

# Water Resources Research®



## RESEARCH ARTICLE

10.1029/2022WR032804

# Irrigation by Crop in the Continental United States From 2008 to 2020

P. J. Ruess<sup>1</sup> , M. Konar<sup>1</sup> , N. Wanders<sup>2</sup> , and M. Bierkens<sup>2</sup> 

<sup>1</sup>Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Champaign, IL, USA, <sup>2</sup>Department of Physical Geography, Utrecht University, Utrecht, The Netherlands

### Key Points:

- A national database of crop-specific irrigation water use by source is developed
- Animal feed uses the most irrigation water compared to other crops across all water sources
- Rice decreased surface water use by 71%, sugar beets increased groundwater use by 232%, and rapeseed increased groundwater depletion by 405% from 2008 to 2020

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

M. Konar,  
[mkonar@illinois.edu](mailto:mkonar@illinois.edu)

### Citation:

Ruess, P. J., Konar, M., Wanders, N., & Bierkens, M. (2023). Irrigation by crop in the Continental United States from 2008 to 2020. *Water Resources Research*, 59, e2022WR032804. <https://doi.org/10.1029/2022WR032804>

Received 13 MAY 2022

Accepted 10 DEC 2022

### Author Contributions:

**Conceptualization:** P. J. Ruess, M. Konar, N. Wanders, M. Bierkens  
**Data curation:** P. J. Ruess, M. Konar, N. Wanders, M. Bierkens  
**Formal analysis:** P. J. Ruess, M. Konar  
**Investigation:** P. J. Ruess, M. Konar  
**Methodology:** P. J. Ruess, M. Konar, N. Wanders, M. Bierkens  
**Resources:** P. J. Ruess, M. Konar  
**Software:** P. J. Ruess, M. Konar, N. Wanders, M. Bierkens  
**Supervision:** N. Wanders, M. Bierkens  
**Visualization:** P. J. Ruess, M. Konar  
**Writing – original draft:** P. J. Ruess, M. Konar, N. Wanders, M. Bierkens  
**Writing – review & editing:** P. J. Ruess, M. Konar, N. Wanders, M. Bierkens

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

**Abstract** Agriculture is the largest user of water in the United States. Yet, we do not understand the spatially resolved sources of irrigation water use (IWU) by crop. The goal of this study is to estimate crop-specific IWU from surface water withdrawals (SWW), total groundwater withdrawals (GWW), and nonrenewable groundwater depletion (GWD). To do this, we employ the PCR-GLOBWB 2 global hydrology model to partition irrigation information from the U.S. Geological Survey Water Use Database to specific crops across the Continental United States (CONUS). We incorporate high-resolution input data on agricultural production and climate within the CONUS to obtain crop-specific irrigation estimates for SWW, GWW, and GWD for 20 crops and crop groups from 2008 to 2020 at county spatial resolution. Over the study period, SWW decreased by 20%, while both GWW and GWD increased by 3%. On average, animal feed (alfalfa/hay) uses the most irrigation water across all water sources: 33 from SWW, 13 from GWW, and 10 km<sup>3</sup>/yr from GWD. Produce used less SWW (43%), but more GWW (57%), and GWD (27%) over the study time-period. The largest changes in IWU for each water source between the years 2008 and 2020 are: rice (SWW decreased by 71%), sugar beets (GWW increased by 232%), and rapeseed (GWD increased by 405%). These results present the first national-scale assessment of irrigation by crop, water source, and year. In total, we contribute nearly 2.5 million data points to the literature (3,142 counties; 13 years; 3 water sources; and 20 crops).

## 1. Introduction

The United States is a major agricultural producer and breadbasket for the world, making it important to understand how specific crops use irrigation water supplies throughout the country and over time. Irrigation water use (IWU) is primarily composed of surface water sources and withdrawals of groundwater, including non-renewable withdrawals of groundwater (i.e., groundwater depletion [GWD]). Increased reliance on unsustainable water use in agriculture poses a potential future risk to the sustainable production of the global agricultural supply (Godfray et al., 2010; Siebert et al., 2015), making IWU particularly important to consider. In this study, we estimate crop-specific surface water withdrawals (SWW), (total) groundwater withdrawals (GWW), and GWD within the Continental United States (CONUS). Of note, we refer to total irrigation of all crops as “crop irrigation,” and irrigation of particular crops (corn, wheat, etc.) as “crop-specific irrigation” for the remainder of this article. The main goal of this study is to better understand how irrigation water supplies (by source) are allocated amongst specific crops throughout the country, as well as any changes that have occurred with time. To achieve our goal, we create a novel database of IWU by crop and water source in the United States, which we make freely available with the paper.

There are several agricultural growing regions in the US with different water demands. The “cornbelt” in the US Midwest is known for growing the majority of US corn and soybeans. Globally, the US is the largest producer of corn, producing ~30% of global corn in 2019 (FAO, 2021). As of 2019 Brazil is the largest global producer of soybeans, but the US is a close second, producing ~29% of soybeans globally in 2019 (FAO, 2021). Grain production in the US Midwest is primarily rainfed, although farm fields are increasingly equipped with irrigation as rainfall patterns shift with a changing climate. California is the “fruit and vegetable basket” for the country, with abundant production of predominantly nuts, seeds, and citrus crops, as well as all varieties of fruits and vegetables. California agriculture relies on conjunctive use of surface and groundwater irrigation. The USA as a country provides 55% of the world's almonds, 37% of pistachios, 13% of walnuts, 13% of lettuce/chicory, 11% of strawberries, 9% of sugar beets, 8% of grapes, and 6% of tomatoes (FAO, 2021).

Only about 15% of harvested land in the US is irrigated, yet this land contributes 40% of the country's agricultural production (Lehrsch et al., 2014). Both surface water and groundwater are important for agricultural production. Groundwater use for irrigation has increased dramatically in the US since the Green Revolution began in the 1940s. Surface water withdrawals for irrigation also increased from 1950 to 1980, but have experienced a general downward trend from 1980 to 2015 (Dieter et al., 2018; Hutson et al., 2004; Kenny et al., 2009; MacKichan, 1951, 1957, 1961; Maupin et al., 2014; Murray, 1968; Murray & Reeves, 1972, 1977; W. B. Solley et al., 1993, 1998; W. R. Solley et al., 1983, 1988). In the last several decades, groundwater irrigation has been physically unsustainable both globally and throughout the US (Famiglietti & Rodell, 2013; Gleeson et al., 2012), occurring at a withdrawal rate that exceeds natural groundwater recharge and consequently leading to GWD. In fact, the US is one of the largest users of groundwater for irrigation globally, with the US ranked fifth behind India, Iran, Pakistan, and China, respectively (Dalin et al., 2017).

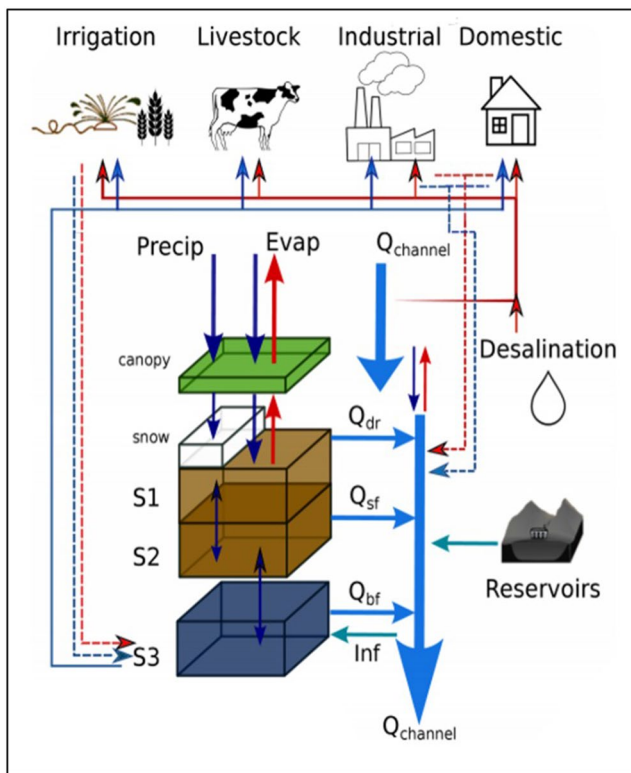
Groundwater irrigation has become increasingly important to crops. In 1950, irrigation water came predominantly from surface supplies (77%), with groundwater contributing 23%. Since then, groundwater has grown in importance, now contributing 48% of total irrigation water (total irrigation includes crops, golf courses, and parks). The fractions are comparable when only irrigated cropland is considered (e.g., golf courses and parks are excluded): surface water accounted for approximately 55% of withdrawals in the US in 2015, with groundwater accounting for the remaining 45%. See Figure S1 in Supporting Information S1 for details on how irrigation water sources have changed over time in the US. Unsustainable GWD has been increasing over time in the US. Irrigated agricultural production supported by groundwater use is particularly concentrated over three aquifers in the US: the Central Valley, High Plains, and Mississippi Embayment. About 50% of GWD is concentrated in the High Plains and Central Valley (Scanlon et al., 2012). Regions that are being pumped at particularly unsustainable rates are the Southern part of the Central Valley Aquifer (i.e., the Tulare Basin) and the Southern High Plains in Texas. Importantly, the federal government in the US provides abundant information on its national agricultural, climate, and water systems.

