}{1000}. \tag{A.35}$$

As the discharge is calculated for each time step within a grid cell, the discharge value is stored so that it can be used to determine the mean annual discharge in future calculations.

When irrigation water is withdrawn from a grid cell's stream, discharge is adjusted to reflect the removal:

$$Q_{j+1}^{t+1} = C_0 Q_j^{t+1} + C_1 Q_{j+1}^t + C_2 Q_j^t \tag{A.36}$$

Grid cells which are defined as open water (e.g., lakes) use a flush routing scheme, in which water is transported immediately between the grid cell and the open water outlet point. In this case, the coefficiencts $C_0$, $C_1$, and $C_2$ are re-defined as open-water coefficients $C_{0_o}$ [-], $C_{1_o}$ [-], and $C_{2_o}$ [-]:

$$C_{0_o} = 1 \tag{A.37}$$

$$C_{1_o} = C_{2_o} = 0 \tag{A.38}$$

An alternative routing scheme can be used in WBMplus: Linear reservoir routing. Linear reservoir routing re-defines the coefficient $C_0$:

$$C_0 = 1/(1 + L\frac{s_d}{V_f/\delta t}) \tag{A.39}$$

where $s_d$ is a conversion factor of 3600 [s d$^{-1}$],

$$L = \begin{cases} \sqrt{A_C} & \text{if } N = 1 \\ \sqrt{A_C} \cdot \frac{1}{sin(\pi/4)} & \text{if } N \neq 1 \end{cases} \tag{A.40}$$

$$V_f = 2.18(1 - 0.077 \cdot log(\frac{S_c}{0.5})) \tag{A.41}$$

where $L$ [m] is the river length, $V_f$ [m h$^{-1}$] is the flow velocity, $S_c$ is the size of the cell expressed in degrees. Note that the equation for $V_f$ applies a correction factor of 2.18 to the ratio of the grid cell size as compared to a grid cell size of 0.5°; this follows Fekete et al. (2001).

Table A2: **Parameter values for large reservoirs.**

| Reservoir type | $D_{min}$ | $S_{opt}$ | $a$ | $b$ |
|---|---|---|---|---|
| Generic | 0.2 | 0.8 | 2/3 | 10 |
| Hyropower | 0.2 | 0.9 | 1 | 40 |
| Irrigation | 0.1 | 0.8 | 0.5 | 10 |
| Natural lake | 0.0 | 0.5 | 2/3 | 1.6 |

## A11   Reservoirs

Reservoirs in WBMplus are divided into two classes: large and small. Large reservoirs interact with the transport of water by altering discharge out of their grid cells, are are explicitly represented as a part of the simulated river network. They represent large lakes as well as constructed water impoundments. This is in contrast to small reservoirs, which in WBMplus are treated as water storage outside the river network that intercepts local runoff in non-irrigated grid-cell areas. Small reservoirs represent farm ponds and tanks (Wisser et al. 2010).

### A11.1   Large Reservoirs

Large reservoirs alter the flow of rivers by storing water and potentially preventing it from traveling downstream. Large reservoir discharge at time $t$, $D_t$ [m$^3$s$^{-1}$], is:

$$D_t = \begin{cases} \bar{Q}(D_{min} + ln(kS^a + 1)) & \text{if } S_l < S_{opt} \\ \bar{Q} \times e^{b(S_l - S_{opt})^2} & \text{if } S_l \geq S_{opt} \end{cases} \tag{A.42}$$

where $\bar{Q}$ is the 5-year average annual discharge [m$^3$ s$^{-1}$], $D_{min}$ is the minmum allowed reservoir release [-], $S_l$ is the reservoir storage at time $t$ [fraction of maximum storage volume], $S_{opt}$ is the optimal level of reservoir storage [fraction of maximum storage volume], $a$ and $b$ are calibrated parameters, and:

$$k = \frac{1}{S_{opt}^\alpha}[exp(1 - D_{min}) - 1] \tag{A.43}$$

Except for $\bar{Q}$, all parameters can be calibrated. WBMplus sets the values of these paramters based on the type and purpose of the reservoir, as defined in the reservoir dataset (Table A2).

### A11.2   Small Reservoirs

WBMplus representation of small reservoirs for the purpose of irrigation was develped by Wisser et al. (2010). Unlike large reservoirs, small reservoirs do not

intersect river segments; rather, they collect a portion of surface runoff from the non-irrigated part of a grid cell and store it as a separate stock. This stored water can then be used to supply water for irrigation to the irrigated portion of a grid cell. See Section A 13.1.1 for further details about irrigation.

$$A_{SR} = \frac{C_{SR}}{d_{SR}} \tag{A.44}$$

where $A_{SR}$ is the small reservoir area [m$^2$], $C_{SR}$ is the small reservoir capacity [m$^3$], and $d_{SR}$ is the depth of the small reservoir [m] (default value 2m). Water is only removed from the small reservoir storage pool when irrigation water is required within the grid cell.

Evaporation from small reservoirs, $E_{SR}$ is calculated as:

$$E_{SR} = 0.6 \times PET_H \times \frac{A_{SR}}{1000} \tag{A.45}$$

where $PET_H$ is potential evapotranspiration (see Section A 4). $E_{SR}$ cannot be larger than the amount of water stored within the reservoir, so there is an upper bound of $S_{SR}$.

## A12   Inter-basin transfers

Inter-basin transfers are canal systems that move water from one river basin into another. In this section, "donor" referes to the location from which water is being moved, and "recipient" refers to the location to which water is being moved by an inter-basin transfer. WBMplus uses eight paramters to simulate these transfers. There parameters are (1) the donor latitude and longitude, (2) the recipient latitude and longitude, (3) a minimum permitted donor river flow, $Q_{min}$ [m$^3$s$^{-1}$], (4) a maximum permited donor river flow, $Q_{max}$ [m$^3$s$^{-1}$], (5) a fraction of flow volumes to be removed from the donor river when discharge is between the minimum and maximum, $F$ [-], (6) the year in which the inter-basin transfer is implemented, (7) the length of the transfer canal, and (8) the width of the transfer canal. For parameters (1) and (2), all relevant latiutudes and longitudes are associated with a specific river network, as hand-checking is required to assure that the correct river is being identified for transfers by the simulation. The volume of water transfered, $Q_T$ [m$^3$s$^{-1}$], through each inter-basin transfer canal is:

$$Q_T = \begin{cases} 0 & \text{if } Q_d \leq Q_{min} \\ (F(Q_d - Q_{min}) & \text{if } Q_{min} < Q_d \leq Q_{max} \\ Q_{max} & \text{if } Q_d > Q_{max} \end{cases} \tag{A.46}$$

where $Q_d$ is the donor river discharge on day $d$. The transfer volume $Q_T$ is adjusted to $Q_{T_a}$ for small transfer volumes (volume measured in m$^3$s$^{-1}$):

$$Q_{T_a} = \begin{cases} 0 & \text{if } Q_T < 0.01 \\ Q_T & \text{if } Q_T \geq 0.01 \end{cases} \tag{A.47}$$

Evaporation from open water along the inter-basin transfer canal is removed from the transfer volume, resulting in the arrival volume $Q_a$:

$$Q_a = \begin{cases} Q_{T_a} - E_c \text{ if } (Q_{T_a} - E_c) > 0.001 \\ 0 \qquad\quad \text{if } (Q_{T_a} - E_c) \leq 0.001 \end{cases} \tag{A.48}$$

where $E_c$ is the canal evaporation volume:

$$E_c = L \cdot W \cdot E_{fw} \tag{A.49}$$

where $L$ [m] is the length of the canal, $W$ [m] is the width of the transfer canal, and $E_{fw}$ is the free-water evaporation volume [m s$^{-1}$], which can be calculated through various free-water evaporation models (Dingman, 2002). If $W$ is not available for the canal, it can be estimated as:

$$W = \begin{cases} \tau \cdot Q_{T_a}^{\phi} \text{ if } (\tau \cdot Q_{T_a}^{\phi}) \geq 0.01 \\ 0 \qquad\quad \text{if } (\tau \cdot Q_{T_a}^{\phi}) < 0.01 \end{cases} \tag{A.50}$$

where $\tau$ is an empirical constant (default value 8.0).

## A13   Land use/land cover types

When provided with a map of land use/land cover (LULC) types, WBMplus calculates each water balance component - canopy interception (optional, based on leaf area index input), soil moisture, potential evapotranspiration, actual evapotranspiration, and runoff - for each LULC type individually. Equations for these components (Equations A.8 –A.14) are modified by providing LULC-specific maximum leaf area index and available water capacity (or root depth). WBMplus has three modes for calculating the LULC-specific water balance components: 1) sub-grid cell processes, 2) average grid cell processes, or 3) dominant LULC type processes.

In (1), sub-grid cell processes, WBMplus calculates each water balance component seperately for each LULC type. Soil moisture values for each sub-grid cell type is tracked at every time step. Grid cell values for runoff and evapotranspiration are output based on the sum of the sub-grid cell components.

For (2), average grid cell processes, WBMplus requires the LULC type maps to be pre-processed. The pre-processor takes an area-weighted average across all grid cell subtypes and outputs a single value for the grid cell average available water

capacity (or root depth) and maximum leaf area index. WBMplus uses these average values to calculate soil moisture, evapotranspiration, and runoff.

For (3), dominant LULC type, WBMplus requires the LULC type maps to be pre-processed. The pre-processor identifies the LULC type with the largest area within each grid cell. Only the available water capaicty (or root depth) and maximum leaf area index values for the maximum LULC type are used within the grid cell. This method is most appropriate for high spatial resolution model simulations in which the dominant LULC type typically occupies more than half of a grid cell.

Impervious areas can be input as part of a general LULC map, or input as a separate map. See Section A 9.1 for details on impervious areas. Croplands, both rainfed and irrigated, can also be input as independent maps. Croplands require additional input datasets; they are described below.

## A13.1   Cropland

Cropland can be irrigated or rainfed. All cropland data inputs are pre-processed to fit WBMplus formatting requirements. The pre-processor makes a daily gridded time series of available water capacity and crop coefficient ($k_c$) for each crop type (both rainfed and irrigated). For all crops, Equations A.10 and A.11 for soil moisture change are modified such that $g(W_s) = 1$. Equation A.14 for actual evapotranspiration is modified such that $PET$ becomes a crop-specific $PET_c$:

$$PET_c = k_c \cdot PET \tag{A.51}$$

Crop coefficients $k_c$ are time-varying, crop-specific parameters. The use of crop coefficients follows the methodology of Allen (1998), in which $k_c$ modifies the reference evapotranspiration ($PET$) over the course of the growing season so that crops require more water as they grow larger, and less water at the end of the growing season once they've reach full size. Each crop has four growth stages: initial, developing, middle, and late. The default crop $k_c$ values for WBMplus are based on Siebert and Döll (2010), which provides crop coefficients and rooting depth (used to calculate available water capcity) for 26 different crop types. Three $k_c$ values are provided by Siebert and Döll (2010) for each crop: one value each for the initial, mid, and end-of-late growth stages. The length of all four crop growth stages for each crop are also provided by Siebert and Döll (2010). The WBMplus crop pre-processor assumes there is a linear increase in $k_c$ from the initial to the middle growth stage, and a linear decrease in $k_c$ from the middle to the end-of-late growth stage. See Figure A3 for examples of time-varying $k_c$ values for typical perennials, vegetables, rice and other types of crops.

## A13.1.1   Irrigated cropland

WBMplus calculates the irrigation water requirements of all irrigated croplands based on crop water requirements and soil moisture content. When the soil moisture
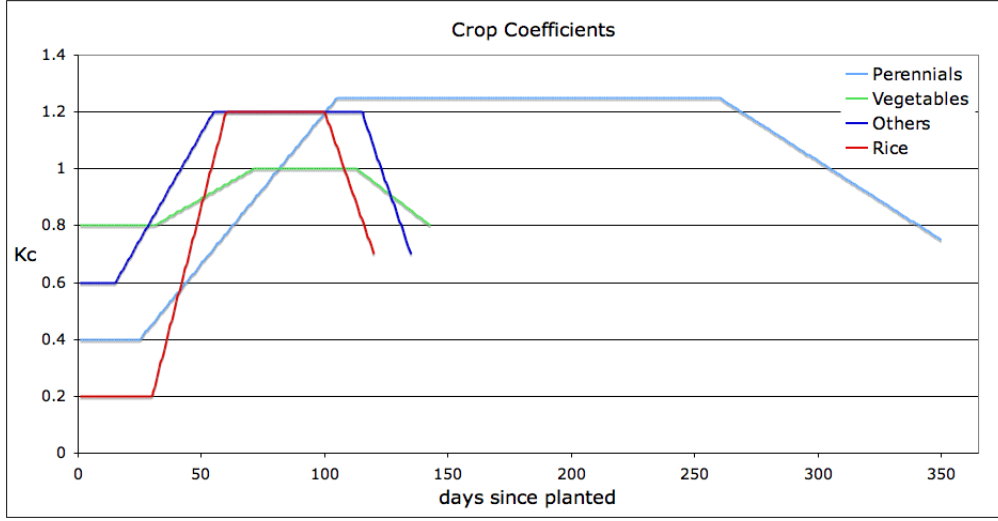
Figure A3: **Crop coefficient, $k_c$, for four example crop categories, plotted against the number of days since the crop was planted. Note that each $k_c$ line here is plotted beginning at 0 days, but the different crop categories' planting dates are not necessarily the same calendar date.**

content of an irrigated cropland area goes below a crop-specific threshold, water is applied to the soil to increase soil moisture up to the available water capacity. The crop-dependent soil moisture threshold, $C_{t_c}$ for each crop $c$ is determined by the field capacity, $F_{cap}$, crop-specific rooting depth, $R_{d_c}$, wilting point, $W_{pt}$, and a crop-specific scalar, $s_c$. The WBMplus crop pre-processor calculates available water capacity, $W_{cap_c}$ for each crop $c$ in each grid cell:

$$W_{cap_c} = (F_{cap} - W_{pt})R_{d_c} \tag{A.52}$$

The crop-specific soil moisture threshold is:

$$C_{t_c} = s_c \cdot W_{cap_c} \tag{A.53}$$

The crop-specific scalar $s_c$ represents a crop's inability to completely remove all water from the soil.

Once soil moisture, $W_{cap_c}$, is below $C_{t_c}$, the amount of irrigation water required to bring the soil back up to available water capacity is calulated:

$$I_{net_c} = W_{cap_c} - W_{s_c} \tag{A.54}$$

Irrigation water extraction, conveyence, and application systems are not perfectly efficient. Rather, water can evaporate from canals and from irrigation sprinklers, it can percolate to groundwater storage due to leaky canals and pipes, and it can become surface runoff and percolation to groundwater storage when more water

is applied to irrigated areas than is required. $I_{net}$ must be modified in each grid cell by an irrigation efficiency scalar, $E_{eff}$, to determine the amount of water that must be withdrawn from water storage pools in order to deliver enough water the irrigated cropland area. The amount of water that must be withdrawn is the gross irrigation amount, $I_{gross}$:

$$I_{gross_c} = \frac{I_{net_c}}{E_{eff}} \tag{A.55}$$

The volume of water $I_{net}$ is applied either to soil moisture or to rice paddies (see Section A 13.1.2 below for a description of rice paddy irrigation water requirements). The remainder, $I_{gross} - I_{net}$ becomes non-beneficial evaporation $E_{nb}$, surface runoff, $R_{ro}$, and percolation, $R_{perc}$:

$$E_{nb} = min(I_a(PET - AET), (I_{gross} - I_{net})) \tag{A.56}$$

$$R_{ro} = r(I_{gross} - I_{net} - E_{nb}) \tag{A.57}$$

$$R_{perc} = (1 - r)(I_{gross} - I_{net} - E_{nb}) \tag{A.58}$$

where $I_a$ is the irrigated cropland area fraction, $PET$ is the potential evapotranspiration of the entire grid cell, $AET$ is the actual evapotranspiration of the entire grid cell, and $r$ is a parameter that splits the return flows between runoff and percolation (default value 0.5).

Like all LULC types, irrigated cropland can be represented as individual sub-grid cell processes, as an average grid-cell, or as the dominant irrigated crop type within each grid cell. When irrigated croplands are represented as individual sub-grid cell processes, grid-cell values of both $I_{net}$ and $I_{gross}$ are the sum of all the individual $I_{net_c}$ and $I_{gross_c}$ values.