In this study, we develop a novel framework to fuse statistical information with a hydrology model to partition IWU by crop. Specifically, we employ the global PCR-GLOBWB 2 model (Sutanudjaja et al., 2018) with highly resolved data inputs for the CONUS to obtain crop-specific water demand estimates. PCR-GLOBWB (the original version, hereafter “PCR-GLOBWB 1”; Van Beek et al., 2011) captures sector-specific water demand, groundwater and surface water withdrawal, water consumption, and return flows, and has been used in previous studies to simulate unsustainable groundwater in irrigation in the United States using a comprehensive water demand and irrigation module (Wada et al., 2011, 2012). To do this, PCR-GLOBWB 1 simulates gross crop water demand for irrigated crops and the irrigation and rainwater that is available to meet this demand. The model simulates groundwater recharge, including return flow from irrigation, to estimate GWD. We input high-resolution data on crop parameters and climate within the CONUS to PCR-GLOBWB 2 (here referred to as PCR-CONUS). We then scale the US Geological Survey Water Use database (USGS, 2021) with our crop-specific values to ensure consistency with widely accepted irrigation information.

The goal of this study is to estimate crop-specific IWU throughout the CONUS. The CONUS has a wealth of freely available information related to crop locations, irrigation water withdrawals, and climate that can inform crop water use in hydrology models. We develop a data-driven approach that fuses statistical information with a global hydrology model to estimate crop-specific IWU within the CONUS. Importantly, our approach to partition water demands by crops relies on a well-established hydrology model (e.g., PCR GLOBWB 2) and scaling to match total water volumes provided in the benchmark USGS Water Use database. Our research is motivated by the following questions: (a) What crops rely on IWU by source of water?, (b) How is crop-specific IWU spatially distributed throughout the country?, and (c) How has crop-specific IWU changed with time? Answering these questions allows us to generate a novel database of crop-specific IWU by water source at county spatial resolutions and annual time scales (2008–2020) for the CONUS. We make our entire crop-specific IWU database publicly available with this study.

## 2. Methods

We use an established global hydrology model (e.g., PCR-GLOBWB 2: Sutanudjaja et al., 2018) with high-resolution data inputs for the US to partition IWU for distinct crops. We then constrain model estimates to



**Figure 1.** Overview of PCR-GLOBWB 2 grid-cell and its modeled states and fluxes; see Figure 1 in Sutanudjaja et al. (2018) for original image as well as abbreviations (reproduced from Sutanudjaja et al. (2018)).

the U.S. Geological Survey (USGS) Water Use database (USGS, 2021) to ensure that our results are consistent with this widely accepted source.

We apply the global PCR-GLOBWB 2 model (Sutanudjaja et al., 2018) to the Continental United States (CONUS). PCR-GLOBWB 2 is a “state-of-the-art grid-based global hydrology and water resources model” natively supporting up to 5 arc-min spatial resolutions and daily hydrology and water use temporal resolutions, with hydrodynamic river routing at sub-daily timesteps (Sutanudjaja et al., 2018). From Sutanudjaja et al., 2018, “PCR-GLOBWB 2 simulates moisture storage in two vertically stacked soil layers (S1 and S2 in Figure 1), as well as the water exchange among the soil, the atmosphere, and the underlying groundwater reservoir (S3 in Figure 1).” PCR-GLOBWB 2 further integrates human water use categorized for irrigation, livestock, industry, and households demands. Note that PCR-GLOBWB 2 focuses on modeling individual grid-cells, therefore limiting the ability to account for water transfers between grid-cells.

Irrigation water specifically is partitioned based on two thresholds for (a) surface water irrigation and (b) fossil groundwater irrigation. Both thresholds are applied to surface water abstraction fractions (Siebert et al., 2013) to define the preference for surface water versus groundwater for irrigation purposes: the surface water threshold maximizes preferential use of surface water when above the threshold, while the fossil groundwater threshold only allows fossil GW above this threshold. Irrigation efficiencies then account for how much water must be supplied for crop water demands to be met.

Groundwater is overall modeled as changes in storage dynamics. GW occur when surface water irrigation does not meet local water demands, resulting in groundwater allocations and changes to modeled groundwater storage. Groundwater depletion is not explicitly modeled and is instead calculated as (modeled) GW minus (modeled) groundwater recharge. Groundwater

recharge is modeled as percolation minus capillary rise, accounting for region-specific aquifer properties consistent with original PCR-GLOBWB 2 model runs. The interested reader is referred to Sutanudjaja et al. (2018) for additional details on the PCR-GLOBWB 2 model.

We estimate water use per 5 arc-min grid-cell using model simulations, which we then aggregate up to counties using zonal statistics methods. Broadly, for each grid-cell in each time-step, we have a particular irrigation demand for each landcover type represented in that grid-cell. This irrigation demand comes from crop curves constructed using crop coefficients and crop calendars. As each grid-cell represents many landcover types, we use a fractional coverage to describe how much of the grid-cell is occupied by each landcover type. These landcover types are then related to the crop specific irrigation water demands, taking into account total area covered within the grid-cell. Climate inputs then determine how much of the irrigation water demand is met naturally (this is almost entirely rainfall), after which we apply available surface water until the irrigation demand is met. If the demand is not met through combined natural sources and surface water irrigation, we then irrigate using groundwater resources. Finally, once we have a modeled estimate of GWW for irrigation in each grid-cell, we then take into account the groundwater recharge modeled over that same grid-cell and we take GWD IWU as the difference between GWW and the sum of long-term net groundwater recharge and river-bed infiltration for each grid-cell.

Importantly, we do not make any changes to the structure of the PCR-GLOBWB 2 model. This is because the novelty of our study is focused in running the model with high-resolution data inputs for the CONUS which we then fuse with the USGS statistical database. Specifically, we collect and implement data on the national agricultural system, including the locations of specific crops, crop growing seasons, irrigation locations, and efficiencies, and crop coefficients for evapotranspiration demands. Additionally, we utilize high-resolution climate forcing data. The new forcing data that we implement in this study is listed in Table 1. The goal of our study is to run the PCR-GLOBWB 2 model with CONUS-specific agricultural and climate forcing data to estimate crop-specific IWU from both surface water and groundwater. We detail the new input data below.

**Table 1**  
*High-Resolution Data Inputs Used in This Study*

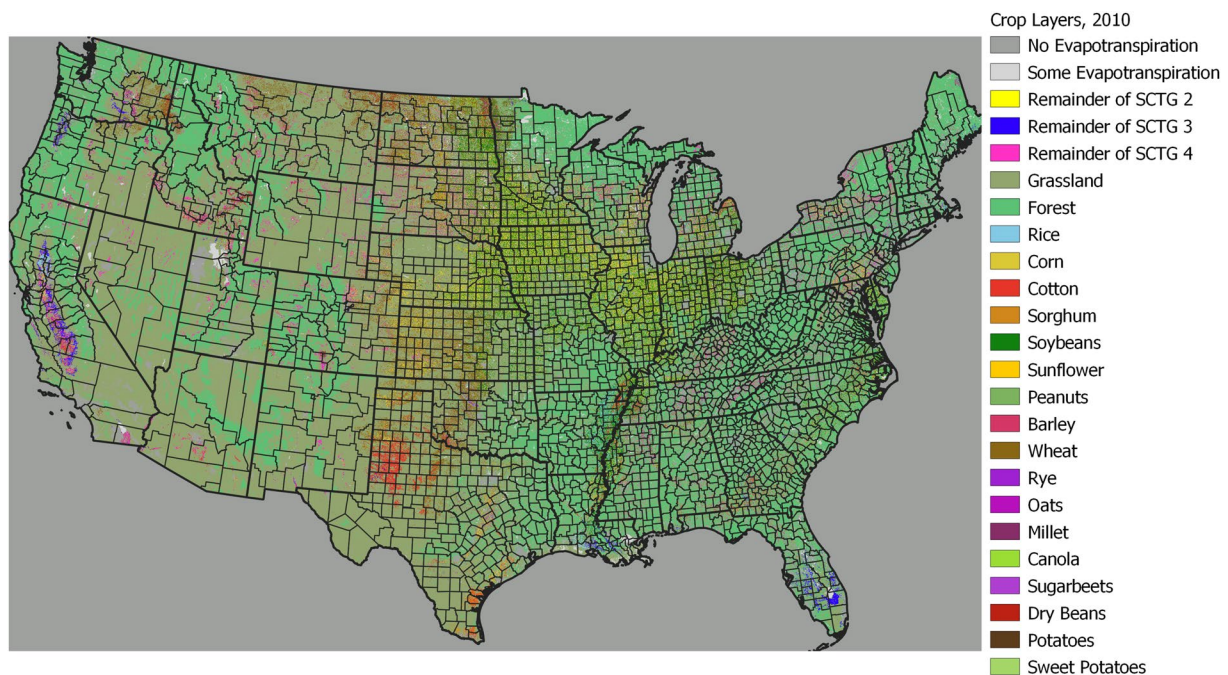
Data	PCR-GLOBWB 2	PCR-CONUS (this study)
Crop locations	MIRCA, Portmann et al. (2010)	CropScape Han et al. (2012)
Climate	WFDE5, Cucchi et al. (2020)	GridMET, Abatzoglou (2013)
Crop calendars	MIRCA, Portmann et al. (2010)	CCD, Sacks et al. (2010)
Crop coefficients	MIRCA, Portmann et al. (2010)	FAO, Allen et al. (1998)
Irrigated areas	MIRCA, Portmann et al. (2010)	USGS, USGS (2021)
Irrigation efficiency	Wada et al. (2014)	Efficiency, Brouwer et al. (1989)

*Note.* This table only lists PCR-GLOBWB 2 data inputs that were changed in this study. For a full list of model data requirements please refer to Sutanudjaja et al. (2018).

### 2.1. High-Resolution Landcover and Crop Locations

Global model runs of PCR-GLOBWB 2 include six unique landcover types: open water, forest, grassland, rain-fed crops, paddy (rice), and non-paddy (all other irrigated crops). To improve model estimates and simultaneously compute crop-specific IWU, we modified the model to include 24 unique landcover types, including 17 unique crops: barley, corn, cotton, millet, oats, peanuts, potatoes, pulses, rapeseed, rice, rye, sorghum, soybeans, sugar beets, sunflower, sweet potatoes, and wheat. The seven remaining categories include forest, grassland, no ET (developed), and some ET (water), as well as three crop categories to capture remaining crops: other grains, other produce, and animal feed. These final three groups correspond to groupings commonly used for food classification (USDOT, 2022).

The original six landcover types are represented as percentage maps in PCR-GLOBWB 2, such that the total percentage of landcover when all are summed together is equal to one for all grid-cells. To compute percent landcover maps, we retrieved crop data from the CropScape Cropland Data Layer (Han et al., 2012). CropScape has annual crop landcover data at 30-m resolution, available for the entire CONUS from 2008 to 2020 at the time of writing. Upscaling from 30 m to 5 arc-min allows us to calculate percent landcover as the fraction of all 30-m



**Figure 2.** Map of the landcover categories included in this study at 30-m spatial resolution for the Continental United States in the year 2010. Crop locations are obtained from the CropScape Cropland Data Layer (Han et al., 2012).

grid-cells of each landcover type in each 5 arc-min grid-cell. We calculate this for all 24 landcover types for each of the 13 years from 2008 to 2020. Figure 2 shows original CropScape data at 30-m resolution after converting to our 24 landcover types in 2010, before calculating crop fractions. For details regarding how CropScape landcover data was translated into our final 24 landcover types, please refer to Table S1 and Figure S2 in Supporting Information S1.

## 2.2. Evapotranspiration Crop-Factors for All Landcover Types

Existing (global) PCR-GLOBWB 2 crop factors were computed using MIRCA2000 crop calendars (Portmann et al., 2010) alongside FAO crop coefficients (Allen et al., 1998) for two agricultural landcover groups: paddy (rice), and non-paddy (everything else).

New crop coefficients were instead calculated for PCR-CONUS using the Crop Calendar Data set (CCD, Sacks et al., 2010) with FAO crop coefficients (Allen et al., 1998). The CCD improves on MIRCA planting dates and harvest dates by combining FAO and United States Department of Agriculture (USDA) crop calendar information (Sacks et al., 2010). FAO data are similar to original PCR-GLOBWB 2 global inputs, except we select values specifically representative of the CONUS climate rather than a global climate (Allen et al., 1998). For example, for Soybeans we selected growing stage durations for “Central USA,” as opposed to the alternatives: “Tropics” or “Japan.”

In addition to 17 unique crops, we have three “other” categories: other grains, other produce, and other animal feed. For each of these cases we had to make assumptions with the data we had available. The “other grains” category is constructed using CCD information for sunflower (closest to growing season of other grain crops), growing stage durations for small grains, and crop coefficients for safflower. The “other produce” category is similarly constructed using CCD information for pulses, and approximated growing stage durations and crop coefficients to try and represent the most significant crops within this group (mostly nuts and grapes, as well as some citrus crops).

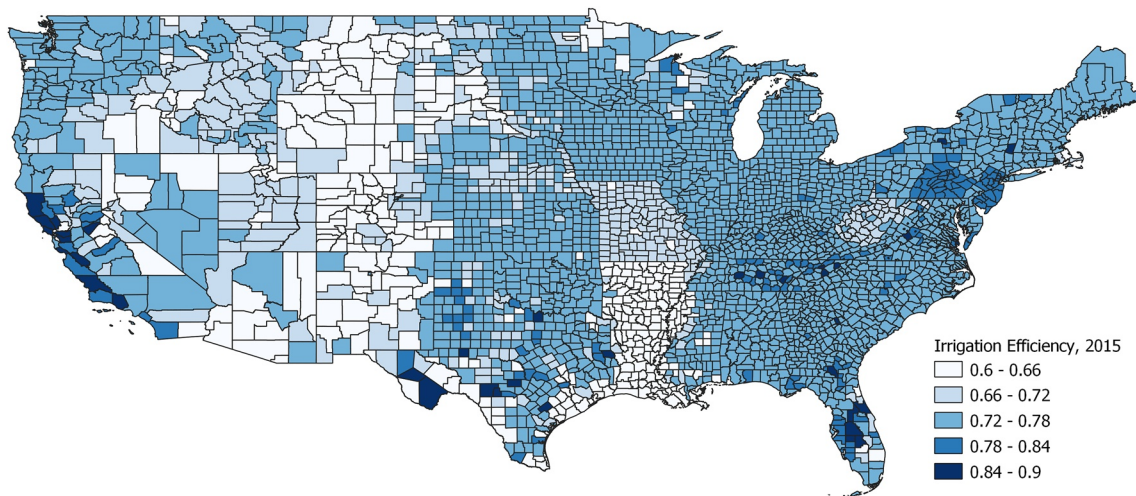
The “other animal feed” category is particularly challenging because it consists of alfalfa and other hay/non-alfalfa. We elected to use alfalfa crop coefficient information to represent this group, as the FAO specifies climate bounds for alfalfa. These climate bounds were useful for approximating an animal feed growing season, as the CCD does not have planting or harvesting dates for alfalfa or hay. We approximate the growing season as being all days between the last  $-4^{\circ}\text{C}$  day in Spring and the first  $-4^{\circ}\text{C}$  day in Fall (Allen et al., 1998). For crop coefficients, we first need to understand that Alfalfa is a perennial and therefore can be harvested many times in 1 year. The first harvest is the longest, while subsequent other harvests follow a slightly smaller crop curve pattern (these subsequent curves are modeled identically). Using this information, we created repeating alfalfa crop curves over the course of the year for each pixel within the CONUS. While theoretically we could reproduce this to have year-specific alfalfa crop coefficients (based on climate data), we took 2008 as a representative year and used our 2008 alfalfa crop coefficients for all subsequent years; this is consistent with all other crops, which are similarly modeled on a singular representative year's crop coefficients (rather than time-variant ones). We selected 2008 as it is closest to the CCD approximations for other crop calendars, which are based on observations of the late 1990s or early 2000s (Sacks et al., 2010). While we are comfortable using this climate-based approach for alfalfa in the mostly temperate climate of the CONUS, this approach may be problematic especially in lower latitudes (Sacks et al., 2010).

The final growth stage durations and crop coefficients used in this study for all 20 crops (including “other” categories) are documented in Table S2 in Supporting Information S1. Note that our crop coefficients are representative of general conditions and are assumed constant for all years in our study.

## 2.3. High-Resolution Irrigation Efficiency

PCR-GLOBWB 2 currently relies on global irrigation efficiency estimates from Rohwer et al. (2007). While this global coverage is necessary for global models, it is limited when applied to more localized models. For example, Rohwer et al. (2007) has only one value for the entire CONUS, set equal to 0.545 everywhere. This value is also time-invariant, based around the year 2000.