Irrigation water can be extracted from five different water storage pools and moved to the soil pool. The amount of water extracted is a function of the soil moisture deficit and irrigation efficiency These five pools are: (1) small reservoirs, (2) groundwater storage pool, (3) rivers, (4) large reservoirs and (5) unsustainable groundwater. Water storage pools (1) and (4) are only used if the associated reservoir database inputs are used in the model simulation. The default order in which water is extracted from these pools is (1) through (5), in numerical order. This order can be re-set. Alternatively, the target ratio, $R_w$ ($R_w \epsilon [0, 1]$), of water extracted from rivers and reservoirs compared to groundwater storage can be input to WBMplus as a map layer. Additionally, water can be extracted to meet the needs of a single grid cell not only from that grid cell, but also from the largest river and reservoir storage of any grid cell within a defined search distance. In the case of a target surface-to-groundwater extraction ratio, water extractions to meet the irrigation water demand of a grid cell are:

$$I_{rr_0} = min(R_w \cdot R_{max}(I_{gross} - I_{sr}), R_{stor}) \tag{A.59}$$

161

where $I_{rr_0}$ is water extracted from river and large reservoir storage within the search distance, $R_w$ is the target ratio, $R_{max}$ is the maximum allowed fraction of all river and reservoir storage that can be extracted (default value 0.8), $I_{sr}$ is irrigation water withdrawan from small reservoir storage, and $R_{stor}$ is the total water stored in rivers and large reservoirs within the grid cell search distance.

$$I_{g_0} = min((1 - R_w)(I_{gross} - I_{sr}), W_g) \tag{A.60}$$

where $I_{g_0}$ is water extracted from groundwater storage, and $W_g$ is the water in groundwater storage in the grid cell.

If $(I_{gross} - I_{sr} - I_{rr_0} - I_{g_0}) > 0$, then additional water is taken from groundwater storage:

$$I_{g_1} = min((I_{gross} - I_{sr} - I_{rr_0} - I_{g_0}), W_g) \tag{A.61}$$

where $I_{g_1}$ is additional water extracted from groundwater storage, regardless of the target ratio. If irrigation water demand $I_{gross}$ is still unfulfilled, then additional water is taken from river and large reservoir storage, regardless of the target ratio:

$$I_{rr_1} = min((I_{gross} - I_{sr} - I_{rr_0} - I_{g_0} - I_{g_1}), R_{max} \cdot R_w) \tag{A.62}$$

where $I_{rr_1}$ is additional water extracted from river and large reservoir storage, regardless of the target ratio.

The total amount of water extracted from rivers and reservoirs for irrigation in a grid cell is:

$$I_{rr} = I_{rr_0} + I_{rr_1} \tag{A.63}$$

The total amount of water extracted from groundwater storage for irrigation in a grid cell is:

$$I_g = I_{g_0} + I_{g_1} \tag{A.64}$$

If water extracted from small reservoirs, rivers and large reservoirs, and groundwater storage still falls short of $I_{gross}$, then additional water is extracted from unsustainable groundwater, $I_{ugw}$, to make up the difference:

$$I_{ugw} = U_l(I_{gross} - I_{sr} - I_{rr} - I_g) \tag{A.65}$$

where $U_l$ ($U_l \epsilon [0, 1]$) is a coefficient that limits the amount of unsustainable groundwater that can be extracted (default value 1).

162

## A13.1.2 Irrigated rice paddies

In addition to water for crop evapotranspiration, irrigated rice paddies also require irrigation water to flood the paddy area. The WBMplus crop pre-processor generates a daily time series of rice paddy flood water requirements. These requirements are: (1) flood water added to the paddy on the day that rice is planted (default value 50 mm), and (2) water added every day of the paddy rice growing season to account for percolation. Percolation occurs at a rate determined by the grid cell's soil drainage class. The WBMplus crop pre-processor assumes the following percolation rates based on input soil drainage classes: WBM estimates percolation rates as: 8 mm day$^{-1}$ for extremely well drained soils, 5 mm day$^{-1}$ for well-, moderately- and imperfectly-drained soils, and 2 mm day$^{-1}$ for poorly and very poorly drained soils. If no soil drainage class input data are provided, WBMplus assumes a global percolation rate of 3 mm day$^{-1}$.

## A13.1.3 Crop rotations

Over the course of a year, the fractional portion of a grid cell occupied by each crop type can change. This change occurs when one crop season ends and the next begins. The amount of water stored within the soil moisture pool is a function of the root depth of the LULC type. Therefore, when the crop areas change and a different set of crops are present, the root depth and soil moisture capacity simulated by WBMplus also change. To accomodate multiple cropping seasons, WBMplus moves water between the soil moisture pool and the groundwater storage pool in order to maintain a constant soil moisture fraction through the crop-switching process:

$$W_{s_1} = W_{s_0} + \delta W_{s_1} \tag{A.66}$$

$$W_{g_1} = W_{g_0} - \delta W_{s_1} \tag{A.67}$$

where:

$$\delta W_s = (R_{d_1} - R_{d_0}) \frac{W_s}{R_{d_0}(F_{cap} - W_{pt})} \tag{A.68}$$

where $R_{d_1}$ is the rooting depth of the $R_{d_0}$ is the rooting depth of the previous crop.

# A14 Tracking water components

WBMplus can track water from a specific source through flows and stocks within the model. Stocks include river storage, large and small reservoir storage, groundwater storage, and soil moisture. Flows include runoff, percolation, river discharge, water discharge from large large reservoirs, water removed from any stock for irrigation,

inter-basin transfers, and evporation. For any water component $w$ in water storage stock $ST$ at time $t$:

$$ST_{w,t} = \frac{(ST_{w,t-1} \cdot ST_{t-1}) + \sum_{i=1,j=1}^{m,n}[(I_{i,w} \cdot I_i) - (O_{j,w} \cdot_O j)]}{ST_t} \tag{A.69}$$

where $ST_{w,t}$ is the water component $w$ [volume] in stock $ST$, $I_i$ are inflows ($i = 1...n$) to and $O_j$ are outflows ($j = 1...m$) from stock $ST$ at time $t$. For each flow (in or out), the water component fraction at time $t$ is:

$$I_w = \frac{w}{I} \tag{A.70}$$

$$O_w = \frac{w}{O} \tag{A.71}$$

All stocks and flows are considered well-mixed, so that the flows out of a stock have the same fractional water source components as the stock itself. WBMplus current has the capability to track the following water component categories:

- Water sources: rain water, snowmelt, glacier melt, and unsustainable groundwater

- Runoff sources: glacier melt, snowmelt, surface runoff, baseflow

- Irrigation return flows: water that has been extracted for irrigation, water that has not been extracted for irrigation

- Land area sources: a map of source areas identifies water to be categorized onto components based on all runoff orriginating from the source regions.

## A15   Water temperature

The temperature of water in the river is a function of the temperature and volume of surface runoff, baseflow, and previous water temperature, as well as the weather conditions air temperature and cloud coverage, and canopy shading. The temperature model is described in Stewart et al (2013).

First, the volume-weighted average of the previous water in the river segment and the incoming runoff is calculated:

$$T_{w,0} = T_{w,t-1} + R_t \delta t T_{R_t} \tag{A.72}$$

where $T_{w,0}$ is the temperature of the water (not yet corrected for equilibration with air or solar radiation; see Eq A.75 below), $R_t \delta t$ is surface runoff volume, and $T_{R_t}$ is the temperature of the surface runoff water:

$$T_{R_t} = \begin{cases} \frac{T_{rAv}(R_s + R_{ro}) + \beta W_g T_b}{R_t} & \text{if } \frac{M_s}{P_a} < 0.1 \\ \frac{\beta W_g T_b}{R_t} & \text{if } \frac{M_s}{P_a} \geq 0.1 \end{cases} \tag{A.73}$$

where $R_s$ is surface runoff, $R_{ro}$ is irrigation surface runoff, $\beta W_g$ is baseflow, $R_t$ is total runoff, $M_s$ is snowmelt, and $P_a$ is the balance of precipitation, snowmelt, and canopy interception (see Eq. A.13), $T_{rAv}$ is the 5-day running average air temperature and $T_b$ is the temperature of baseflow:

$$T_b = T_{W_g} + \omega(T_{bAv} - T_{W_g}) \tag{A.74}$$

where $T_{W_g}$ is the groundwater temperature (input time series), $\omega$ is a groundwater temperature scaling factor (default value 0.59), and $T_{bAv}$ is the 15-day running average air temperature.

Water within the river segment is then adjusted for equilibration with water previously in the river, energy exchange with the air, and incoming solar radiation:

$$T_{w,t} = T_e + (T_{w,0} - T_e) \cdot e^{(-E_c*\cdot L)/(\tilde{n} C_w h V_w)} \tag{A.75}$$

where $T_{w,t}$ is the temperature of river water at time $t$ (minum value allowed is 0°C), $L$ is the length of the stream segment within the grid cell [m], $\tilde{n}$ is the density of water [kg m$^{-3}$], $C_w$ is the specific heat of water [KJ kg$^{-1}$°C$^{-1}$], $h$ is the water depth [m], $V_w$ is the stream velocity [m d$^{-1}$], and $T_e$ is the in-stream equilibrium temperature including the wet bulb correction:

$$T_e = \frac{273 \cdot E_s}{17.27 - E_s} \tag{A.76}$$

where:

$$E_s = log(H_r \cdot 10^{7.5 S_e/(237.3+S_e)}) \tag{A.77}$$

where $H_r$ is relative humidity, and

$$S_e = T + \frac{E_R - E_0}{E_c} \tag{A.78}$$

where $T$ is the mean daily air temperature, $E_R$ is the net incoming solar radiation [KJ m$^{-2}$d$^{-1}$], $E_o$ is the heat loss rate [KJ m$^{-2}$d$^{-1}$], and $E_c$ is the energy exchange coefficient [KJ m$^{-2}$d$^{-1}$ °C$^{-1}$]:

$$E_0 = \begin{cases} 105 + 23 V_a & \text{if } CS_{fr} < 0.95 \\ -73 + 9.1 V_a & \text{if } CS_{fr} \geq 0.95 \end{cases} \tag{A.79}$$

$$E_c = \begin{cases} 35 + 4.2 V_a & \text{if } CS_{fr} < 0.95 \\ 37 + 4.6 V_a & \text{if } CS_{fr} \geq 0.95 \end{cases} \tag{A.80}$$

where $V_a$ is wind speed and $CS_{fr}$ is the cloud cover fraction, considering canopy shading:

$$CS_{fr} = C_{fr} + (1 - C_{fr})\frac{LAI}{max(LAI)}\frac{H_C}{w} \tag{A.81}$$

where $C_{fr}$ is the cloud cover fraction, $LAI$ is the leaf area index, $H_C$ is the canopy height [m], and $w$ is the stream width [m].

## A16 Notation
### In order of appearance:

| | |
|---|---|
| $PET$ | potential evapotranspiration [mm day$^{-1}$] |
| $PET_H$ | Hamon-method PET [mm day$^{-1}$] |
| $T$ | mean daily air temperature [°C] |
| $\Lambda$ | day length [rad], $\pi = 12$ hrs |
| $\rho_{sat}$ | saturated vapor density [g m$^{-3}$] |
| $P_{sat}$ | saturated vapor pressure [kg m$^{-1}s^{-2}$] |
| $P$ | total precipitation (snow + rainfall) [mm day$^{-1}$] |
| $P_s$ | snow fall [mm day$^{-1}$] |
| $S_p$ | snow pack [mm] |
| $M_S$ | snow melt [mm day$^{-1}$] |
| $T_s$ | snowfall threshold temperature [°C] |
| $T_m$ | snowmelt threshold temperature [°C] |
| $I_c$ | canopy interception [mm day$^{-1}$] |
| $s$ | canopy storage capacity [mm] |
| $LAI$ | leaf area index [-] |
| $W_s$ | soil moisture [mm] |
| $P_r$ | precipitation as rain [mm day$^{-1}$] |
| $\alpha$ | soil drying function constant [-] |
| $W_{cap}$ | available water capacity [mm] |
| $F_{cap}$ | field capacity [-] |
| $W_{pt}$ | wilting point [-] |
| $R_d$ | rooting depth [mm] |
| $P_a$ | balance of precipitation, snowmelt, and canopy interception [mm day$^{-1}$] |
| $AET$ | actual evapotransiration [mm day$^{-1}$] |
| $S$ | surplus water to become runoff [mm day$^{-1}$] |
| $\gamma$ | groundwater storage infiltration constant [-] |
| $R_s$ | runoff from surface [mm day$^{-1}$] |
| $W_g$ | groundwater storage [mm] |
| $\beta$ | baseflow constant [-] |
| $R_t$ | total runoff [mm day$^{-1}$] |
| $R_{imp}$ | runoff from impervious surfaces [mm day$^{-1}$] |
| $C_{imp}$ | impervious surface scalar [-] |
| $A_{imp}$ | impervious surface area [m$^2$] |
| $Q_j^t$ | instantaneous discharge at time t and grid j [m$^3$ s$^{-1}$] |
| $C_0, C_1, C_2$ | coefficients for flow routing [-] |
| $C_{0o}, C_{1o}, C_{2o}$ | coefficients for open water flow routing [-] |
| $C$ | Courant number [m$^2$] |
| $D$ | Reynolds number [m$^{-2}$] |
| $U_w$ | speed of wave propogation, i.e. wave celerity [m$^3$ s$^{-1}$] |
| $V_m$ | mean fluid velocity [m s$^{-1}$] |

| | |
|---|---|
| $L$ | river length across grid cell [m] |
| $A_c$ | grid cell area [m$^2$] |
| $Y_m$ | mean annual channel depth [m] |
| $W_m$ | mean annual channel width [m] |
| $S_0$ | riverbed slope [m km$^{-1}$] |
| $\sigma$ | shape coefficient [-] |
| $Q_m$ | mean annual discharge [m$^3$ s$^{-1}$] |
| $\eta, \nu, \tau, \phi$ | channel shape parameters [-] |
| $N$ | flow direction paramter [-] |
| $V_f$ | flow velocity [m h$^{-1}$] |
| $s_d$ | conversion factor of 3600 [s d$^{-1}$] |
| $S_c$ | grid cell size in degrees [°] |
| $D_t$ | large reservoir discharge at time $t$ [m$^3$ s$^{-1}$] |
| $\bar{Q}$ | 5-year average discharge [m$^3$ s$^{-1}$] |
| $D_{min}$ | minimum allowed reservoir release [-] |
| $S_l$ | reservoir storage level [-] |
| $S_opt$ | optimal reservoir storage level [-] |
| $a, b$ | reservoir constants [-] |
| $A_{SR}$ | small reservoir area [m$^2$] |
| $C_{SR}$ | small reservoir capacity [m$^3$] |
| $d_{SR}$ | small reservoir depth [m] |
| $E_{SR}$ | evaporation from small reservoirs [mm] |
| $Q_T$ | discharge through inter-basin transfer [m$^3$ s$^{-1}$] |
| $Q_d$ | donor river discharge [m$^3$ s$^{-1}$] |
| $Q_{min}$ | minimum allowed donor river flow [m$^3$ s$^{-1}$] |
| $Q_{max}$ | maximum allowed donor river flow [m$^3$ s$^{-1}$] |
| $F$ | fraction parameter for inter-basin transfer flows [-] |
| $Q_{Ta}$ | adjusted $Q_T$ for evaporation [m$^3$ s$^{-1}$] |
| $Q_a$ | inter-basin transfer arrival water volume [m$^3$ s$^{-1}$] |
| $E_c$ | evaporation volume from canal [m$^3$ d$^{-1}$] |
| $L$ | canal length [m] |
| $W$ | canal width [m] |
| $E_{fw}$ | free-water evaporation from canals [m s$^{-1}$] |
| $PET_c$ | crop-specific potential evapotranspiration [mm] |
| $k_c$ | crop coefficient [-] |
| $s_c$ | crop-specific scaler [-] |
| $C_{t_c}$ | crop-specific soil moisture threshold [-] |
| $I_{net}$ | net irrigation [mm] |
| $I_{gross}$ | gross irrigation [mm] |
| $E_{eff}$ | irrigation efficiency [-] |
| $E_{nb}$ | non-benficial evaporation from irrigation water [mm] |
| $I_a$ | irrigated cropland area fraction [-] |
| $R_{ro}$ | irrigation return flow as surface runoff [mm] |

$R_{perc}$    percolation to groundwater storage from irrigation water [mm]

$I_{rr}$    irrigation water withdrawn from rivers and reservoirs [mm]

$R_w$    target extraction ratio between surface water and ground water [-]

$R_{max}$    maximum allowed river water withdrawal [-]

$I_{ugw}$    irrigation water withdrawn from unsustainable groundwater [mm]

$U_l$    unsustainable groundwater coefficient [-]

$ST_w$    water component w in water storage stock ST [-]

$I_w$    inflow of component w [-]

$O_w$    outflow of component w [-]

$T_{w,0}$    water temperature, not corrected for equilibration with air or solar radiation [°C]

$T_{R_t}$    temperature of total runoff [°C]

$T_{rAv}$    5-day average air temperature [°C]

$T_b$    baseflow temperature [°C]

$T_{W_g}$    groundwater temperature [°C]

$\omega$    groundwater temperature scaling factor [-]

$T_{bAv}$    15-day running average air temperature

$T_w$    water temperature [°C]

$\tilde{n}$    density of water [kg m$^{-3}$]

$C_w$    specific heat of water [KJ/kg/C] [KJ m$^{-2}$ kg$^{-1}$ °C$^{-1}$]

$h$    water depth [m]

$V_w$    stream velocity [m d$^{-1}$]

$T_e$    in-stream equilibrium temperature, including wet bulb correction [°C]

$E_s$    wet bulb correction factor [°C]

$H_r$    relative humidity [-]

$S_e$    in-stream equilibrium temperature [°C]

$E_c$    energy exchange coefficient [KJ m$^{-2}$ d$^{-1}$ °C$^{-1}$]

$V_a$    wind speed [m$^{-2}$ d$^{-1|}$]

$CS_{fr}$    cloud cover fraction corrected for canopy shading [-]

$C_{fr}$    cloud cover fraction [-]

$w$    stream width [m]

$I_{sr}$    irrigation water withdrawn from small reservoirs [mm]

$I_g$    irrigation water withdrawn from groundwater storage [mm]

This appendix contains a comparison of Water Balance Model simulated river discharge to available discharge data for each of the three dissertation chapters. Model fit is assessed using the Nash-Sutcliffe efficiency (*NSE*) metric (Nash and Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum(Obs - Sim)^2}{\sum(Obs - Obs_{mean})^2} \qquad \text{(B.1)}$$

where *Obs* is the observation data, *Sim* is the model simulated data, and *Obs_{mean}* is the mean of the observation data. The *NSE* compares model simulations against the mean of the observation data. If the *NSE* < 0, then the observed mean is a better predictor than the model; if the *NSE* > 0, then the model is a better predictor than the observed mean; an *NSE* = 1 indicates a perfect fit between the model and observation data.