In contrast, water-use data has been published every 5 years at the US county scale since the year 2000 (USGS, 2021). These publications include irrigated area information for three different irrigation types: sprinkler,



**Figure 3.** Map of irrigation efficiencies used as input to the PCR-GLOBWB 2 model for this study. Values represent a percentage of irrigation water that gets to a crop, ranging from 60% to 90% depending on the irrigation methods used. This map is constructed from published water-use data (USGS, 2021) and irrigation efficiencies (Brouwer et al., 1989).

micro, and surface irrigation. Compiling this information with approximate irrigation efficiencies from Brouwer et al. (1989), we calculated weighted-average irrigation efficiencies for every CONUS county. To calculate irrigation efficiencies for unreported years, we assumed a linear trend between years with available data. For years after 2015, we used the 2015 value, as this is the most recent published data available. Our calculated irrigation efficiencies for the year 2015 are shown in Figure 3. Note that all values are between 0.6 and 0.9 because these are the irrigation efficiencies for surface and micro irrigation, respectively, while sprinkler irrigation efficiency falls in-between at 0.75 (Brouwer et al., 1989).

#### 2.4. High-Resolution Climate Forcing Data

PCR-GLOBWB 2 requires three meteorological data inputs that we have changed in this study: precipitation, reference evapotranspiration (for grass), and mean daily temperature.

In global model runs, PCR-GLOBWB 2 uses meteorological data from WFDE5 (see Table 1), which is published at a spatial resolution of 30 arc-min. For 5 arc-min PCR-GLOBWB 2 global model runs, it is therefore necessary to statistically downscale this resolution, which of course means that the spatial distribution of meteorological data within each  $6 \times 6$  block of 5 arc-min grid-cells remains static over time.

To improve on this spatial resolution, we used GridMET data (Abatzoglou, 2013), which is available at  $4 \times 4$  km resolution. For consistency with our other data sets, we upscale to 5 arc-min resolution, meaning that each of our 5 arc-min grid-cells represents the higher-resolution grid-cells used to construct them. This is an improvement over the downscaling approach used for global 5 arc-min PCR-GLOBWB 2 model runs, as our 5 arc-min grid-cells are all unique. Generally, our GridMET-based meteorological forcing variables are smaller than WFDE5 when averaged (temperature) or summed (precipitation and evapotranspiration, separately) over the CONUS, see Figure S3 in Supporting Information S1 for comparisons.

#### 2.5. High-Resolution Inputs

Using these high-resolution data inputs, we ran various versions of the model to assess the impact of specific changes individually, namely climate forcings, landcover, and irrigation efficiency data. To compare outputs, we focused specifically on groundwater resources over aquifer regions, summing GWW and GWD individually over the High Plains Aquifer, Central Valley Aquifer, and Mississippi Embayment Aquifers. We ran four new model iterations with our new landcover types compared against PCR-GLOBWB 2 original results: (a) All original PCR-GLOBWB 2 inputs, (b) New irrigation efficiency otherwise original PCR-GLOBWB 2 inputs, (c) New forcing data otherwise original PCR-GLOBWB 2 inputs, and (d) New irrigation efficiency and new forcing data otherwise original PCR-GLOBWB 2 inputs (these are the results discussed throughout this publication).

## 2.6. Constraining Estimates to Match Total Volumes Reported by USGS Water Use Database

After implementing all changes, PCR-CONUS estimates of SWW, GWW, and GWD for all defined crops and crop categories. We then scaled these final PCR-CONUS estimates to USGS Census data to resolve any differences. This has the combined benefit of 1). Allowing us to parse out crop-specific IWU values, and 2). Ensuring our total IWU estimates sum to credible USGS data values.

To accomplish this, we first summed PCR-CONUS raster outputs of SWW and GWW for all crops to each CONUS county. Using these county sums, we calculated individual SWW and GWW percent contributions by all crops for each county, then applied this percentage distribution to available USGS SWW and GWW water-use estimates. Specifically, we use USGS county-level crop irrigation census data from 2010 to 2015, scaling the years 2008–2014 to the 2010 census data, and scaling the years 2015–2020 to the 2015 census data. We scaled only to the years 2010 and 2015 because the USGS census data is only collected in 5-year intervals, and the 2020 data has not yet been published (USGS, 2021). 2010 and 2015 data were scaled directly to USGS data in these two years. For all other years in our 13-year period (2008–2020), we estimated USGS county-level data using the relationship of PCR-CONUS estimates between the year we are estimating and the nearest available year with USGS data. We then use these modified USGS estimates in off-years with crop-specific PCR-CONUS percent contributions to estimate each crop's contribution to SWW and GWW for every year.

Finally, because GWD data is not available from USGS, we used PCR ratios to upscale modeled GWD results. To do this, we took the ratio between PCR-CONUS GWD estimates and PCR-CONUS GWW estimates for all counties and years, and we multiplied this ratio by our scaled GWW values to get county-level GWD estimates.

Some complications arose with scaling PCR-CONUS data to USGS data, most notably in two cases. The first complication occurred when PCR-CONUS crop-specific estimates were extremely small for all years, meaning that when scaling in reference to a USGS year, sometimes the ratio (“scaling factor”) between SWW or GWW estimates was very large. This would result in unrealistic estimated USGS values. To avoid this issue, we constrained this scaling factor (of the current year we were estimating USGS data for, compared to the USGS reference year which was always 2010 or 2015) between 0.5 and 2. For county/crop pairs where the scaling factor was outside this range, we set the ratio to the average of all other counties with scaling factors within the acceptable range. So if a ratio was greater than 2 or less than 0.5, we set it equal to the average of all other counties with scaling factors within the 0.5–2 range. Note that this impacted a significant portion of counties: approximately 25% of counties had scaling factors below 0.5, while nearly 40% had scaling factors larger than 2.

The second complication occurred when PCR-CONUS estimates for all crops were zero in a county that had a positive USGS value. In these cases, to best match accepted USGS values, we calculated a ratio between the average current year PCR-CONUS values and the average reference year (2010 or 2015) PCR-CONUS values, and we applied this ratio to these counties where PCR estimated zero withdrawals. Effectively this means we apply an averaged crop fractional distribution to these counties where no crop-specific IWU was modeled by PCR-CONUS, to better match the USGS values. This complication was far less frequent, occurring in only about 5% of counties.

## 3. Results

Here we address our research questions and provide estimates of SWW, (total) GWW, and GWD by crop within the CONUS. We also compare our findings with other estimates in the literature.

### 3.1. What Crops Rely on Irrigation Water Use by Source of Water?

Surface water withdrawals, (total) GWW, and GWD values are summarized for all crops modeled in the CONUS in Table 2. “Other animal feed” (alfalfa/hay) has the largest IWU across the board, making up approximately 37% of total IWU, including 49% of SWW, 27% of groundwater abstractions (GWA), and 29% of GWD. “Other produce” is the next largest contributor to IWU (18% overall), making up about 14% of SWW, 20% of GWW, and 22% of GWD. Wheat, corn, rice, cotton, barley, and soybeans follow in that order when considering average IWU across years. Soybeans are interesting because their large GWA (7th largest) contrast with their significantly smaller surface water (14th) and GWD (14th) demands; most other crops are more consistent across water types. All crops not mentioned use significantly less water, summing to less than 7% of total crop IWU.

**Table 2**

*Surface Water Withdrawals, Groundwater Withdrawals, and Groundwater Depletion Mean Values, Volumetric Changes, and Percent Changes for All 20 Crops and Crop Groups Over the Continental United States*

Crop	Surface water withdrawals			Groundwater withdrawals			Groundwater depletion		
	Mean (km <sup>3</sup> )	Change (km <sup>3</sup> )	Change (%)	Mean (km <sup>3</sup> )	Change (km <sup>3</sup> )	Change (%)	Mean (km <sup>3</sup> )	Change (km <sup>3</sup> )	Change (%)
Barley	3.05	-0.66	-18.63	1.19	0.06	5.21	0.71	-0.12	-17.41
Corn	5.79	-2.70	-37.08	5.65	0.62	11.48	3.52	-0.17	-4.61
Cotton	1.76	-0.74	-41.74	2.00	-0.08	-5.19	1.42	0.15	15.10
Millet	0.06	-0.04	-37.50	0.07	0.02	27.21	0.04	0.00	-6.61
Oats	0.66	-0.55	-62.14	0.44	-0.21	-36.84	0.31	-0.14	-32.69
Other grains	0.59	0.12	16.92	0.56	0.47	105.50	0.38	0.25	78.26
Other produce	9.39	-4.95	-43.13	9.66	2.38	26.50	7.00	3.40	57.11
Other animal feed	32.93	1.98	5.78	12.90	4.20	34.47	9.57	2.77	30.03
Peanuts	0.06	0.00	1.54	0.29	-0.05	-14.73	0.03	-0.03	-49.00
Potatoes	1.03	0.23	28.82	0.57	0.12	28.45	0.45	0.06	16.71
Pulses	0.58	-0.21	-33.24	0.26	0.01	4.18	0.15	0.03	32.14
Rapeseed	0.03	0.01	42.55	0.01	0.01	206.58	0.01	0.01	405.45
Rice	2.73	-4.16	-71.38	5.14	-3.89	-48.07	3.85	-3.61	-59.69
Rye	0.08	-0.02	-29.74	0.09	0.06	95.39	0.04	0.01	28.40
Sorghum	0.39	0.05	13.28	0.95	0.09	9.04	0.64	-0.21	-27.58
Soybeans	0.38	-0.27	-64.11	1.71	-1.51	-72.59	0.08	-0.07	-69.16
Sugar Beets	0.41	0.04	16.84	0.24	0.22	231.61	0.19	0.19	210.75
Sunflower	0.08	-0.01	-6.26	0.08	-0.01	-7.34	0.04	0.00	3.87
Sweet Potatoes	0.01	0.00	36.60	0.02	0.00	-26.51	0.00	0.00	-97.61
Wheat	7.35	-3.99	-39.75	6.30	-0.91	-13.23	4.10	-1.57	-30.69
All Crops	67.38	-15.87	-20.17	48.16	1.60	3.21	32.52	0.96	2.82

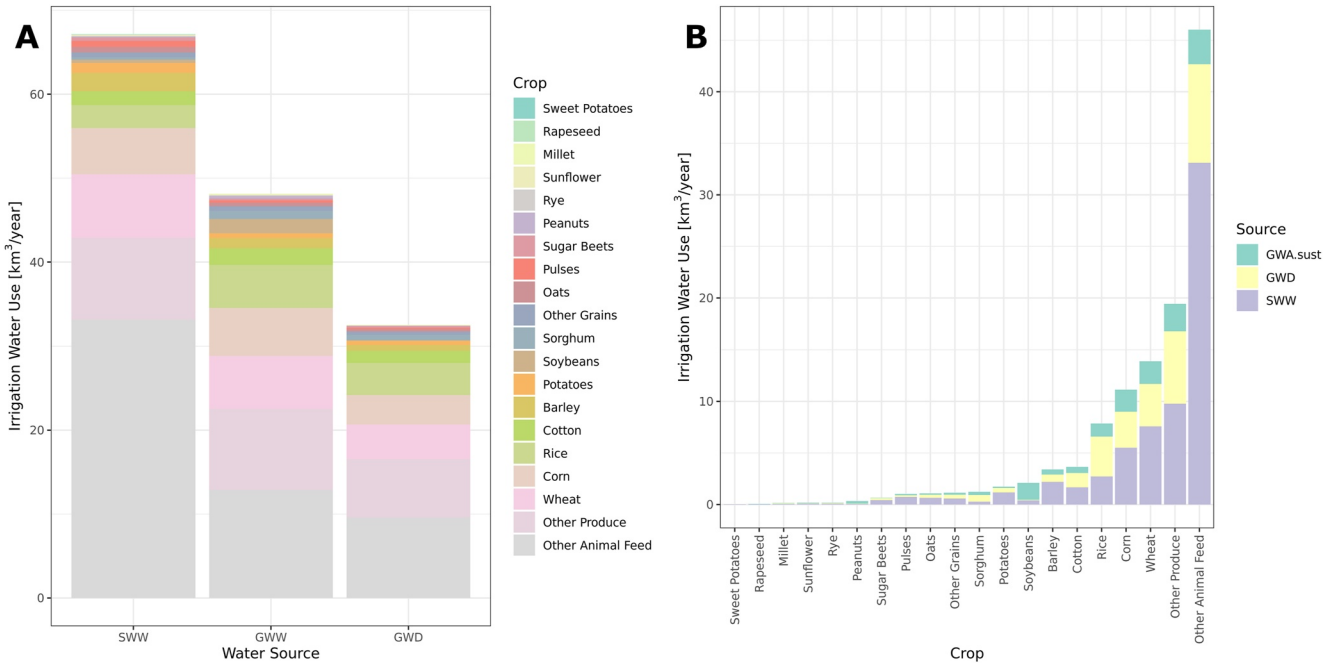
*Note.* Mean values include all years from 2008 to 2020, while volumetric change and percent change represent comparisons of the year 2020 to the year 2008.

Total IWU values averaged across the 13-year period are 67 (SWW), 48 (GWW), and 33 km<sup>3</sup> (GWD). The percent change in IWU between 2008 and 2020 across all crops shows a notable decrease in SWW (20%), while GWW and GWD instead both increase during this time-period by approximately 3% each. SWW and GWW values of course match USGS data due to our method of scaling model outputs to match USGS data. Notably the match is slightly different due to rounding errors and the capping of large scaling factors, but these differences are negligibly small. Part of our justification for scaling to USGS data is that PCR-CONUS has been shown to over-estimate GWW in certain locations in the CONUS, such as the Mississippi Embayment aquifer system (Scanlon et al., 2018). Scaling to USGS data helps mitigate this over-estimation problem.

Visualizations of averaged IWU estimates organized by water source (Figure 4a) and crop (Figure 4b) allow us to better see where IWU is being used. SWW is the largest user when averaging across the 13-year period, with GWW contributing about 70% as much IWU as compared to SWW (Figure 4a). GWD is a fraction of GWW, so Figure 4a also shows us that about 65% of all GWW are unsustainable (GWD), where the sustainable portion is only the remaining 35% of total GWW. This is because  $GWW = GWW_{\text{sustainable}} + GWW_{\text{unsustainable}}$ , where  $GWW_{\text{unsustainable}}$  is equal to GWD. Looking at unique crops, we see that “other animal feed” is the most significant user of IWU by far, with SWW making up ~70% of its total IWU, setting it apart from other crops (Figure 4b). When looking only at the groundwater variables ( $GWW_{\text{sustainable}}$  and GWD), there is notably less variability between crops. As noted previously, soybeans are also interesting due to their lack of reliance on SWW: almost all of their IWU comes from groundwater resources, most of which is sustainable (e.g., not contributing to GWD).

When looking at GWD as a fraction of GWW within the CONUS we see that soybeans on average have the lowest ratio, with GWD making up only about 5% of total GWW. Peanuts are closest, at 10%, while most other crops are



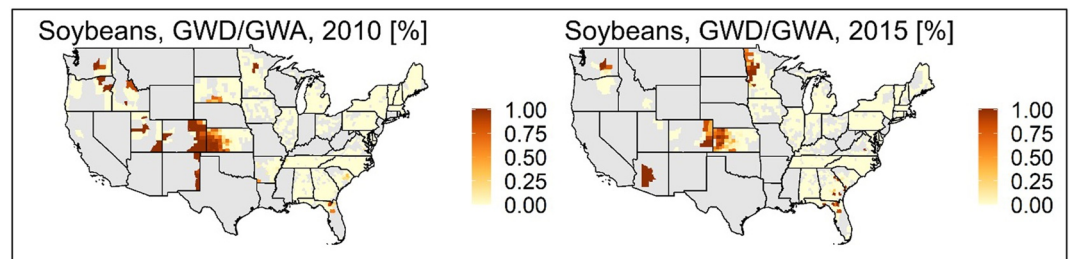


**Figure 4.** Irrigation Water Use averaged for all years, 2008–2020, and organized by panel (a). Water Source and panel (b). Crop. Of note, panel (b) includes a “sustainable” groundwater withdrawals (GWW) variable (GWW\_sust), which is the difference between total GWW and groundwater depletion.