For all analysis shown below, both *Obs* and *Sim* are monthly mean river discharge [m$^3$ s$^{-1}$]. Observation and simulation values are only compared for months in which both values are present; where the observation data time series has gaps, the gaps are left out of the analysis.

## B1. River discharge simulations in Chapter 1: China

Observation data are from the Global Runoff Data Center (GRDC), http://www.bafg.de/GRDC/EN/Home/homepage_node.html, and were downloaded in April of 2012. The GRDC compiles river discharge time series from gauge stations that have at least 20 years of data and capture a basin area > 10,000 $km^2$. Model simulations are described in Chapter 1; WBM uses the MERRA climate data product, and simulates the years 1981-2000. At the end of this section (B1), all monthly hydrographs are shown, comparing model and observation river discharge.
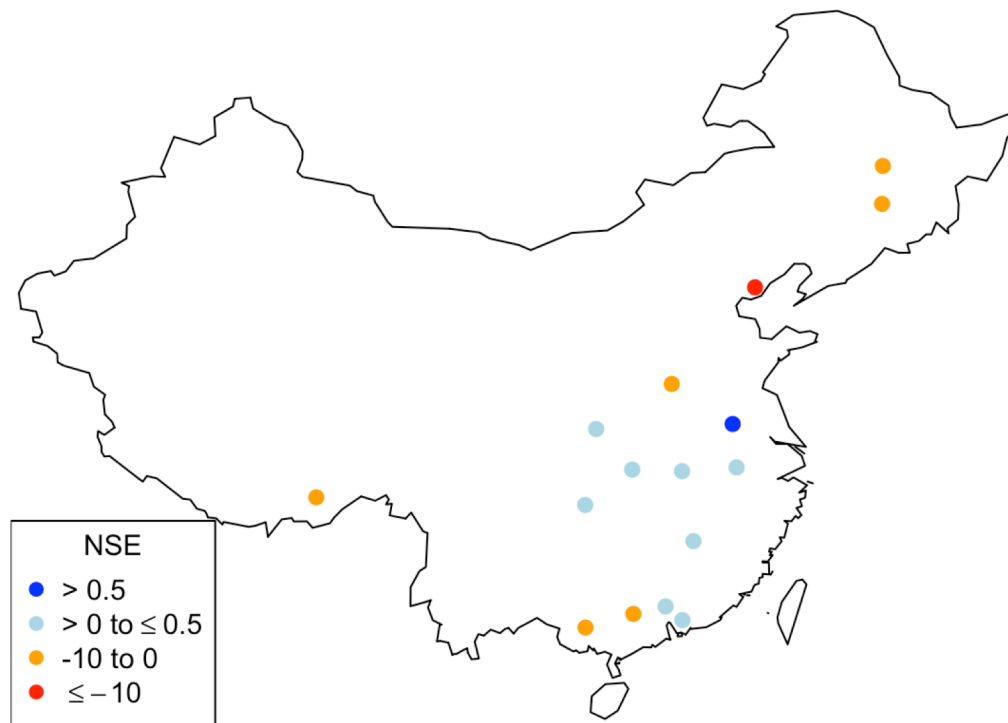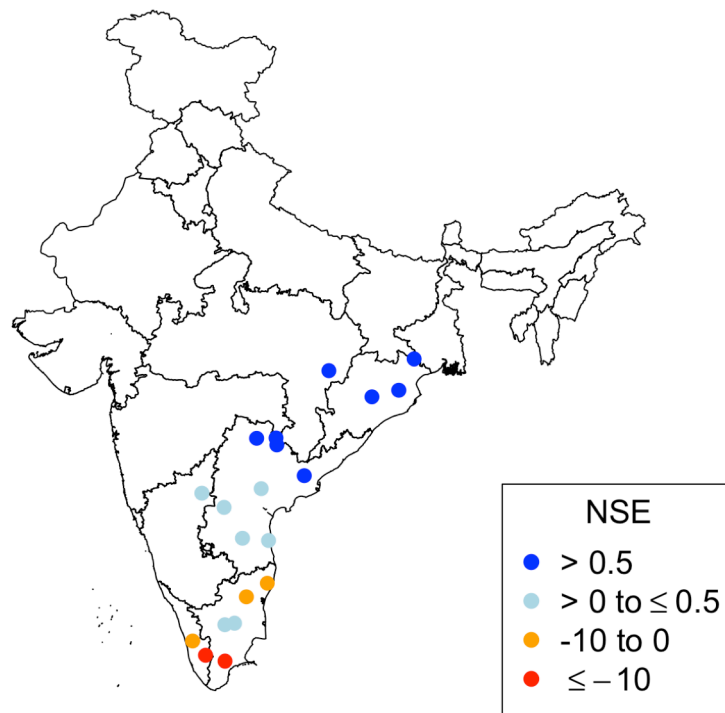


**Fig B 1.1** Nash-Sutcliffe efficiency values for modeled versus WBM-simulated river discharge for the 16 GRDC river gauges that have data in the modeled time period.
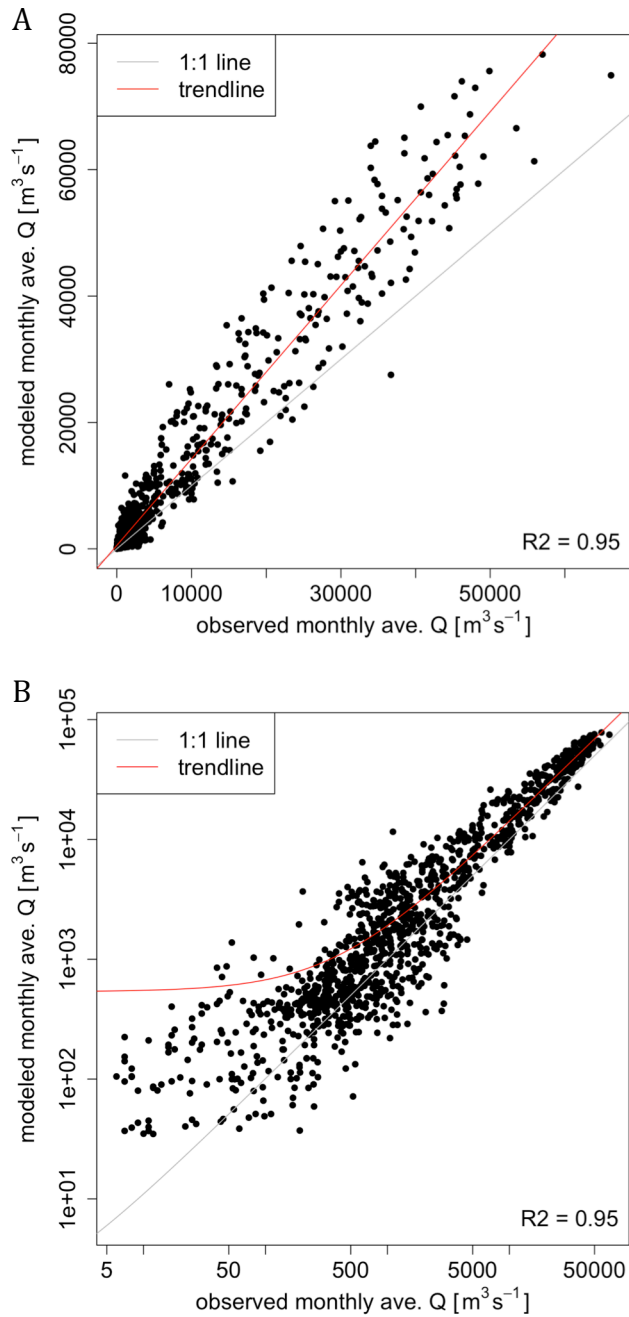
**Fig B 1.2** Modeled versus observed monthly average discharge from all 16 GRDC river gauge stations in China (A) on a linear scale, and (B) on a log-log plot. Each data point is one month's value; only months with both data and observation are shown. Comparison of the model trend line (red) to the 1:1 line (grey) shows that the model typically over-estimates monthly average discharge. The $R^2$ value is 0.95.
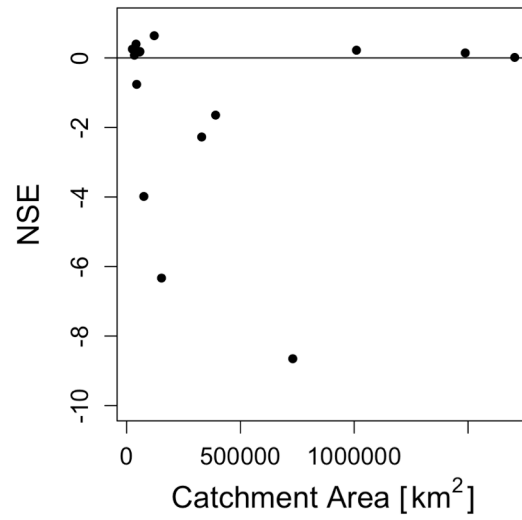
**Fig B 2.3** Nash-Sutcliffe efficiency (NSE) plotted against catchment area for each of the 16 GRDC river gauge stations in China.

**Table B 1.1** Summary of Nash-Sutcliffe efficiency (NSE) values for each of the 16 GRDC gauge stations with available data in China during the model simul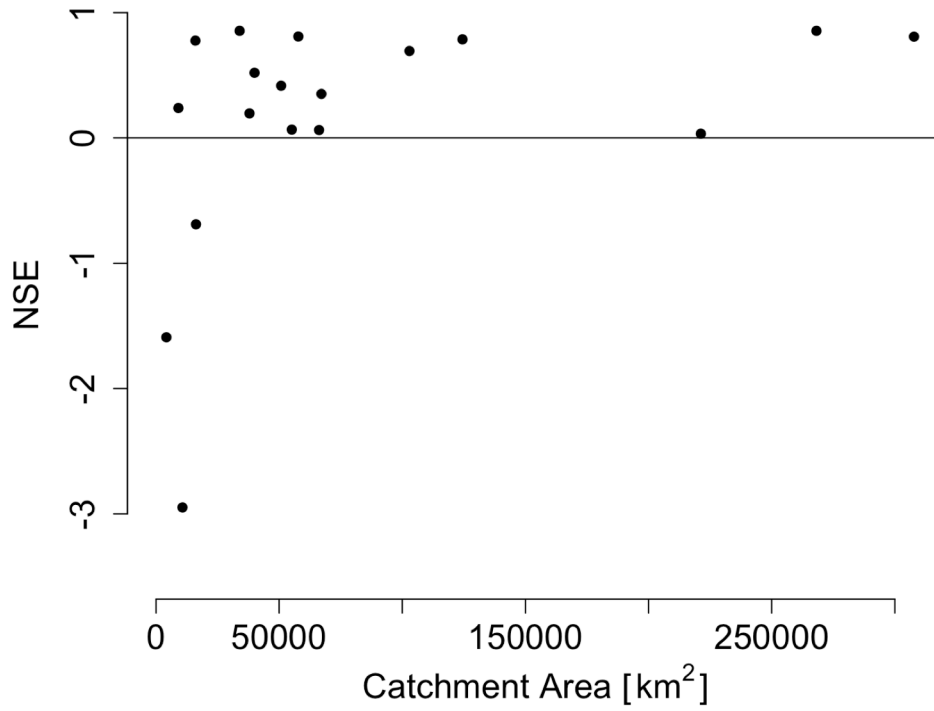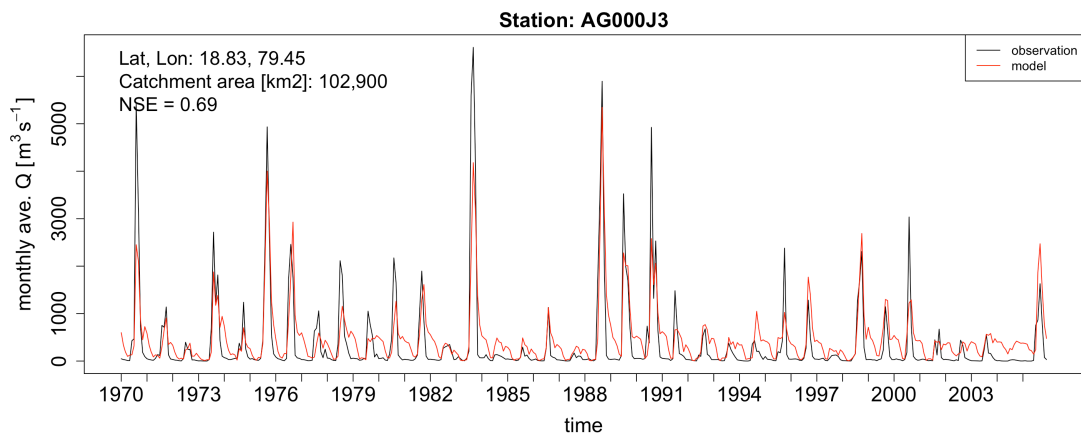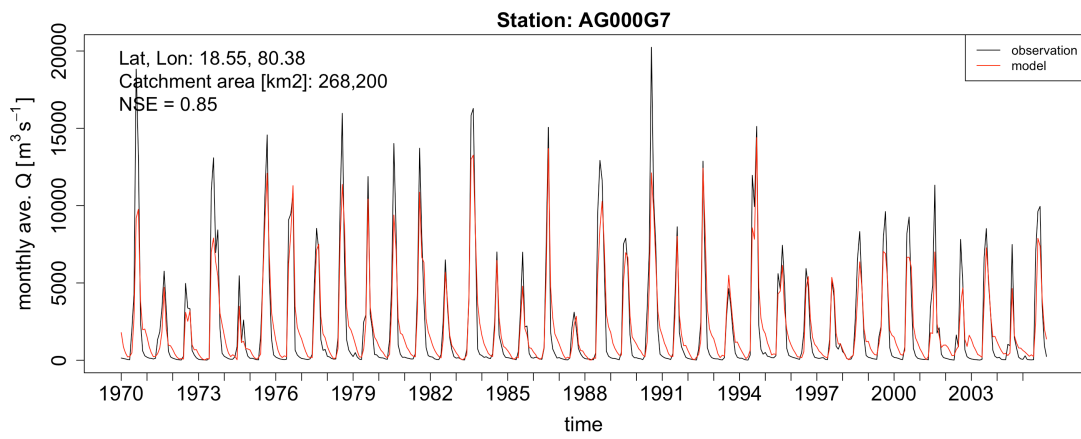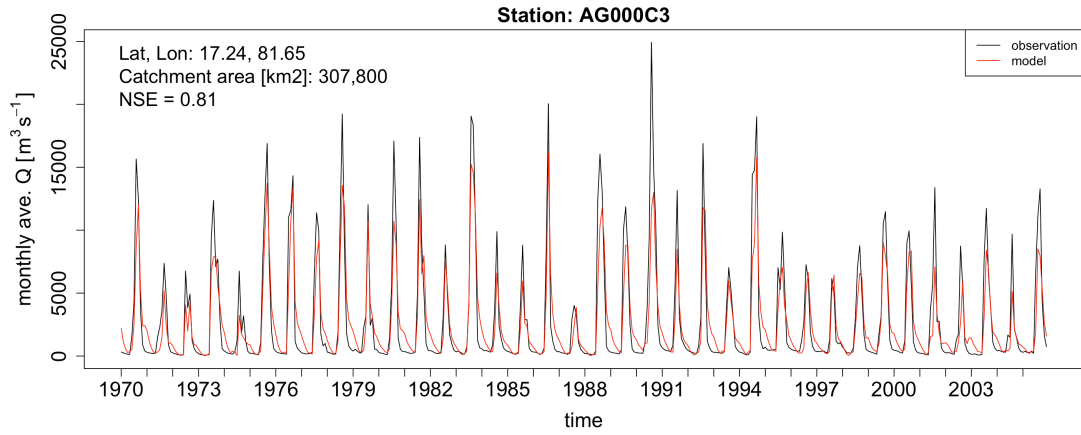ation period. The **% Data Cov.** shows the % of non-missing monthly data points within the given date range (**Start Date** to **End Date**).  For gauges with **% Data Cov.** < 100%, there are data gaps within the river discharge time series.  Model results which coincide with gaps in the river discharge time series are not used in the NSE calculation.