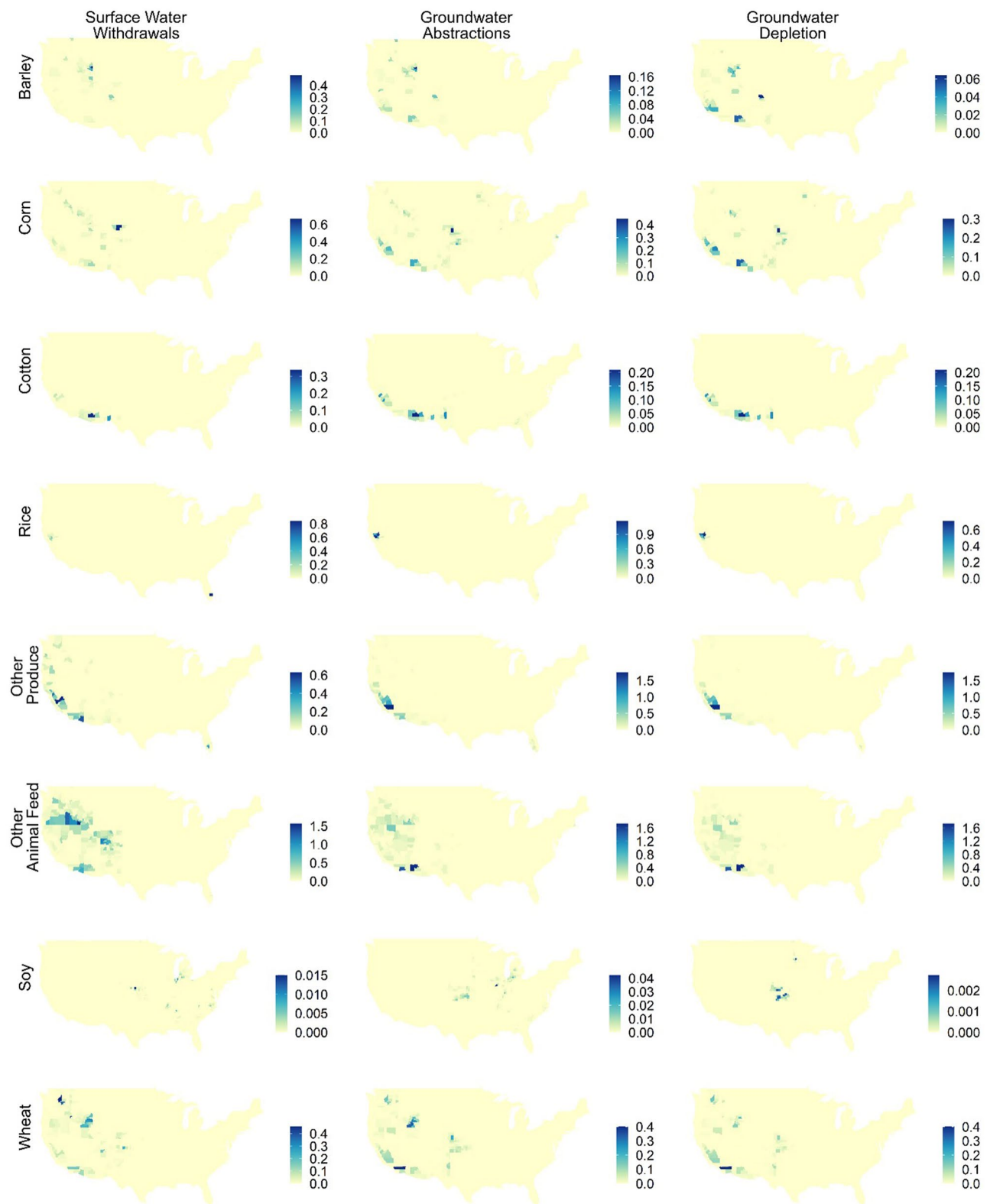
generally between the 50% and 80% range. The soybean GWD/GWA ratio is quite small along the midwestern corn-belt where soybeans are mostly grown (Figure 5), due to high recharge in these locations, leading soy to have the smallest GWD/GWA fraction.

### 3.2. How Is Crop-Specific Irrigation Water Use Spatially Distributed Throughout the Country?

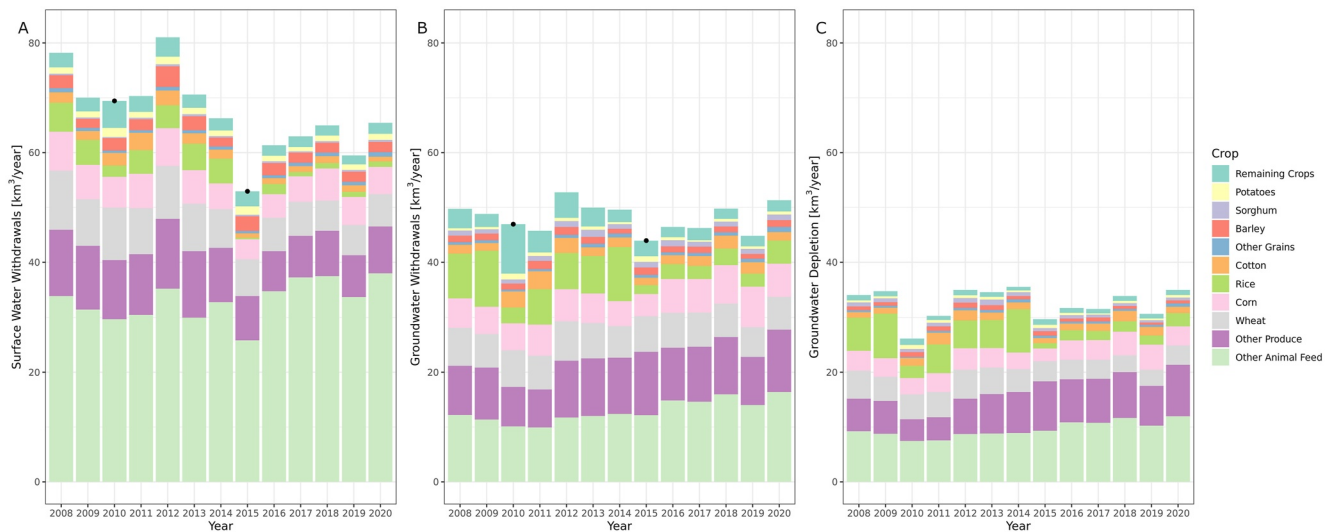
Crop-specific IWU (km<sup>3</sup>) for 2020 is mapped in Figure 6 for the top eight crops/categories. These maps are intentionally scaled individually to show where each crop contributes to SWW, GWW, and GWD. SWW and GWW are related in that GWW is only modeled when SWW is fully allocated (or was unavailable to begin with), while GWD is always smaller than GWW because GWW is the total amount of GWW while GWD is only the unsustainable amount of GWW. From these maps we see the top contributors to IWU, specifically barley, corn, “other animal feed,” and wheat are scattered throughout the West. Cotton interestingly sees highest IWU values in the Southwest, though cotton is of course historically grown in the Southeast; this pattern is likely due to the comparatively dry climate in the Southwest. Rice is irrigated most heavily in California, though a significant portion of the US rice supply comes from Southern states adjacent to the Mississippi river; this is similarly most likely due to climate differences. “Other produce” is unsurprisingly most heavily irrigated in California, where most US produce is grown. Soybeans are scattered throughout the Midwest as would be expected. Interestingly, we see soybean irrigation in the Midwest and not much corn irrigation in the Midwest; this is because corn is



**Figure 5.** Ratio of soybean groundwater depletion to soybeans groundwater abstractions for 2010 and 2015.



**Figure 6.** Maps of crop-specific irrigation water use by water source in 2020 in  $\text{km}^3/\text{yr}$ . The columns show irrigation water source: surface water withdrawals, groundwater withdrawals, and groundwater depletion. The rows show specific crops: barley, corn, cotton, rice, produce, other animal feed (alfalfa/hay), soy, and wheat. Note that all maps are scaled individually to their unique values, so they can not be compared directly to each other.



**Figure 7.** Time series of irrigation water use by crop and water source across the Continental United States from 2008 to 2020. (a) Crop-specific surface water withdrawals (SWW), (b) crop-specific total groundwater withdrawals (GWW), and (c) crop-specific groundwater depletion (GWD). Black dots on the SWW and GWW plots show corresponding U.S. Geological Survey (USGS) census withdrawals, which align with our estimates by design. There are no USGS values to compare GWD to in panel (c). Legends are stacked in order of top 10 based on GWD values in 2020, with an additional “other” category representing the sum of all un-specified crops in this plot.

grown expansively in the West as well, and these climates demand more irrigation. There is of course corn irrigation in the Midwest as well, it is just not as visible due to corn irrigation being comparatively larger in the West.

### 3.3. How Has Crop-Specific Irrigation Water Use Changed With Time?

As shown in Table 2 comparing 2020–2008, the largest percent changes in IWU are generally seen in more minor crops, such as rapeseed and sugar beets; this makes sense because these crops have smaller IWU values to begin with, so any difference will register as a larger fractional change. Interestingly, apart from “other animal feed,” most major crops have a decrease in SWW over the period. We can also see informative trends when observing particular crops. For example, “other animal feed” increases across all water types, while wheat and rice both decrease across all water types; this may be a result of changes to cropping patterns.

Figure 7 shows total SWW, GWW, and GWD summed over the CONUS by crop categories. For clarity, we have listed only the top 10 IWU based on 2020 GWD ranked values, with all other crops summed into the “other” category. This “other” category is always plotted on top for clarity, though because it includes many crops, it is often larger than some crops that have less IWU, such as potatoes and sorghum. We also see significantly larger “other” values particularly in GWW in 2010, which is likely due to soybeans having on average the 7th largest IWU from GWW sources but only the 14th largest IWU from GWD sources (see Table 2).

Comparing our scaled SWW and GWW values to USGS data of course shows very negligible differences because we have scaled to the USGS data. Very slight differences that do exist are due to rounding errors. For reference, the largest county difference we see when subtracting scaled PCR estimates from USGS values across all variables and years is  $2.22e^{-16}$  for GWW in 2015. Mapped comparisons of USGS minus scaled PCR estimates are available in Figure S4 in Supporting Information S1.

## 4. Discussion

### 4.1. Limitations

Due to a lack of available data detailing IWU by crop (a major motive for this research), we do not have any data to validate our results against. Instead, we constrain our PCR-CONUS estimates to match published USGS water use data by water source and county. This requires summing the volumetric use of water across crops to match these values. This means that our main novelty is the addition of modeled crop dimensionality (across both space

**Table 3**  
*Error Quantification for Surface Water and Groundwater Values in 2010 and 2015*

Error quantification		Surface water 2010	Surface water 2015	Groundwater 2010	Groundwater 2015
Raw	Mean error	−0.019	−0.014	−0.012	−0.011
	Standard deviation	0.113	0.085	0.065	0.076
	Root mean squared error (RMSE)	0.115	0.086	0.066	0.076
Normalized	Mean error	−0.030	−0.030	−0.024	−0.005
	Standard deviation	0.043	0.063	0.038	0.035
	Root mean squared error (RMSE)	0.044	0.064	0.039	0.035

*Note.* Errors here are defined as the difference when subtracting USGS reported county withdrawals from our PCR estimates aggregated to the county scale. All errors are calculated across all counties.

and time) to existing USGS water use data, which is generally considered the best available data when considering the entirety of the CONUS area. We believe our approach offers the best available solutions given current data limitations.

It is important to be aware of the limitations of this data set. First, we are restricted to the years 2008–2020 due to data input restrictions, most especially from the CropScape data set (Han et al., 2012). Additionally, other data sets are available only partially throughout the study period, such as irrigation efficiency (USGS, 2021) which is only available for select years; we applied methodologies that we consider reasonable to expand these data sets to other years, but of course it would be preferable to use primary data for these intermittent years. We also suffer from any existing limitations of the original PCR-GLOBWB 2 model structure, as this is left unchanged. For example, PCR-GLOBWB 2 always meets irrigation water demands using surface water resources first, then meets any remaining demands using groundwater (Sutanudjaja et al., 2018). Also, aside from the data inputs that we have explicitly changed, all other inputs to the model are original global data set inputs. As such, there is likely room to improve on our estimates by further changing some of these remaining global data sets to better represent the CONUS area specifically. Finally, our approach ultimately requires agricultural water use data to be available, for scaling our crop-specific model estimates to.

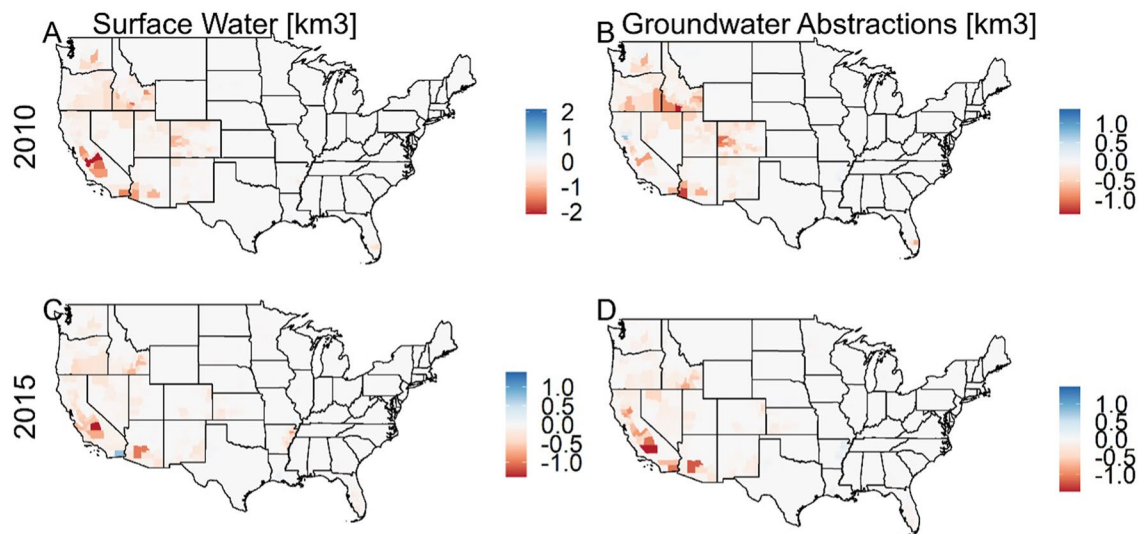
#### 4.2. Discrepancies Between PCR Estimates and USGS Irrigation Data

The main goal of this work is to use PCR-CONUS estimates to disaggregate USGS irrigation data into crop-specific irrigation water allocations. We constrain crop-specific water use volumes to sum to county-level water use data provided by USGS. This means that our total surface and groundwater withdrawal volumes in the final model estimates match USGS data. However, we recognize that understanding how our unscaled PCR-CONUS estimates compare with USGS irrigation data is valuable for future modeling efforts.

To quantify these differences, we first calculate mean error across all counties as PCR values minus USGS values to assess the error in PCR-CONUS estimates. Seeing as all results are negative (see Table 3), we can conclude that the PCR model consistently under-estimates withdrawals when compared with USGS values. This trend is also visible in scatterplot comparisons of the two variables, where each of the two withdrawals variables are normalized to be between 0 and 1 individually and we see that USGS values are almost exclusively larger than our PCR-CONUS estimates (see Figures S5–S10 in Supporting Information S1).

In all instances the standard error and Root Mean Squared Error (RMSE) are nearly identical (Table 3), this being because USGS values are scales of magnitude larger than PCR values, resulting in differences frequently being nearly equal to the USGS values. We can also see that the standard deviation and RMSE both are largest for surface water values in 2015, followed by 2010 surface water values which in turn are slightly larger than groundwater values in both years. This means simply that groundwater estimates are generally better than surface water estimates, particularly in the year 2015.

For an initial assessment of spatial variability, we can plot PCR estimates (Figure 6) and USGS reported values against each other (see Figures S5–S10 in Supporting Information S1). From these plots, spatial trends seem to be largely similar in California and along the western side of the Mississippi in 2010, though most other Western



**Figure 8.** Difference between PCR estimates and U.S. Geological Survey (USGS) reported withdrawals, calculated as PCR values minus USGS values. Differences are provided for surface water and groundwater in 2010 and 2015, in  $\text{km}^3$ .

counties seem to have significantly smaller PCR estimates when compared to the USGS values. Interestingly, we seem to see considerable differences between the 2010 versus 2015: while 2010 seems to be largely mismatched even in California, the Mississippi region seems to match better; contrastingly 2015 seems to spatially match better in California, while the Mississippi seems over-estimated.

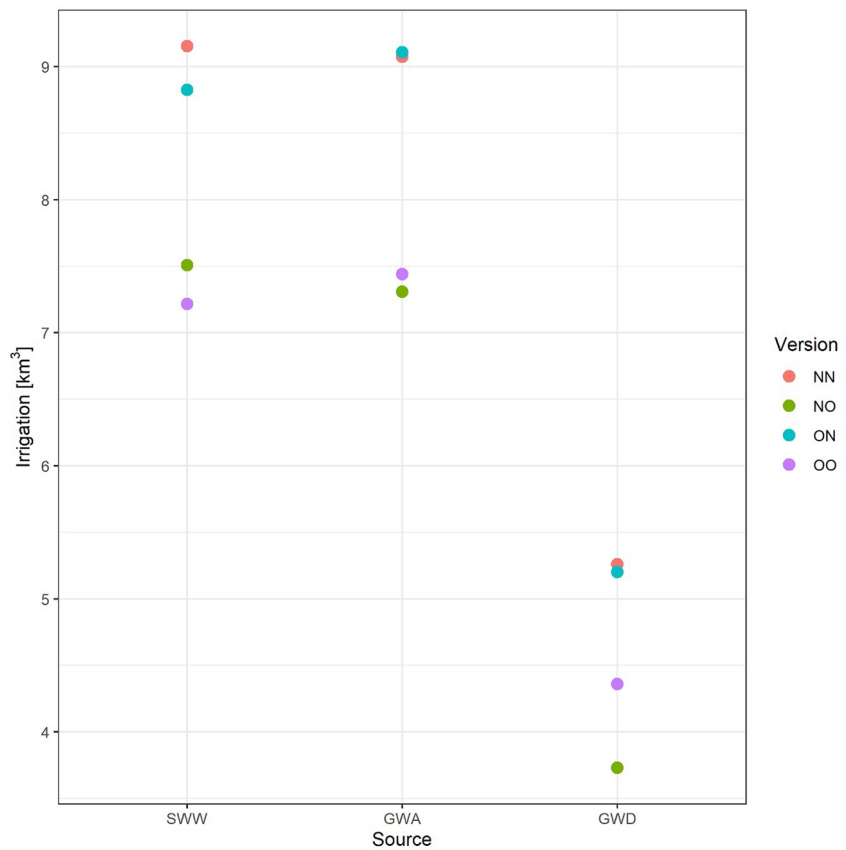
To better understand some of these differences, we can further consider spatial variability between volumes modeled by PCR and those provided by USGS (Figure 8). From Figure 8, we see PCR largely underestimates in the Western half of the country, particularly in California across years and water types. Groundwater sees larger differences outside California, particularly in Idaho and Oregon in 2010. We also see small positive values showing up in Imperial, CA and Colusa, CA, as well as along the Arkansas side of the Mississippi River, though these differences are volumetrically smaller than most other differences seen.