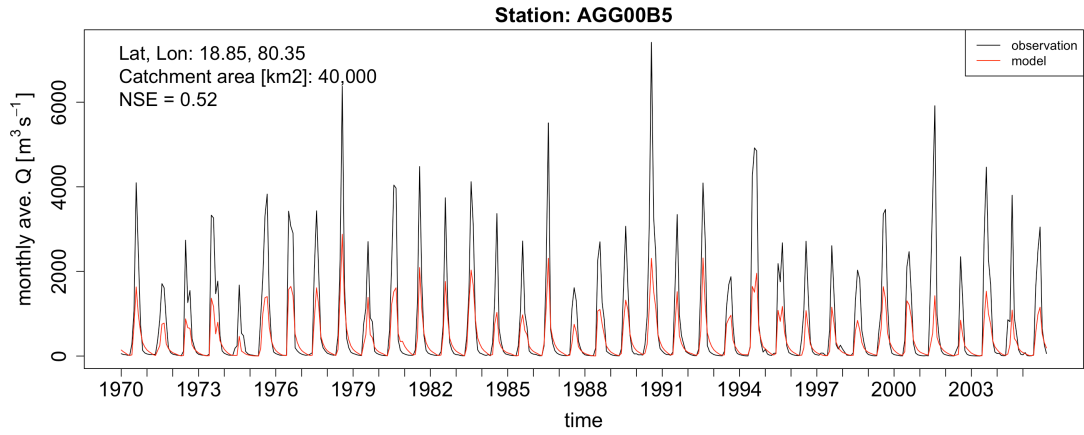| River | GRDC Station | Lat | Lon | Catchment Area [km2] | start date | end date | % data Cov. | NSE |
|---|---|---|---|---|---|---|---|---|
| YANGTZE | DATONG | 30.77 | 117.62 | 1,705,383 | 1981 | 2000 | 40 | 0.02 |
| YANGTZE | HANKOU | 30.58 | 114.28 | 1,488,036 | 1981 | 1986 | 100 | 0.15 |
| YANGTZE | YICHANG | 30.66 | 111.23 | 1,010,000 | 1981 | 1986 | 100 | 0.22 |
| HUANG | HUAYUAN KOU | 34.92 | 113.65 | 730,036 | 1981 | 2000 | 40 | -8.65 |
| SONGHUA | HAERBIN | 45.77 | 126.58 | 391,000 | 1981 | 1987 | 100 | -1.64 |
| XI | WUZHOU | 23.48 | 111.3 | 329,705 | 1981 | 2000 | 30 | -2.27 |
| BRAHMA-PUTRA | YANGCUN | 29.28 | 91.88 | 153,191 | 1981 | 1982 | 100 | -6.33 |
| HUAI | BENGBU | 32.93 | 117.38 | 121,330 | 1981 | 2000 | 45 | 0.64 |
| YU | NANNING | 22.8 | 108.37 | 75,500 | 1981 | 1984 | 100 | -3.98 |
| WU | GONGTAN | 28.9 | 108.35 | 58,300 | 1981 | 1982 | 100 | 0.19 |
| GAN | JIAN | 27.1 | 114.98 | 56,200 | 1981 | 1984 | 100 | 0.17 |
| SONGHUA | JILIN | 43.88 | 126.53 | 44,100 | 1981 | 1986 | 100 | -0.76 |
| LUAN | LUAN | 39.73 | 118.75 | 44,100 | 1981 | 2000 | 40 | -25.08 |
| HAN | ANKANG | 32.68 | 109.02 | 41,400 | 1981 | 1986 | 100 | 0.40 |
| BEI | HENGSHI | 23.85 | 113.27 | 34,013 | 1981 | 1987 | 100 | 0.08 |
| DONG | BOLUO | 23.17 | 114.3 | 25,325 | 1981 | 2000 | 35 | 0.26 |

River SONGHUA, Station GRDC_2106500, climate merra

Lat, Lon: 45.77, 126.6
Catchment area [km2]: 391,000
NSE = -1.6



River SONGHUA, Station GRDC_2106600, climate merra

Lat, Lon: 43.88, 126.5
Catchment area [km2]: 44,100
NSE = -0.76



River BRAHMAPUTRA, Station GRDC_2151100, climate merra

Lat, Lon: 29.28, 91.88
Catchment area [km2]: 153,191
NSE = -6.3

River LUAN, Station GRDC_2178500, climate merra

Lat, Lon: 39.73, 118.8
Catchment area [km2]: 44,100
NSE = -25

River HUANG, Station GRDC_2180800, climate merra

Lat, Lon: 34.92, 113.6
Catchment area [km2]: 730,036
NSE = -8.7

River WU, Station GRDC_2181400, climate merra

Lat, Lon: 28.9, 108.4
Catchment area [km2]: 58,300
NSE = 0.19

River YANGTZE, Station GRDC_2181600, climate merra

Lat, Lon: 30.66, 111.2
Catchment area [km2]: 1,010,000
NSE = 0.22



River YANGTZE, Station GRDC_2181800, climate merra

Lat, Lon: 30.58, 114.3
Catchment area [km2]: 1,488,036
NSE = 0.15



River GAN, Station GRDC_2181850, climate merra

Lat, Lon: 27.1 , 115
Catchment area [km2]: 56,200
NSE = 0.17

176

River YANGTZE, Station GRDC_2181900, climate merra

Lat, Lon: 30.77, 117.6
Catchment area [km2]: 1,705,383
NSE = 0.017



River HUAI, Station GRDC_2181950, climate merra

Lat, Lon: 32.93, 117.4
Catchment area [km2]: 121,330
NSE = 0.64



River HAN, Station GRDC_2182100, climate merra

Lat, Lon: 32.68, 109
Catchment area [km2]: 41,400
NSE = 0.4

**River YU, Station GRDC_2186500, climate merra**

Lat, Lon: 22.8, 108.4
Catchment area [km2]: 75,500
NSE = -4



**River XI, Station GRDC_2186800, climate merra**

Lat, Lon: 23.48, 111.3
Catchment area [km2]: 329,705
NSE = -2.3



**River BEI, Station GRDC_2186900, climate merra**

Lat, Lon: 23.85, 113.3
Catchment area [km2]: 34,013
NSE = 0.08

**River DONG, Station GRDC_2186950, climate merra**

Lat, Lon: 23.17, 114.3
Catchment area [km2]: 25,325
NSE = 0.26

## B2. River discharge simulations in Chapter 2: India

Observation data for river discharge in India is from the Water Resources Information System of India (India-WRIS). This data was downloaded from: http://www.india-wris.nrsc.gov.in in August of 2015, and includes monthly mean river discharge from 20 different gauge stations. Only observation data that overlaps with the model historical period (1970-2005) are used for the comparisons shown below. At the end of this section (B2), all monthly hydrographs are shown, comparing model and observation river discharge.



**Fig B 2.1** Nash-Sutcliffe efficiency values for modeled versus WBM-simulated river discharge for the 20 river gauges with available data in India.
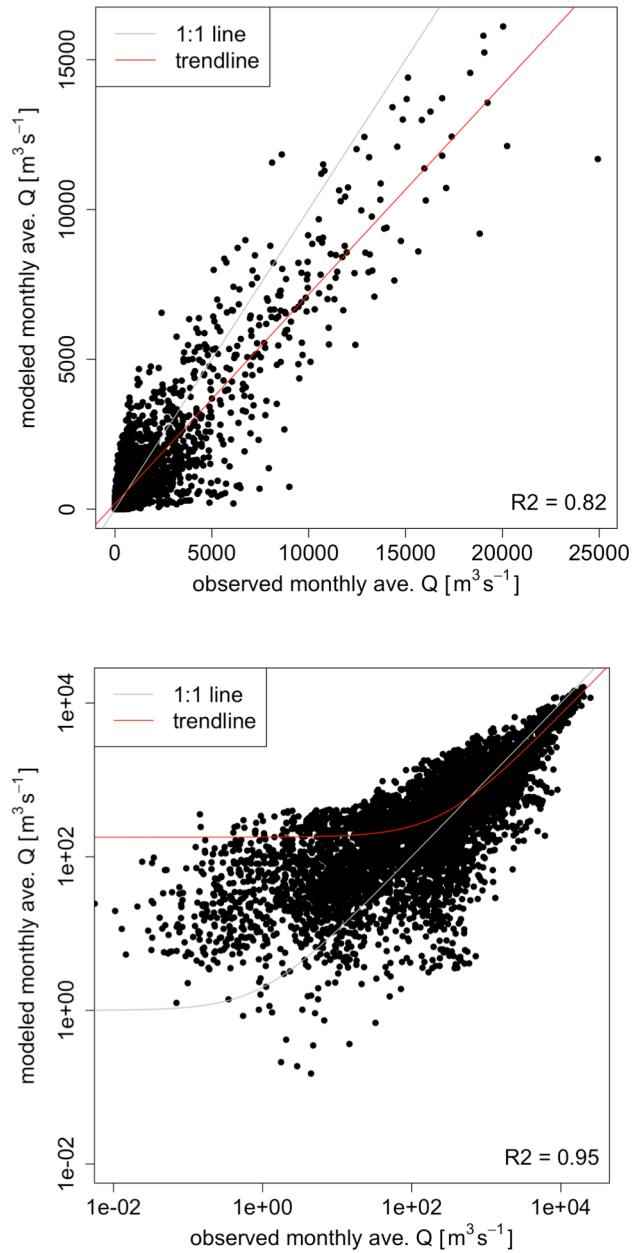
**Fig B 2.2** Modeled versus observed monthly average discharge from all 20 river gauge stations on (A) linear axes, and (B) on a log-log plot. Each data point is one month's value; only months with both data and observation are shown. Comparison of the model trend line (red) to the 1:1 line (grey) shows that the model typically under-estimates monthly average discharge. The $R^2$ value is 0.82.

**Fig B 2.3** Nash-Sutcliffe efficiency (NSE) plotted against catchment area for each of the 18 river gauge stations NSE > -3. The two stations with NSE < -3 have catchment areas of 712 km$^2$ (NSE = -63) and 3,721 km$^2$ (NSE = -18).

**Table B 2.1** Summary of Nash-Sutcliffe efficiency (NSE) values for each of the 20 gauge stations with available data through the IWRIS. The **% Data Coverage** shows the % of non-missing monthly data points within the given date range (**Start Date** to **End Date**). For gauges with **% Data Coverage** < 100%, there are data gaps within the river discharge time series. Model results which coincide with gaps in the river discharge time series are not used in the NSE calculation.

| IWRIS Gauge Name | Latitude | Longitude | Catchment Area [km²] | Start Date | End Date | % Data coverage | NSE |
|---|---|---|---|---|---|---|---|
| AG000C3 | 17.2425 | 81.6486 | 307,800 | 1970 Apr | 2005 Sep | 100 | 0.81 |
| AG000G7 | 18.5478 | 80.3844 | 268,200 | 1970 Apr | 2005 Sep | 100 | 0.85 |
| AK000E7 | 16.6844 | 79.6589 | 221,220 | 1970 Apr | 2005 Sep | 100 | 0.03 |
| EM000G5 | 20.6017 | 84.7758 | 124,450 | 1970 Apr | 2005 Sep | 100 | 0.79 |
| AG00OJ3 | 18.8333 | 79.4483 | 102,900 | 1975 Apr | 2005 Sep | 97 | 0.69 |
| AKL00B8 | 15.8842 | 77.9475 | 67,180 | 1976 Apr | 2005 Sep | 100 | 0.35 |
| CC000G4 | 10.9444 | 78.4336 | 66,243 | 1970 Apr | 2005 Sep | 99 | 0.06 |
| EM000R2 | 21.7219 | 82.7894 | 57,780 | 1987 Apr | 2005 Sep | 50 | 0.81 |
| AK000N2 | 16.4906 | 76.92 | 55,150 | 1989 Apr | 2005 Sep | 87 | 0.07 |
| AP000A6 | 14.4708 | 79.9994 | 50,800 | 1971 Apr | 2005 Sep | 95 | 0.42 |
| ACG00B5 | 18.8483 | 80.3478 | 40,000 | 1977 Apr | 2005 Sep | 45 | 0.52 |
| AP000I1 | 14.5678 | 78.8 | 37,981 | 1978 Apr | 2005 Sep | 93 | 0.2 |
| EB000G6 | 20.8847 | 86.0164 | 33,955 | 1977 Apr | 2005 Sep | 51 | 0.85 |
| CQ000C1 | 12.6511 | 79.9425 | 16,230 | 1989 Apr | 2005 Sep | 37 | -0.69 |
| JAMSHOLAG-HAT | 22.2192 | 86.7228 | 16,000 | 1979 Apr | 2005 Sep | 98 | 0.78 |
| CP000H2 | 12.0667 | 78.9775 | 10,780 | 1971 Apr | 2005 Sep | 96 | -2.95 |
| CCD00D5 | 10.8817 | 77.9842 | 9,080 | 1971 Apr | 2005 Sep | 99 | 0.24 |
| KS000F3 | 10.1831 | 76.4975 | 4,234 | 1976 Apr | 2005 Sep | 3 | -1.59 |
| CY000K7 | 9.3236 | 77.9919 | 3,721 | 1971 Apr | 2005 Sep | 51 | -17.96 |
| KS000U1 | 9.5739 | 77.0894 | 712 | 1999 Apr | 2005 Sep | 80 | -63.11 |

Station: AG000C3

Lat, Lon: 17.24, 81.65
Catchment area [km2]: 307,800
NSE = 0.81

Station: AG000G7

Lat, Lon: 18.55, 80.38
Catchment area [km2]: 268,200
NSE = 0.85

Station: AG000J3

Lat, Lon: 18.83, 79.45
Catchment area [km2]: 102,900
NSE = 0.69

184

**Station: AGG00B5**

Lat, Lon: 18.85, 80.35
Catchment area [km2]: 40,000
NSE = 0.52

**Station: AK000E7**

Lat, Lon: 16.68, 79.66
Catchment area [km2]: 221,220
NSE = 0.034

**Station: AK000N2**

Lat, Lon: 16.49, 76.92
Catchment area [km2]: 55,150
NSE = 0.067

185

Station: CC000G4

Lat, Lon: 10.94, 78.43
Catchment area [km2]: 66,243
NSE = 0.063

Station: CCD00D5

Lat, Lon: 10.88, 77.98
Catchment area [km2]: 9,080
NSE = 0.24

Station: CP000H2

Lat, Lon: 12.07, 78.98
Catchment area [km2]: 10,780
NSE = -2.9

**Station: KS000F3**



Lat, Lon: 10.18, 76.5
Catchment area [km2]: 4,234
NSE = -1.6

**Station: KS000U1**



Lat, Lon: 9.574, 77.09
Catchment area [km2]: 712
NSE = -63

## B3. River discharge simulations in Chapter 3: Global

Observation data are from the Global Runoff Data Center (GRDC), http://www.bafg.de/GRDC/EN/Home/homepage_node.html, and were downloaded in April of 2012. The GRDC compiles river discharge time series from gauge stations that have at least 20 years of data and capture a basin area > 10,000 km$^2$.

Chapter 3 uses 4 different climate input data sets for WBM to simulate river discharge. These four datasets are described in Chapter 3, and will be referred to here as ERA, MERRA, NCEP, and UDEL. Only GRDC data that was available for the model simulation time period (1980-2009) were used; in total, 599 GRDC gauge stations were used.
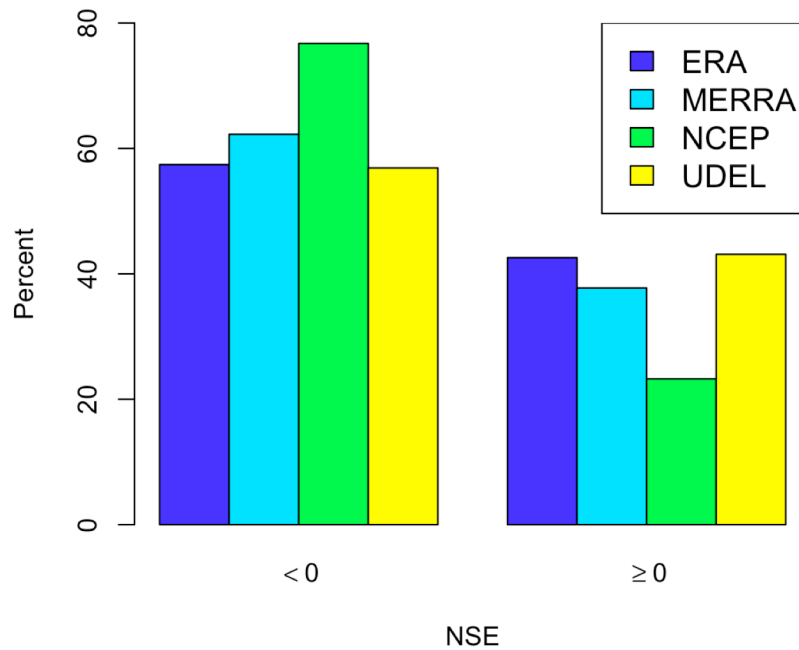


**Fig B 3.1** Percent of Nash-Sutcliffe Efficiency (NSE) values < 0 and ≥ 0 for each of the 4 input climate data sets.
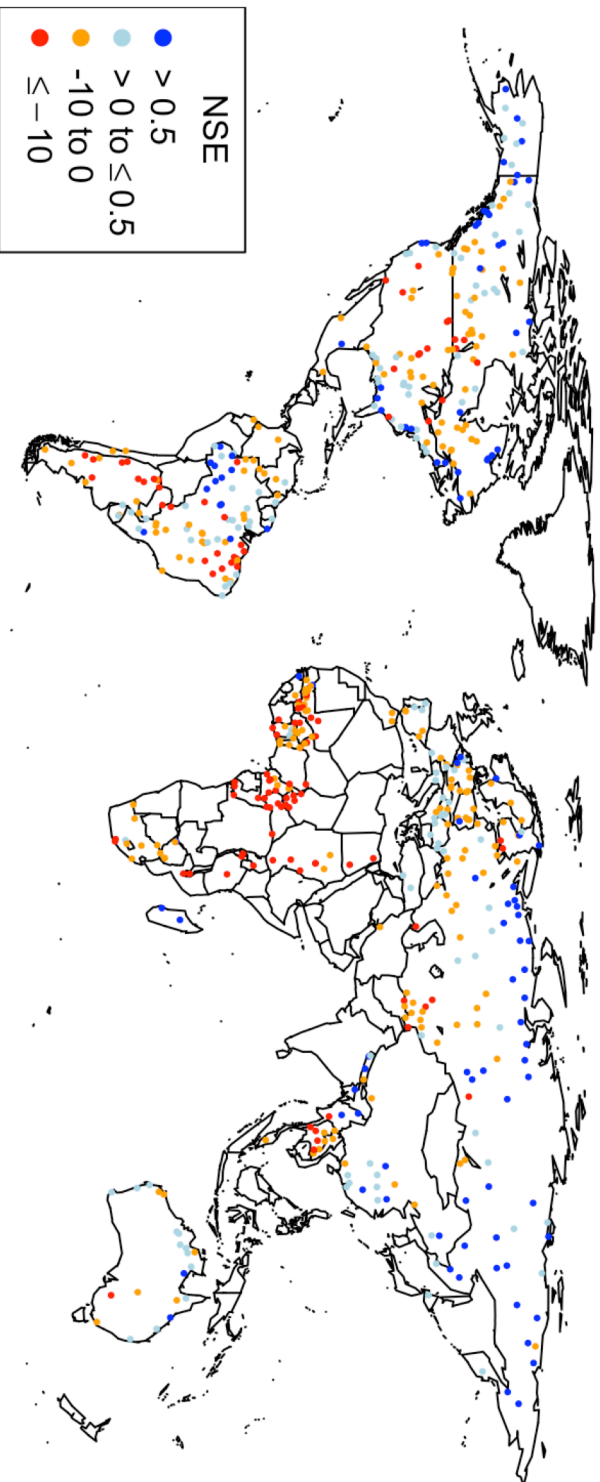
**Fig B 3.2** Nash-Sutcliffe Efficiency (NSE) values for average monthly river discharge of WBM simulations with the ERA climate, compared to 599 GRDC gauge stations.
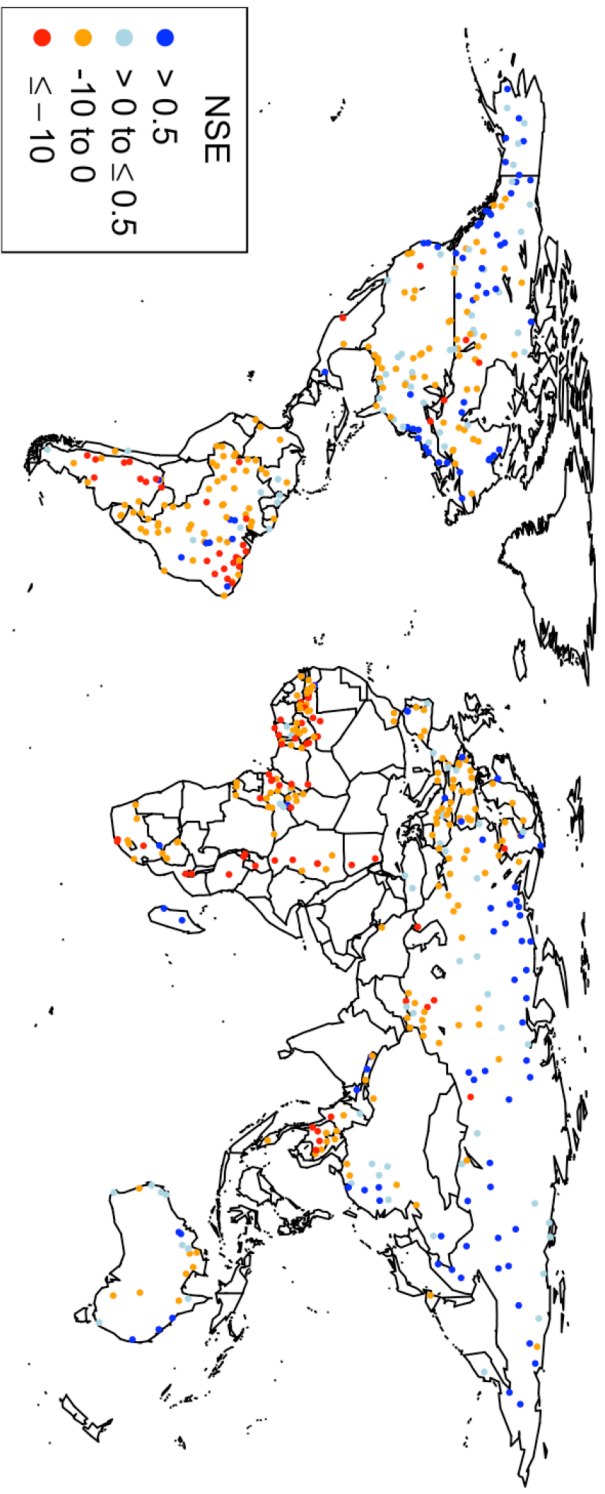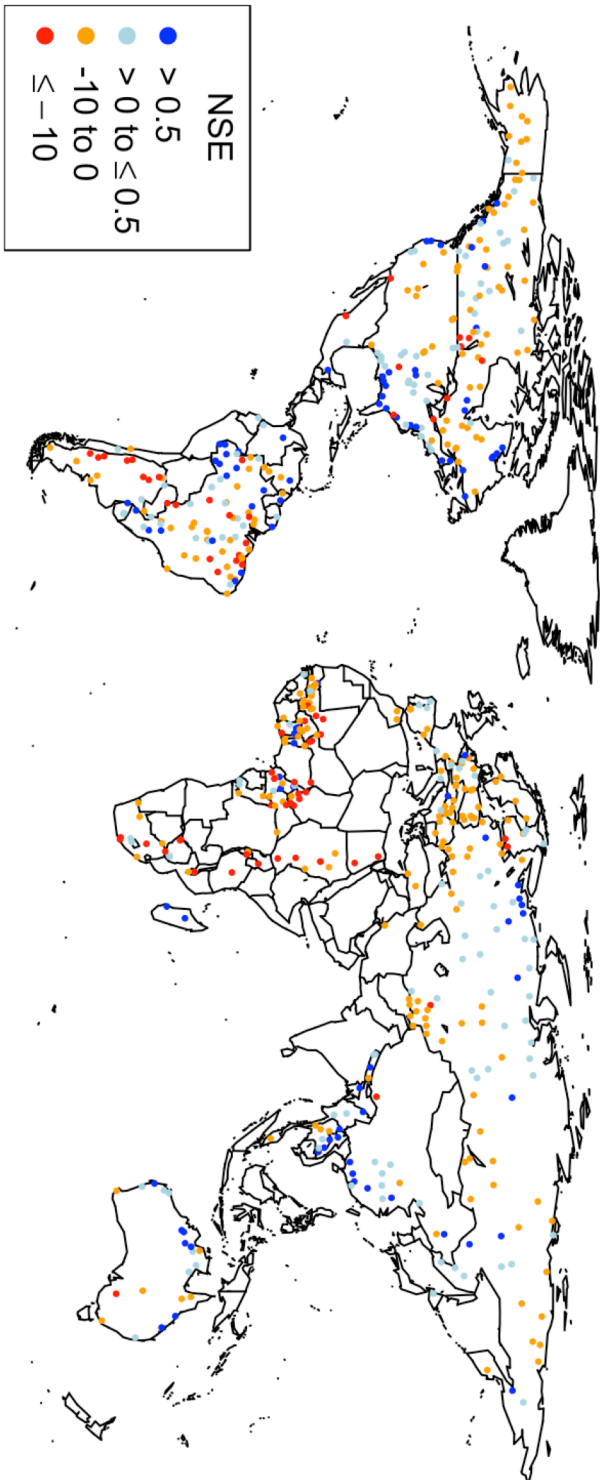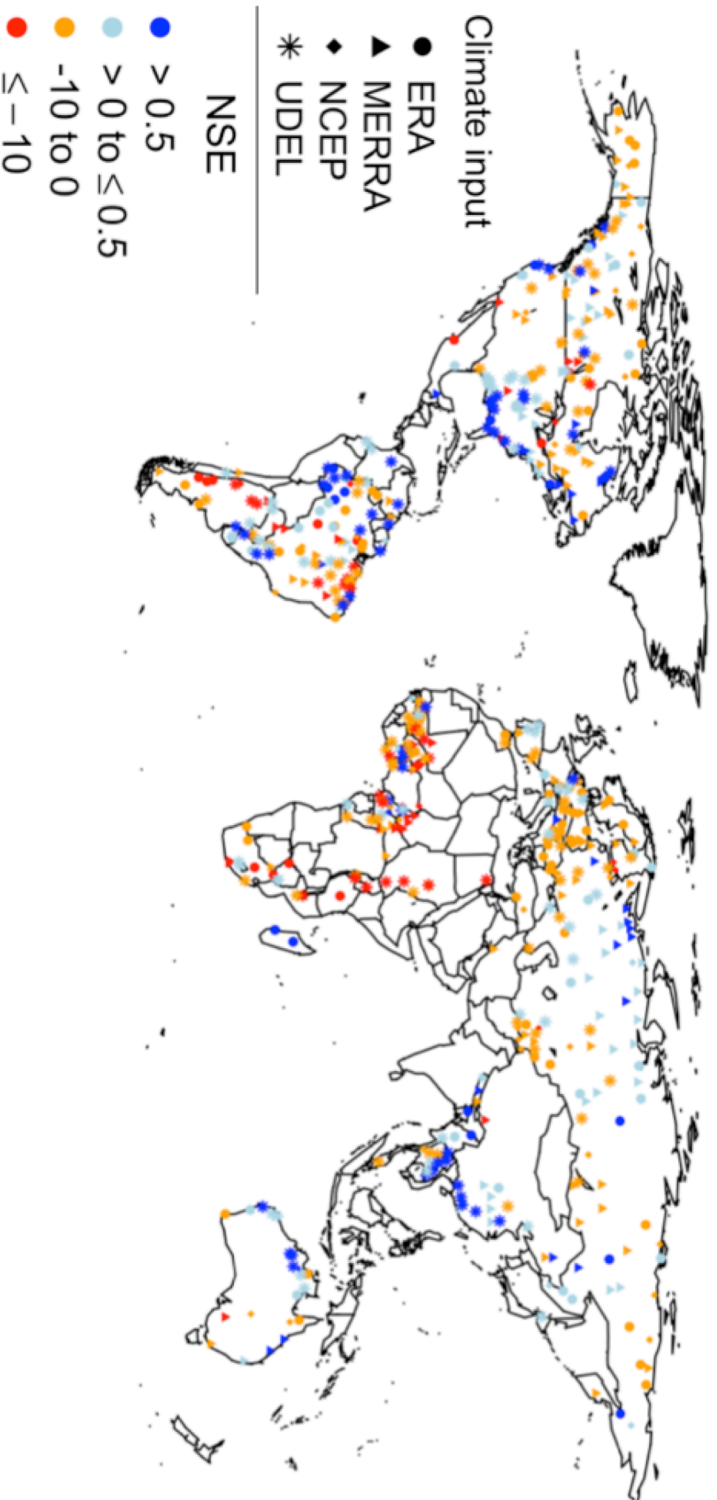
**Fig B 3.3** Nash-Sutcliffe Efficiency (NSE) values for average monthly river discharge of WBM simulations with the MERRA climate, compared to 599 GRDC gauge stations.

193
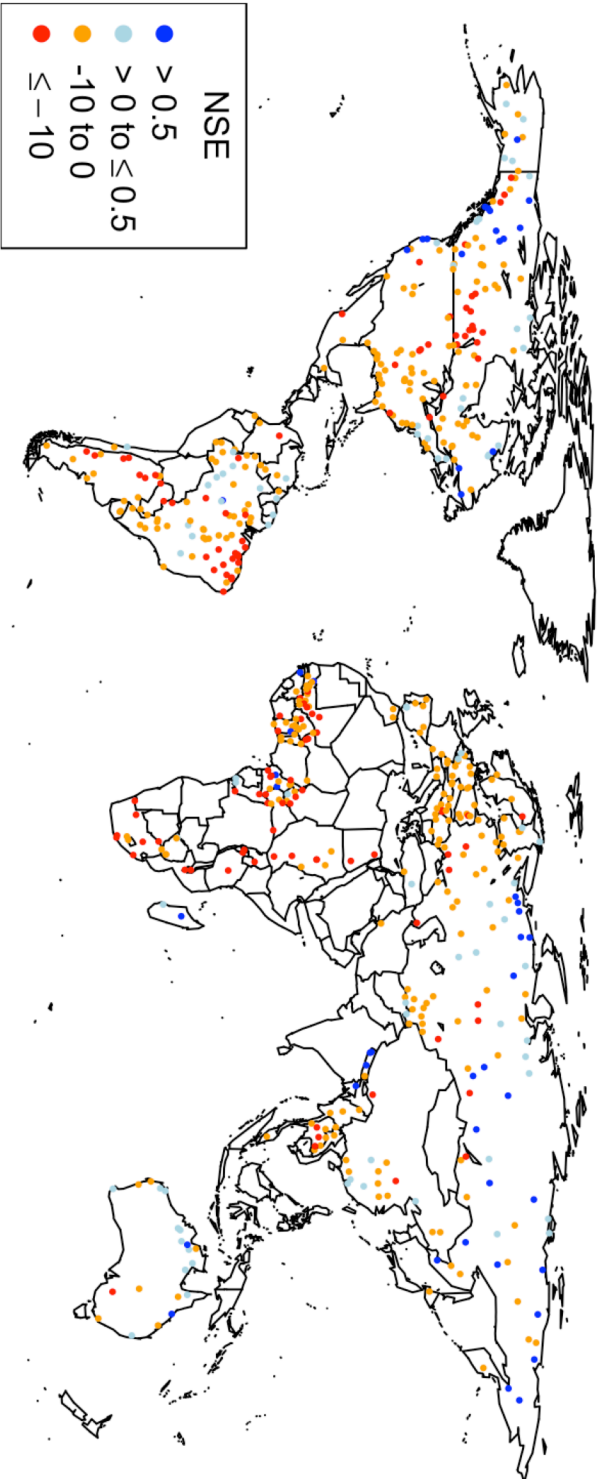
**Fig B 3.4** Nash-Sutcliffe Efficiency (NSE) values for average monthly river discharge of WBM simulations with the NCEP climate, compared to 599 GRDC gauge stations.

194

**Fig B 3.5** Nash-Sutcliffe Efficiency (NSE) values for average monthly river discharge of WBM simulations with the UDEL climate, compared to 599 GRDC gauge stations.

**Fig B 3.6** The maximum Nash-Sutcliffe Efficiency (NSE) values for average monthly river discharge of the four WBM simulations are shown. The color of each map symbol shows the maximum NSE value, and the shape indicates which climate input produced the maximum value. The WBM simulation with ERA climate (circles) produced 121 maximum NSE values; MERRA (triangles) produced 173 maximum NSE values; NCEP (diamonds) produced 50 maximum NSE values; UDEL (asterisk) produced 212 maximum NSE values.
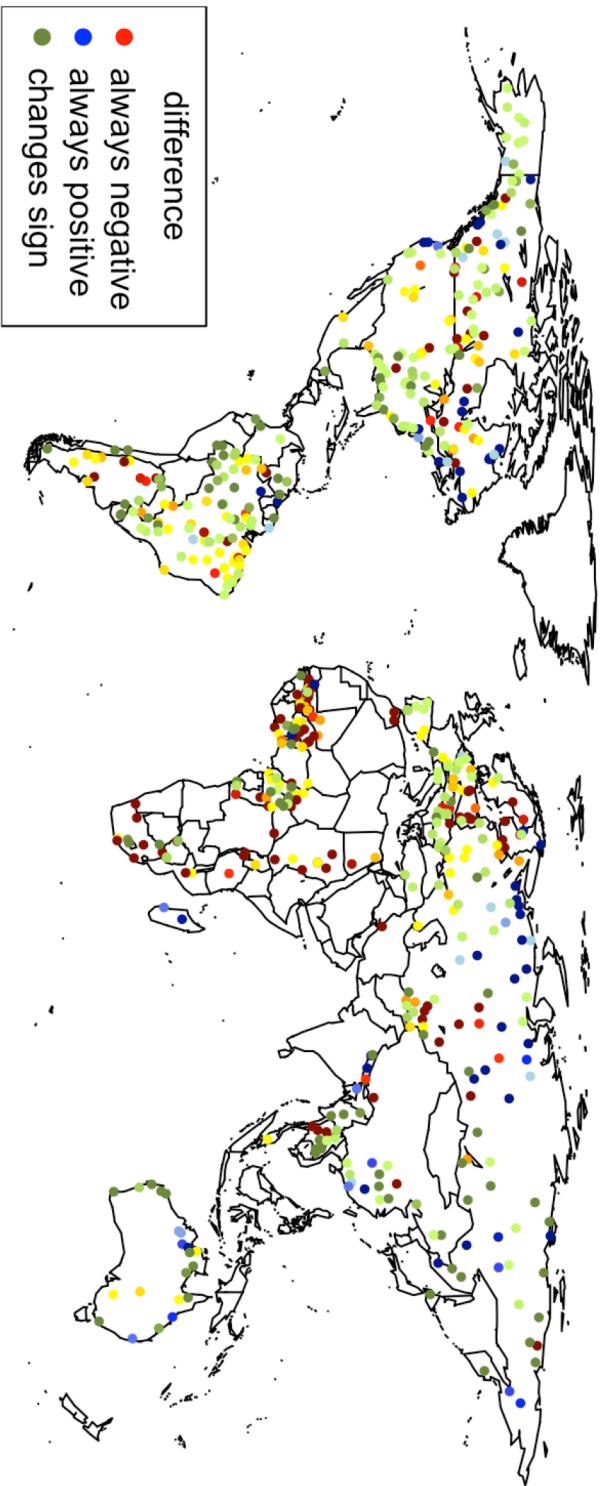
**Fig B 3.7** The maximum difference in Nash-Sutcliffe Efficiency (NSE) values resulting from the four climate inputs. Red points show where the NSE value is negative for all four WBM simulations; blue points show where NSE values are positive for all four simulations; green points show gauges for which the minimum NSE value is negative and the maximum NSE value is positive. Color darkness shows the range of difference values, with darker colors indicating larger differences.
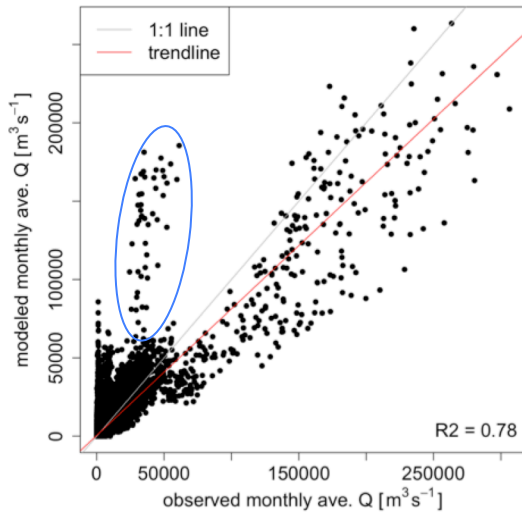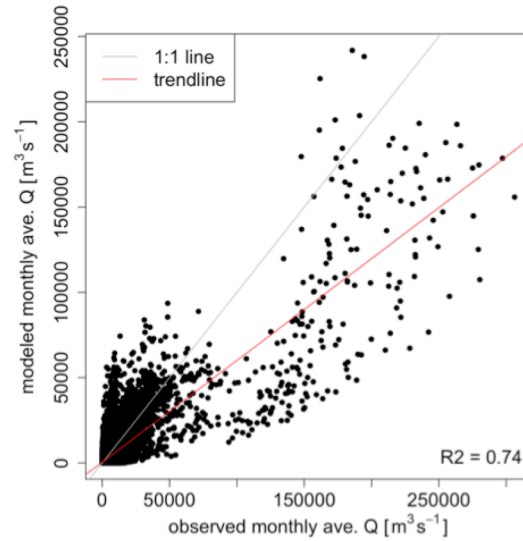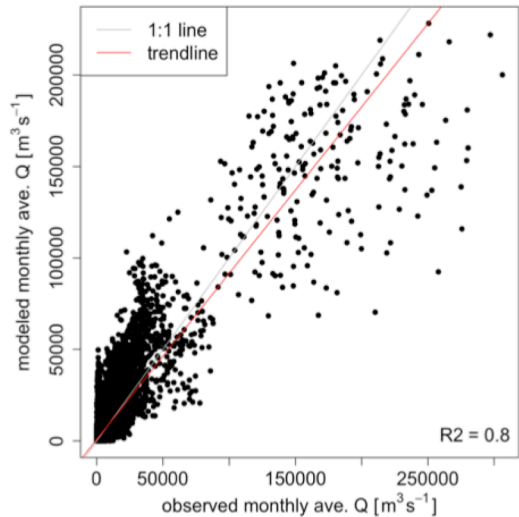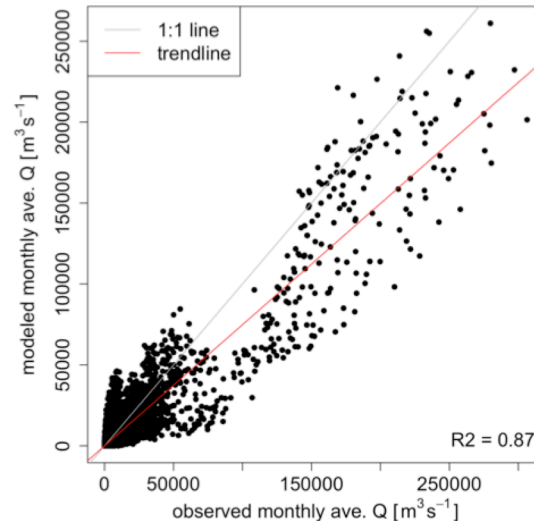
**Fig B 3.8** Modeled versus observed monthly average discharge. Each data point is one month's value; only months with both data and observation are shown. Comparison of the model trend line (red) to the 1:1 line (grey) shows that the model typically under-estimates monthly average discharge. The $R^2$ value for the WBM simulation using (a) ERA climate is 0.78; (b) MERRA climate is 0.74; (c) NCEP climate is 0.80; (d) UDEL climate is 0.87. All data points in the blue oval in (a) are from the Congo River.
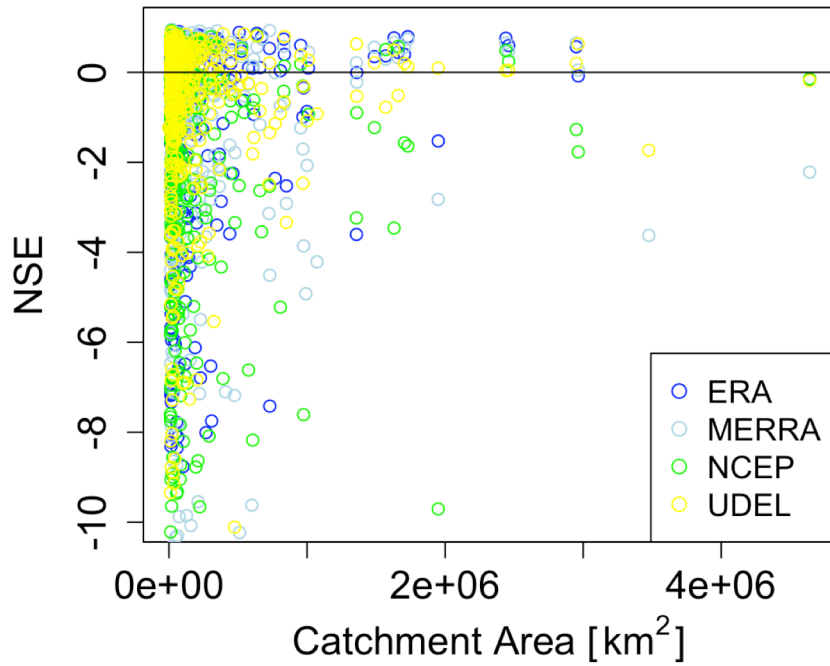
**Fig B 3.8** Nash-Sutcliffe efficiency (NSE) plotted against catchment area for each of the 18 river gauge stations NSE > -10. All catchments with NSE values < -10 have catchment areas < 2,000,000 $km^2$.
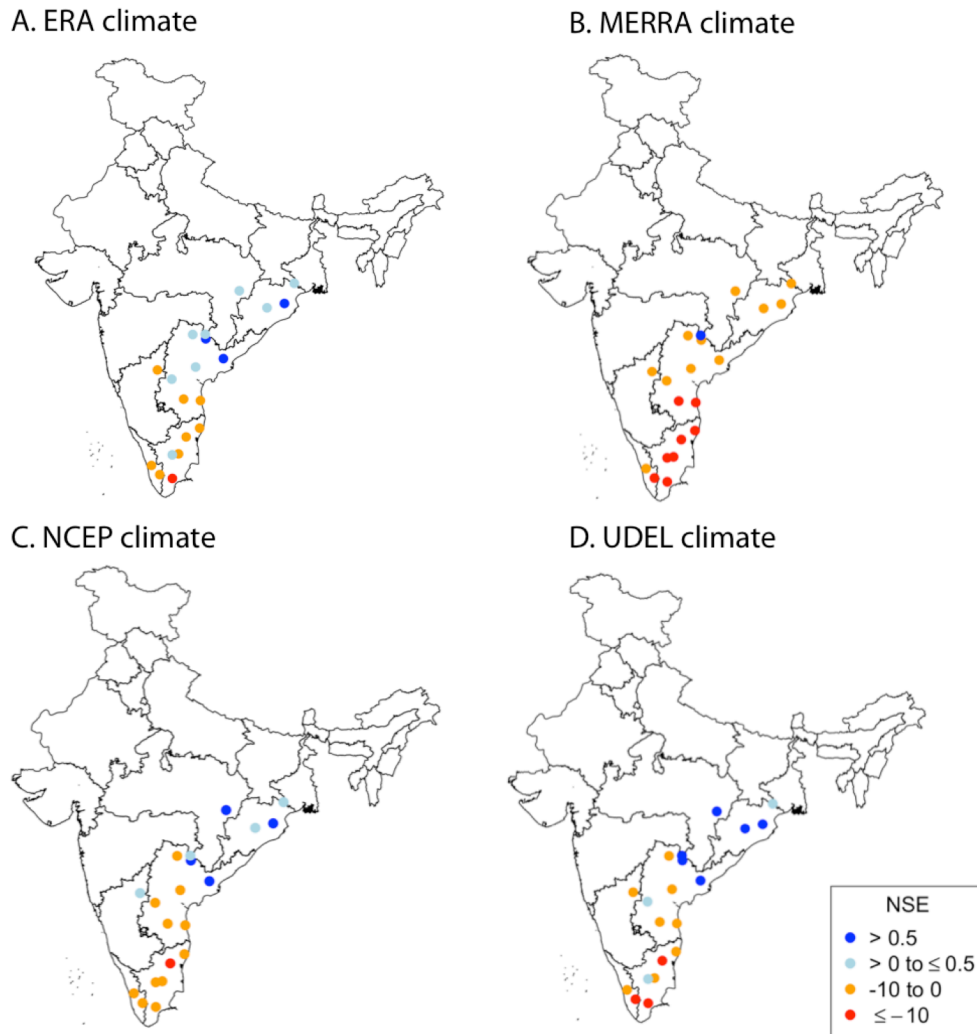
**Fig B 3.9** Nash-Sutcliffe efficiency (NSE) values of each of the global WBM simulations over India, using the IWRIS river discharge data (see Section B2 above) as the observation data set. The IWRIS data for rivers in India is not part of the GRDC; the maps shown here fill in an important gap in the GRDC dataset for the purposes of analyzing the WBM simulations for Chapter 3, as India is an important agricultural producer and groundwater user. None of the four global simulations perform as well over India as the India-specific WBM simulations analyzed in Section B2. The model input data used for Chapter 2 and Section B2 for both climate and agricultural land use are specific to India; the APHRODITE climate product claims to be the best gridded climate product for use over the monsoon regions of Asia, and the agricultural land use maps were produced by the International Crops Research Institute for the Semi-Arid Tropics. WBM-simulated river discharge better matches observed discharge (higher NSE value) when the region-specific input data is used; this result indicates that input data error is one significant source of uncertainty in the model, potentially leading to poor NSE values.

## Bibliography

Aeschbach-Hertig, W. and Gleeson, T. (2012) Regional strategies for the accelerating global problem of groundwater depletion, *Nature Geoscience,* **5**(12) 853-861, doi: 10.1038/NGEO1617.

Adhikari, B., Verhoeven, R. and Iroch, P. (2009) Inter-basin transfer of Nepal's water resources for sustainable benefits *Water and Urban Development Paradigms*, ed J Feyen J, K Shannon and M Neville (London: Taylor & Francis Group) pp 647-53.

Allen, R.G., Pereira, L.S., Raes, D. & Smith, M. (1998) Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements, *FAO Irrigation and Drainage Paper No. 56* (FAO, Rome).

Amarasinghe, U. A., Shah, T. and Malik, R. P. (2009) *Strategic Analyses of the National River Linking Project (NRLP) of India Series 1, India's Water Future: Scenarios and Issues* (Sri Lanka : International Water Management Institute, Colombo) http://publications.iwmi.org/pdf/H042029.pdf

AQUASTAT, FAOs global information system of water and agriculture 2008 http://www.fao.org/nr/water/aquastat/irrigationdrainage/index.stm.

Ashfaq, M., Shi, Y., Tung, W. W., Trapp, R. J., Gao, X., Pal, J. S. and Diffenbaugh, N. S. (2009) Suppression of south Asian summer monsoon precipitation in the 21st century *Geophys. Res. Lett.* **36** L01704.

Auffhammer, M., Ramanathan, V. and Vincent, J. R. (2012) Climate change, the monsoon, and rice yield in India *Clim. Change* **111**(2) 411-24.

Auffhammer ,M., Hsiang, S. M., Schlenker, W. and Sobel, A. (2013) Using weather data and climate model output in economic analyses of climate change *Review of Environmental Economics and Policy* **7**(2) 181-98.

Badiani, R., Jessoe, K. K.  and Plant, S. (2012) Development and the environment: the implications of agricultural electricity subsidies in India *J. Env. Dev.* **21** 244-62.

Banzhaf H S and Lavery N 2010 Can the land tax help curb urban sprawl? Evidence from growth patterns in Pennsylvania *J. Urban Econ.* **67**(2) 169-79.

Bhat, G. S. (2006) The Indian drought of 2002-a sub-seasonal phenomenon? *Q. J. R. Meteorol. Soc.* **132** 2583-2602.

Biemans, H., Speelman, L. H., Ludwig, F., Moors, E. J., Wiltshire, A. J., Kumar, P., Gerten, D. and Kabat, P. (2013) Future water resources for food production in five South Asian river basins and potential for adaptation - A modeling study *Sci. Total Environ.* **468** S117-31.

Brouwer, C. and Heibloem, M. (1986) *Irrigation Water Management Training Manual No. 3: Irrigation Water Needs* (Rome: Food and Agriculture Organization of the United Nations).

Burke, M. and Emerick, K. (2015) Adaptation to climate change: Evidence from US agriculture *Am. Econ. J. Econ. Policy,* in press http://web.stanford.edu/~mburke/papers/Burke_Emerick_2015.pdf

Cai, X., D. Molden, M. Mainuddin, B. Sharma, A. Movin-ud-Din, and P. Karimi (2011) Producing more food with less water in a changing world: assessment of water productivity in 10 major river basins*, Water International,* **36**, 42-62.

Calow, R. C., Howarth, S. E. and Wang, J. (2009) Irrigation Development and Water Rights Reform in China*, Int. J. Water Resour. Dev.,* **25**(2) doi: 10.1080/07900620902868653.

Cao, G., Zheng, C., Scanlon, B.R., Liu, J. and Li, W. (2013) Use of flow modeling to assess sustainability of groundwater resources in the North China Plain, *Water Resources Research*, **49,** 159-175, doi:10.1029/2012WR011899.

Chamberlain, G. (1984) *Handbook of Econometrics* vol 2, ed Z Griliches and M D Intriligator (Amsterdam: North-Holland) pp 1247–1318.

Chaturvedi, R. K., Joshi, J., Jayaraman, M., Bala, G. and Ravindranath, N. (2012) Multi-model climate change projections for India under representative concentration pathways *Curr. Sci.* **103** 791-802.

Chellaney, B. (2011) *Water: Asia's New Battleground* (Washington, DC: Georgetown University Press) pp 210-15.

Contor, B.A. & Taylor, R.G. (2013) Why improving irrigation efficiency increases total volume of consumptive use. *Irrig. Drain.* **62** (3), 273-280.

Cui, D., Liu, H., Min, J. and He, J. (1994) *Atlas of Phenology for Major Crops in China*, Meteorological Press, Beijing.

Currell, M., Han, D., Chen, Z. and Cartwright, I. (2012) Sustainability of groundwater usage in northern China: dependence on paleowaters and effects on water quality, quantity, and ecosystem health, *Hydrological Processes*, **26**, 4050-4066.

Currell, M. (2014) Mega-scale groundwater quality challenges and the need for an inter-disciplinary approach, *Hydrogeology Journal*, **22**, 745-748, doi:10.1007/s10040-014-1119-z.

de Graaf, I.E.M., van Beek, L.P.H., Wada, Y., & Bierkens, M.F.P. (2014) Dynamic attribution of global water demand to surface water and groundwater resources: Effects of abstractions and return flows on river discharges. *Advances in Water Resources*, **64**, 21-33.

Dee, D.P., et al. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553-597.

Dell, M., Jones,, B. and Olken, B. (2014) What Do We Learn from the Weather? The New Climate-Economy Literature *J. Econ. Lit.* **52**(3) 740-98.

Deng, J., Zhou, Z., Zhu, B., Zheng, X., Li, C., Wang, X. and Jian, Z. (2011) Modeling nitrogen loading in a small watershed in southwest China using a DNDC model with hydrological enhancements*, Biogeosciences, 8*(10) 2999-3009, doi: 10.5194/bg-8-2999-2011.

Dhawan, B. D. (1989) Mounting antagonism towards big dams *Econ. Polit. Wkly.* **24** 1096-99.

Dickinson, R.E. (1984) Modeling evapotranspiration for three-dimentional global climate models. *Climate Processes and Climate Sensitivity, Geophys. Monogr.* No. 29, Maurice Ewing Volume, Amer. Gephys. Union, 58-72.

Dingman, L. S. (2002) *Physical Hydrology, 2nd edition* (Upper Saddle River: Prentice Hall) p 646.

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G. and Scanlon, B. R. (2012) Impact of water withdrawals from groundwater and surface water on continental water storage variations*, Journal of Geodynamics, 59*, 143-156, doi: 10.1016/j.jog.2011.05.001.

Döll, P. and Siebert, S. (2002) Global modeling of irrigation water requirements*, Water Resour. Res., 38*(4) 1037, doi: 10.1029/2001WR000355.

Edwards, M. (1989) Global gridded elevation and bathymetry (etopo5), digital raster data on a 5-minute geographical grid. *NOAA National Geophyicsal Data Center, Boulder, Colorado.*

Elliott, J. et al. (2014) Constraints and potentials of future irrigation water availability on agricultural production under climate change*, Proc. Natl. Acad. Sci. U. S. A.,* **111**(9) 3239-3244, doi: 10.1073/pnas.1222474110.

ESRI (1992) Arc/world 1:3m digital data on CD-ROM: Users guide and data reference. Environmental Systems Research Institute, Redlands, CA.

Famiglietti, J .S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., Swenson, S. C., de Linage, C. .R and Rodell, M. (2011) Satellites measure recent rates of groundwater depletion in California's Central Valley *Geophys. Res. Lett.* **38** L03403.

FAO/UNESCO 2003 Digital Soil Map of the World and Derived Soil Properties Version 3.6. CD ROM.

Federer, C. A., Vörösmarty, C.  J. and B. Fekete (1996) Intercomparison of methods for calculating potential evaporation in regional and global water balance models*, Water Resour. Res., 32*(7) doi: 10.1029/96WR00801.

Fekete, B.F., Vörösmarty, C. J., Lammers, R.B. (2001) Scaling gridded river networks for macroscale  hydrology: Development, analysis, and control of error, *Water Resour. Res.,* 37(7), 1955-1967.

Fishman, R. M., Siegfried, T., Raj, P., Modi, V. and Lall, U. (2011) Over‐extraction from shallow bedrock versus deep alluvial aquifers: Reliability versus sustainability considerations for India's groundwater irrigation *Water Resour. Res.* **47** W00L05.

Fishman, R. M. (2012) Climate change, rainfall variability, and the adaptation through irrigation: evidence from Indian agriculture *George Washington University mimeo* https://sites.google.com/site/ramfishman/RFishman_JMP.pdf?attredirects=0

Foley, J. A. et al. (2011) Solutions for a cultivated planet*, Nature,* **478**(7369) doi: 10.1038/nature10452.

Food and Agriculture Organization of the United *Nations Food and Agriculture Organization of the United Nations Statistical Database (FAOSTAT*) (Rome: FAO) faostat.fao.org

Food and Agriculture Organization of the United Nations, International Fund for Agricultural Development, and World Food Programme (2015) *The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress* (Rome: FAO)  http://www.fao.org/3/a-i4646e.pdf

Foster, A. and Sekhri, S. (2008) Can expansion of markets for groundwater decelerate the depletion of groundwater resources in rural India? *Brown Univ. Working Paper* http://people.virginia.edu/~ss5mj/water_mkt.pdf).

Foster, S.S.D. & Perry, C.J. (2010) Improving groundwater resource accounting in irrigated areas: a prerequisite for promoting sustainable use. *Hydrogeology journal*, **18**, 291-294.

Frolking, S., Yeluripati, J. B. and Douglas, E. (2006) New district-level maps of rice cropping in India: a foundation for scientific input into policy assessment *Field Crops. Res.* **98**(2-3) 164-77.

Gadgil, S. and Kumar, K. R. (2006) The Asian Monsoon- agriculture and economy *The Asian Monsoon* ed B Wang (Honolulu: Springer/Praxis Publishing) pp 660-70.

Ghassemi, F. and White, I. (2007) *Inter-Basin Water Transfer: Case Studies from Australia, United States, Canada, China and India* (Cambridge: Cambridge University Press).

Gleeson, T., Wada, Y., Bierkens M.F.P, & van Beek, L.P.H. (2012) Water balance of global aquifers revealed by groundwater footprint. *Nature* **488**, 197-200.

Gleick, P. Making every drop count. *Scientific American* (2001).
Food and Agriculture Organization of the United *Nations Food and Agriculture Organization of the United Nations Statistical Database (FAOSTAT)* (Rome: FAO) faostat.fao.org (2015).

Greene, W. (2004) Fixed effects and bias due to the incidental parameters problem in the Tobit model *Econom. Rev.* **23**(2) 125-47.

Grogan, D. S., Zhang, F., Prusevich, A., Lammers, R. B., Wisser, D., Glidden, S., Li, C., and Frolking, S. (2015) Quantifying the link between crop production and mined groundwater irrigation in China *Sci. Total Environ.* **511** 161-75.

Global Runoff Data Centre (2013): Long-Term Mean Monthly Discharges and Annual Characteristics of GRDC Station / Global Runoff Data Centre. Koblenz, Germany: Federal Institute of Hydrology.

Government of India (2014) *Agricultural Statistics at a Glance* (New Delhi: Ministry of Agriculture).

Government of India (2014) *Key Indicators of Situation of Agricultural Households in India* (New Delhi: National Sample Survey Office, Ministry of Statistics and Programme Implementation ) http://mospi.nic.in/mospi_new/upload/KI_70_33_19dec14.pdf

Guerra, L. C., Bhuiyan, S. I., Tuong, T. P. and Barker, R. (1998) Producing more rice with less water from irrigated systems *System-Wide Initiative on Water Management Pap. 5* (Colombo, Sri Lanka: IWMI).

Hafeez, M.M., Bauman, B.A.M., Van de Giesen, N., & Vlek, P. (2007) Scale effects on water use and water productivity in a rice-based irrigation system (UPRIIS) in the Philippines. *Agric. Water Manage.* **92** (1-2), 81-89.

Hamon, W. R. (1963) Computation of direct runoff amounts from storm rainfall, *International Association of Sciences Hydrology Publications, 63*, 52-62.

Han, J., Z. Jia, W. Wu, C. Li, Q. Han, and J. Zhang (2014) Modeling impacts of film mulching on rainfed crop yield in Northern China with DNDC, *Field Crops Res., 155*, 202-212, doi: 10.1016/j.fcr.2013.09.004.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J. and Piontek, F. (2013) A trend-preserving bias correction - the ISI-MIP approach *Earth. Syst. Dynam*. **4** 219-36.

Honoré, B., Vella, F. and Verbeek, M. (2008) *The Econometrics of Panel Data,* ed L Mátyás and P Sevestre (Berlin Heidelberg: Springer) pp 385-418.

Hsiang, S. M. (2010) Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America *Proc. Natl. Acad. Sci. USA* **107**(35) 15367-72.

Huke, R. E. and Huke, E. H. (1997) *Rice Area by Type of Culture: South, Southeast, and East Asia: A Revised and Updated Data Base* (Los Baños: International Rice Research Institute).

Jain, S. K., Reddy, N. S. R. K. and Chauge, U. C. (2005) Analysis of a large inter-basin water transfer system in India *Hydrolog. Sci. J.* **50** 1-137.

Jack, B. K. (2011) Constraints on the adoption of agricultural technologies in developing countries. *White paper, Agricultural Technology Adoption Initiative* (Cambridge, MA; Berkeley, CA: J-PAL, MIT and CEGA, UC Berkeley) http://www.atai-research.org/sites/default/files/ATAIwhitepaper12102011.pdf

Kang, Y., S. Khan, and X. Ma (2009) Climate change impacts on crop yield, crop water productivity and food security - A review, *Progress in Natural Science, 19*(12) doi: 10.1016/j.pnsc.2009.08.001.

Knighton, D. (1998) *Fluvial Forms and Processes: A New Perspective.* Oxford University Press, Inc., Arnold, London.

Konikow, L.F. (2011) Contribution of global groundwater depletion since 1900 to sea-level rise. *Geophys. Res. Lett.* **38**, L17401.

Konikow, L.F., and S.A. Leake (2014) Depletion and capture: Revisiting "the source of water derived from wells", *Groundwater*, 52, 100-111, doi:10.1111/gwat.12204.

Kumar, K. K., Kumar, R. K., Ashrit, R. G., Deshpande, N. R. and Hansen, J. W. (2004) Climate impacts on Indian agriculture *Int. J. Climatol.* **24** 1375-1393.

Lancaster, T. (2000) The incidental parameters problem since 1948 *J. Econom.* **95**(2) 391-413.

Lehner, B. et al. (2011) High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management*, Frontiers in Ecology and the Environment, 9*(9) 494-502, doi: 10.1890/100125.

Li, C. S., S. Frolking, and T. A. Frolking (1992) A Model of Nitrous-Oxide Evolution from Soil Driven by Rainfall Events .1. Model Structure and Sensitivity*, Journal of Geophysical Research-Atmospheres, 97*(D9) 9759-9776.

Li, C. S., S. Frolking, and T. A. Frolking (1992) A Model of Nitrous-Oxide Evolution from Soil Driven by Rainfall Events .2. Model Applications*, Journal of Geophysical Research-Atmospheres, 97*(D9) 9777-9783.

Li, C. S., S. Frolking, and R. Harriss (1994) Modeling Carbon Biogeochemistry in Agricultural Soils*, Global Biogeochem. Cycles, 8*(3) 237-254, doi: 10.1029/94GB00767.

Li, C. S., J. J. Qiu, S. Frolking, X. M. Xiao, W. Salas, B. Moore, S. Boles, Y. Huang, and R. Sass (2002) Reduced methane emissions from large-scale changes in water management of China's rice paddies during 1980-2000*, Geophys. Res. Lett., 29*(20) 1972, doi: 10.1029/2002GL015370.

Li, C. (2007) Quantifying greenhouse gas emissions from soils: Scientific basis and modeling approach*, Soil Sci. Plant Nutr., 53*(4) 344-352, doi: 10.1111/j.1747-0765.2007.00133.x.

Liu, Q. and Z. Yang (2010) Quantitative estimation of the impact of climate change on actual evapotranspiration in the Yellow River Basin, China*, Journal of Hydrology, 395*(3-4) 226-234, doi: 10.1016/j.jhydrol.2010.10.031.

Liu, C., and H. Zheng (2004) Changes in components of the hydrological cycle in the Yellow River basin during the second half of the 20th century, *Hydrological Processes,* 18, 2337-2345, doi:10.1002/hyp.5534.

Liu, Y., S. Li, F. Chen, S. Yang, and X. Chen (2010) Soil water dynamics and water use efficiency in spring maize (Zea mays L.) fields subjected to different water management practices on the Loess Plateau, China*, Agric. Water Manage., 97*(5) 769-775, doi: 10.1016/j.agwat.2010.01.010.

Ma, J., A. Y. Hoekstra, H. Wang, A. K. Chapagain, and D. Wang (2006) Virtual versus real water transfers within China, *Philosophical Transactions of the Royal Society B-Biological Sciences, 361*(1469) 835-842, doi: 10.1098/rstb.2005.1644.

Maidmen, D.R. (1992) *Handbook of Hydrology*. McGraw-Hill, Inc., Columbus, Ohio.

Majumdar, D. K. (2013) *Irrigation Water Management: Principles and Practice, 2nd Edition* (Delhi : PHI Learning Pvt. Ltd.) pp 295-399.

Menon, A., Levermann, A., Schewe, J., Lehmann, J. and Frieler, K. (2013) Consistent increase in Indian monsoon rainfall and its variability across CMIP-5 models *Earth Sys. Dynam.* **4** 287-300.

Moiwo, J.P., Y. Yang, H. Li, S. Han, and Y. Yang (2010) Impact of water resource exploitation on the hydrology and water storage in Baiyangdian Lake, *Hydrological Processes*, 24, 3026-3039, doi:10.1002/hyp.7716.

Molden, D. and T. Y. Oweis (2007) Pathways for increasing agricultural productivity, in *Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture*, edited by D. Molden, pp. 279-310, IMWI, Earthscan, London.

Molden, D. & Bos, M.G. (2005) Improving basin water use in linked agricultural ecological and urban systems: seeing new flow paths and avoiding dead ends. *Water Sci. Technol.* **51** (8), 147-154.

Mueller, C. and R. D. Robertson (2014) Projecting future crop productivity for global economic modeling, *Agricultural Economics, 45*(1) 37-50, doi: 10.1111/agec.12088.

Mukherji, A., Shah, T. and Giordano, M. (2012) *Managing energy irrigation nexus in India: a typology of state interventions* (Sri Lanka: International Water Management Institute Water Policy Research Highlight http://www.iwmi.cgiar.org/iwmi-tata/PDFs/2012_Highlight-36.pdf)

Nash, J.E. and Sutcliffe, J.V. (1970) River flow forecasting through conceptual models part I – A discussion of principles. *J. Hydrol.,* **10**, 282-290.

Neyman, J. and Scott, E. (1948) Consistent Estimates Based on Partially Consistent Observations. *Econometrica* **16**(1) 1-32.

Nickell, S. (1981) Biases in dynamic models with fixed effects *Econometrica* **49**(6) 1417-26.

NBS (National Bureau of Statistics of China) (2008) *China Statistical Yearbook*, National Bureau of Statistics of Chin, Beijing.

Pang, Z., T. Huang, and Y. Chen (2010) Diminished groundwater recharge and circulation relative to degrading riparian vegetation in the middle Tarim River, Xinjiang Uygur, Western China, *Hydrological Processes*, 24, 147-159, doi:10.1002/hyp.7438.

Park, C. (1977) World-wide variations in hydraulic geometry exponents of stream channels – Analysis and some observations *J. Hydrol. 33* 133-46.

Pierce, L.T. (1958) Estimating seasonal and short-term fluctuations in evapotranspiration from meadow crops. *B. Am. Meteorol. Soc.* **39**, 73-38.

Portmann, F. T., S. Siebert, and P. Döll (2010) MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, *Global Biogeochem. Cycles,* 24, GB1011, doi: 10.1029/2008GB003435.

Prusevich, A. A., A. I. Shiklomanov, S. Frolking, S. Glidden, R. B. Lammers, and D. Wisser (2013) Log-Exponential Reservoir Operating Rules for Global And Regional Hydrological Modeling, *Eos Transactions, AGU, Fall Meeting Supplement, abstrct GC21B-0827,* **94**.

Qiu, J., T. Huajun, S. Frolking, S. Boles, C. Li, X. Xiangming, J. Liu, Y. Zhuang, and X. Qin (2003) Mapping Singe-, Double-, and Triple-crop Agriculture in China at $0.5^o$ x $0.5^o$ by Combining County-scale and Census Data with a Remote Sensing-derived Land Cover Map, *Geocarto Internationa,* **18**.

Qiu, J., C. Li, L. Wang, H. Tang, H. Li, and E. Van Ranst (2009) Modeling impacts of carbon sequestration on net greenhouse gas emissions from agricultural soils in China, *Global Biogeochem. Cycles,* **23**, GB1007, doi: 10.1029/2008GB003180.

Rienecker, M. M. et al. (2011) MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.,* **24**(14) 3624-3648, doi: 10.1175/JCLI-D-11-00015.1.

Rodell, M., Velicogna, I., and Famiglietti, J. S. (2009) Satellite-based estimates of groundwater depletion in India *Nature* **460** 999-1002.

Rohwer, J., Gerten, D. and Lucht, W. (2007) *Development of Functional Irrigation Types for Improved Global Crop Modelling* PIK Report 104 (Potsdam: Potsdam Institute for Climate Impact Research) https://www.pik-potsdam.de/research/publications/pikreports/.files/pr104.pdf

Russo T *personal communication.*

Saha, S., et al. (2011) The NCEP climate forecast system version 2. *J. Climate,* **27**, 2185-2208.

Schlenker, W. (2006) The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions *Rev. Econ. Stat.* **88**(1) 113-125.

Seckler, D., Amarasinghe, U., Molden, D., De Silva, R. and Barker, R. (1998) *World Water Demand and Supply, 1990 to 2025: Scenarios and Issues* (Colombo: International Water Management Institute).

Seckler, D., Molden, D. & Sakthivadivel, R. (2003) The concept of efficiency in water resources management and policy.  In: Kijne, J.W., Barker, R., & Modlde, D. (Eds.), Water Productivity in Agriculture: Limits and Opportunities for Improvement.  CAB international.

Sekhri, S. (2011) Public Provision and Protection of Natural Resources: Groundwater Irrigation in Rural India *Am. Econ. J. Appl. Econ.* **3** 29-55.

Sekhri, S. (2012) *India Policy Forum* vol 9*,*ed *S* Shah et al (New Delhi: Sage) pp 149-86.

Sekhri, S. (2014) Wells, Water, and Welfare: The Impact of Access to Groundwater on Rural Poverty and Conflict *Am. Econ. J. Appl. Econ.* **6** 76-102.

Shah, T. and Verma, S. (2008) Co-management of electricity and groundwater: an assessment of Gujarat's Jyotirgram Scheme *Econ. Polit. Wkly.* **43** 59-66.

Shah, T. (2010) *Taming the Anarchy: Groundwater Governance in South Asia* (Washington, DC: Resources for the Future Press) pp 5-33, 59-90, 110.

Sharma, B. R., Rao, K. V., Vittal, K. P. R. and Amarasinghe, U. A. (2015) Realizing the potential of rainfed agriculture in India,  *Drafts prepared for the IWMI-CPWF project on 'Strategic Analysis of India's National River-Linking Project'* (Colombo, Sri Lanka: IWMI).

Sheffield, J., G. Goteti, and E. F. Wood (2006) Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, *J. Climate*, **19**(13) 3088-3111.

Shi, X., D. Yu, E. D. Warner, X. Pan, G. W. Petersen, Z. G. Gong, and D. C. Weindorf (2004) Soil database of 1:1,000,000 digital soil survey and reference system of the Chinese genetic soil classification system*, Soil Survey Horizons,* **45**, 129-136.

Shiklomanov, I. A. and J. C. Rodda (Eds.) (2003) *World Water Resources at the Beginning of the 21st Century, UNESCO International Hydrology Press Ed.*, Cambridge University Press.

Siebert, S., et al., Groundwater use for irrigation – a global inventory. *Hydrol. Earth Syst. Sci.*, **14**, 1863-1880 (2010).

Siebert, S. and P. Döll, (2010) Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation*, Journal of Hydrology, 384*(3-4) doi: 10.1016/j.jhydro1.2009.07.031.

Siegfried, T., Sobolowski, S., Raj, P., Fishman, R., Vasquez, V., Narula, K., Lall, U. and Modi, V. (2010) Modeling irrigated area to increase water, energy, and food security in semiarid India *Wea. Climate Soc.* **2** 255-70.

Simons, G.W.H., Bastiaanssen, W.G.M, & Immerzeel, W.W. Water reuse in river basins with multiple users: A literature review. *J. Hydrol.* **522**, 558-571 (2015).

Singh, R. B., Kumar, P. and Woodhead, T. (2002) *Smallholder farmers in India: Food Security and Agricultural Policy* (Bangkok: FAO) ftp://ftp.fao.org/docrep/fao/005/ac484e/ac484e00.pdf

Smilovic, M., Gleeson, T. and Siebert, S. (2015) The limits of increasing food production with irrigation in India *Food Secur.* **7** 835-56.

Sooraj, K. P., Terray, P. and Mujumdar, M. (2014) Global warming and the weakening of the Asian summer monsoon circulation: assessments from the CMIP5 models *Clim. Dyn.* **45** 233-52.

Stewart, R.J., Wollheim, W.M., Miara, A., Vörösmarty, C. J., Fekete, B., Lammers, R.B., Rosenzweig, B. (2013) Horizontal cooling towers: riverine ecosystem services and the fate of thermoelectric heat in the continental Northeast US *Environ. Res. Lett.,* **8**, doi:10.1088/1748-9326/8/2/025010.

Sue Wing, I. and De Cian, E. (2014) Integrated assessment: Modelling agricultural adaptation *Nat. Clim. Chang.* **4** 535-36.

Syed, T. H., J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson (2008) Analysis of terrestrial water storage changes from GRACE and GLDAS*, Water Resour. Res.,* **44**(2) W02433, doi: 10.1029/2006WR005779.

Tang, C., J. Chen, S. Shindo, Y. Sakura, W. Zhang, and Y. Shen (2004) Assessment of groundwater contamination by nitrates associated with wastewater irrigation: A case study in Shijiazhuang region, China, *Hydrological Processes*, **18,** 2303-2312, doi:10.1002/hyp.5531.

Tang, H., J. Qiu, E. Van Ranst, and C. Li (2006) Estimations of soil organic carbon storage in cropland of China based on DNDC model, *Geoderma,* **134**(1-2) 200-206, doi: 10.1016/j.geoderma.2005.10.005.

Tang, Q., X. Zhang, and Y. Tang (2013) Anthropogenic impacts on mass change in North China, *Geophys. Res. Lett.,* **40**(15) 3924-3928, doi: 10.1002/grl.50790.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A. (2012) An Overview of CMIP5 and the experiment design *B. Am. Meteorol. Soc.* **93** 485-98.

Tilman, D., K. G. Cassman, P. A. Matson, R. Naylor, and S. Polasky (2002) Agricultural sustainability and intensive production practices, *Nature,* **418**(6898) 671-677, doi: 10.1038/nature01014.

Törnqvist, R. & Jarsjö, J. (2011) Water savings through improved irrigation techniques: basin-scale quantification in semi-arid environments. *Water Resour. Manage.* **26** (4), 949-962.

United Nations, Department of Economic and Social Affairs, Population Division (2013), *World Population Prospects: The 2012 Revision*, New York.

Vörösmarty, C. J., Federer, C. A. and Schloss, A. L. (1998) Evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling, *Journal of Hydrology,* **207**(3-4) doi: 10.1016/S0022-1694(98)00109-7.

Vörösmarty, C.J., Green, P., Salisbury, J. & Lammers, R.B. (2000a) Global water resources: Vulnerability from climate change and population growth. *Science,* **289**, 284.

Vörösmarty, C. J., Fekete, B. M., Meybeck, M. and Lammers R. B. (2000b) Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution, *Journal of Hydrology,* **237**(1-2) 17-39, doi: 10.1016/S0022-1694(00)00282-1.

Vörösmarty, C. J., Green, P., Salisbury, J. & Lammers, R.B. (2010c) A simulated topological network representing the global system of rivers at 30-min spatial resolution (STN-30) *Global Biogeochem. Cycles.* **14** 599–621.

Vörösmarty, C. J., Dougles, E. M., Gren, P. A. and Revenga, C. (2005) Geospatial indicators of emerging water stress: An application to Africa, *Ambio*, **34**(3), 230-236.

Wada, Y., Wisser, D. and Bierkens, M. F. P. (2014) Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, *Earth System Dynamics,* **5**, 15-40.

Wada, Y., van Beek, L. P. H. and Bierkens M. F. P. (2012) Nonsustainable groundwater sustaining irrigation: A global assessment, *Water Resour. Res.,* **48**, W00L06, doi: 10.1029/2011WR010562.

Wada., Y., Gleeson., T, & Esnault, L. (2014) Wedge approach to water stress. *Nature Geoscience,* **7**, 615 -617.

Wang, L., Qiu, J., Tang, H., Li, H., Li, C. and Van Ranst E. (2008) Modelling soil organic carbon dynamics in the major agricultural regions of China, *Geoderma,* **147**(1-2) 47-55, doi: 10.1016/j.geoderma.2008.07.009.

Ward, F.A. & Pulido-Velazquez, M. (2008) Water conservation in irrigation can increase water use. *Proc. Natl. Acad. Sci. USA,* **105** (47), 18215-18220.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O. and Schewe, J. (2014) The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project framework *Proc. Natl. Acad. Sci. USA* **111**(9) 3228-32.

Willmott, C.J. & Matsuura, K. (2001) Terrestrial air temperature and precipitation: Monthly and annual time series (1950-1999), http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html.

Wisser, D., B. M. Fekete, C. J. Vörösmarty, and A. H. Schumann (2010) Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network-Hydrology (GTN-H) *Hydrology and Earth System Sciences,* **14**(1) 1-24.

Wisser, D., Frolking, S., Douglas, E. M., Fekete, B. M., Vörösmarty, C. J. and Schumann A. H. (2008) Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, *Geophys. Res. Lett.,* **35**(24) L24408, doi: 10.1029/2008GL035296.

Wisser, D., Frolking, S., Douglas, E. M., Fekete, B. M., Schumann, A. H. and Vörösmarty, C. J. (2010) The significance of local water resources captured in small reservoirs for crop production - A global-scale analysis *J. Hydrol .(Amst.)* **384** 264-75.

Wollheim, W.M., Peterson, B.J., Thomas, S.M., Hopkinson, C.H., Vörösmarty, C. J. (2008a) Dynamics of N removal over annual time periods in a suburban river network. *Journal of Geophysical Research*, **113**.

Wollheim, W.M., Vörösmarty, C. J., Bouwman, A.F., Green, P., Harrison, J., Linder, E., Peterson, B.J., Seitzinger, S.P., Syvitski, J.P.M. (2008ab) Global N removal by

freshwater aquatic systems using a spatially distributed, within-basin approach. *Global Biogeochemical Cycles,* **22**(2).

Wooldridge, J. M. (2010) *Econometric analysis of cross section and panel data* (Cambridge, MA: MIT press).

World Bank (1998) *India - Water resources management sector review: groundwater regulation and management report* (Washington, DC: World Development Sources).

World Bank (2015) *Global Monitoring Report 2014/2015: Ending Poverty and Sharing Prosperity* (Washington, DC: World Bank and the International Monetary Fund).

Yasutomi, N., Hamada, A. and Yatagai, A. (2011) Development of a long-term daily gridded temperature dataset and its application to rain/snow discrimination of daily precipitation *Global Environmental Research* **15**(2) 165-72.

Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I. and Hamada, A. (2012) A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges *SOLA* **5** 137-40.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi N. and Kitoh, A. (2012): APHRODITE: Constructing a Long-term Daily Gridded Precipitation Dataset for Asia based on a Dense Network of Rain Gauges, *Bulletin of American Meteorological Society*, **93**, 1401-1415, doi:10.1175/BAMS-D-11-00122.1.

Ye, L. et al. (2013) Climate change impact on China food security in 2050*, Agronomy for Sustainable Development,* **33**(2) 363-374, doi: 10.1007/s13593-012-0102-0.

Yaun, Z, and SHen, Y. (2013) Estimation of Agricultural Water Consumption from Meteorological and Yield Data: A Case Study of Hebei, North China, *PLoS ONE*, 8(3): e58685.  doi:10.1371/journal.pone.0058685.

Yu, C., Li, C., Xin, Q., Chen, H., Zhang, J., Zhang, F., Li, X., Clinton, N., Huang, X., Yue, Y. and Gong, P. (2014). Dynamic assessment of the impact of drought on agricultural yield and scale-dependent return periods over large geographic regions, *Environmental Modelling & Software,* published online 4 September 2014.

Zhang, Y., Kendy, E., Qiang, Y., Changming, L., Yanjun, S. and Hongyong, S. (2004) Effect of soil water deficit on evapotranspiration, crop yield, and water use efficiency in the North China Plain, *Agricultural Water Management*, **64**, doi:101016/S0378-3774(03)00201-4.

Zhang, Y., Li, C. S., Trettin, C. C., Li, H. and Sun G. (2002) An integrated model of soil, hydrology, and vegetation for carbon dynamics in wetland ecosystems*, Global Biogeochem. Cycles,* **16**(4) 1061, doi: 10.1029/2001GB001838.