In addition to these spatial differences, we also see noticeable differences when comparing across the years 2010 and 2015. For example, the Mississippi Embayment surface water difference is negligible in 2010, but negative in 2015; the groundwater abstraction difference in this region is positive in 2015. This reiterates previous findings showing that PCR over-estimates water use over the Mississippi Embayment aquifer region (Scanlon et al., 2018).

Overall, these comparisons shed light on some important modeling issues. Volumetric PCR estimates are generally smaller than USGS values across the Western U.S.; this should be better addressed in future similar modeling efforts. Additionally, PCR estimates of surface water have worse (larger) standard deviations and RMSE, meaning that improvements to surface irrigation could be addressed in future PCR model runs for the CONUS.

### 4.3. Model Sensitivity

To better understand uncertainty and sensitivity of our model results, we ran four different model versions. All four of these model versions are distinguishable from previously published PCR-GLOBWB 2 results (Sutanudjaja et al., 2018) in that they incorporate the 20 unique crop types we constructed from CropScape data (Han et al., 2012). The differences between our four PCR-CONUS model versions are in some of the data inputs, namely the climate data and the irrigation efficiency data. Our baseline PCR-CONUS model uses all PCR-GLOBWB 2 data dependencies coupled with our new cropland information; we term this the “old climate, old irrigation efficiency” model (OO). Our newest and most complete PCR-CONUS model instead uses all new data inputs mentioned in this work, so named the “new climate, new irrigation efficiency” model (NN). Finally, we have two intermediary model versions to better quantify and assess the impacts of each of these data sets independently; these models are the “new climate, old irrigation efficiency” model (NO), and the “old climate, new irrigation efficiency” model (ON).

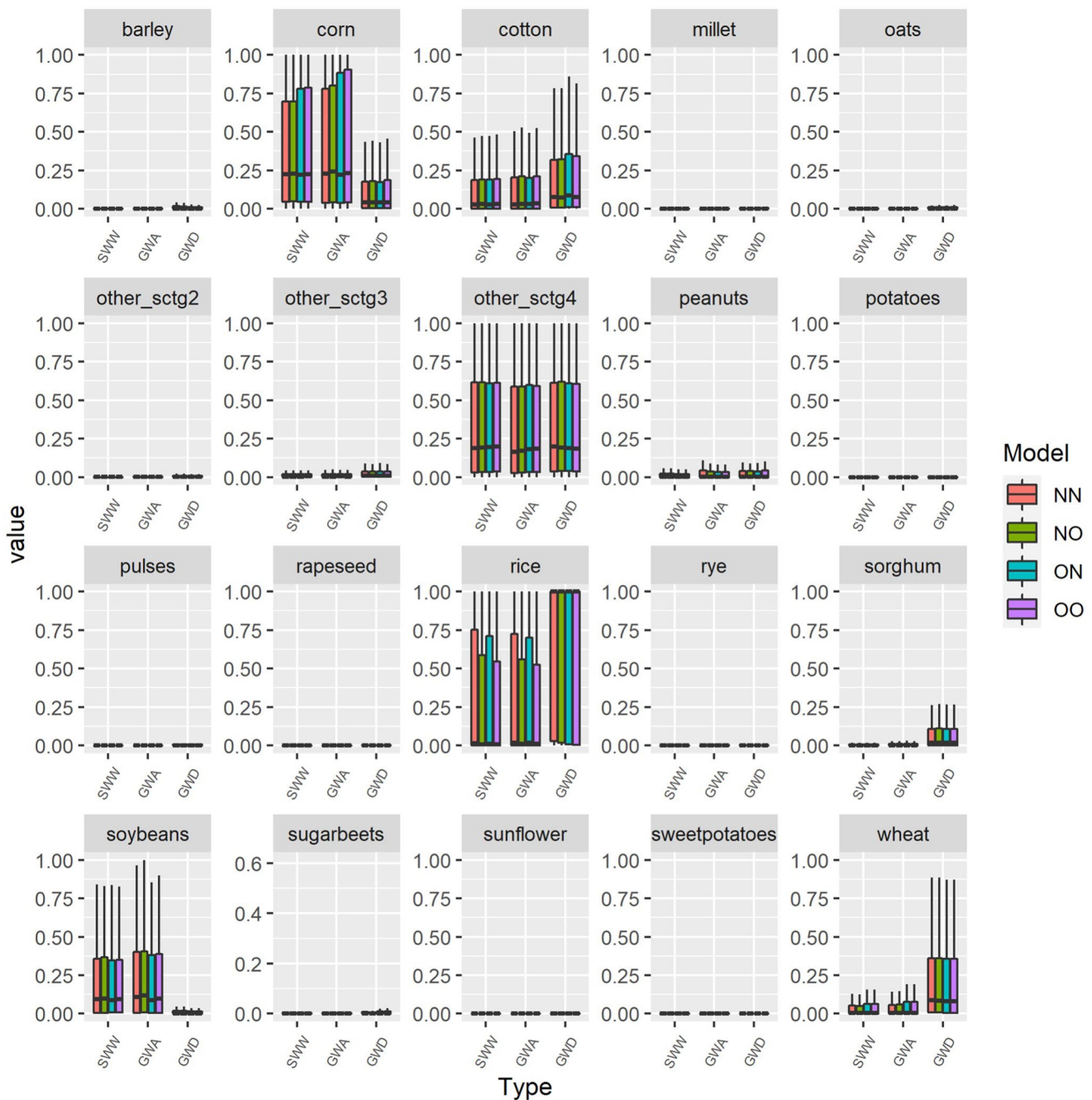


**Figure 9.** Total irrigation water use summed over all counties for four different model versions and all three unique water sources, where each different model version relies on different data inputs. The different versions are “NN” (new climate data and new irrigation data), “NO” (new climate data and old irrigation data), “ON” (old climate data and new irrigation data), and “OO” (old climate data and old irrigation data). Please note that “old” data refers to PCR-GLOBWB 2 global data inputs, while “new” data refers to PCR-CONUS data inputs.

The summed irrigation estimates across all crops for the entire CONUS spatial area are plotted and compared in Figure 9. From the figure we can see that the NN model (used for all values analyzed and published in this paper) generally results in a larger PCR-CONUS irrigation estimate as compared to other model versions. Additionally, we find that the new irrigation efficiency data in particular leads to larger irrigation estimates (NN and ON models). The differences between climate inputs also is noticeably smaller than the differences between the irrigation efficiencies, and this is consistent across all modeled water types. Another clear difference seen in Figure 9 is that the spread of the model estimates is largest for SWW and smallest for GWD, with GWD estimates having a comparably more even spread.

To investigate each crop individually, we constructed boxplots (Figure 10) showing the fractional contribution of each crop to the three different water sources (SWW, GWA, and GWD) in each county. These boxplots show the distribution of these fractional contributions to irrigation water estimates according to the same four unique model runs described previously. Note that these plots are visualized without outliers, and we’ve also removed any zero and NA values for clarity; this likely results in rice having such high crop fractions in GWD when rice actually relies on non-renewable groundwater irrigation in only a very small number of counties.

By comparing these boxplots in Figure 10, we can see that some crops have significantly more variability between models than other crops. Corn, for example, seems to have the smallest standard deviations with the new climate data inputs as compared to the old climate data inputs, across all water types, and the new climate input data also seems to fractionally allocate less surface and (renewable) groundwater to corn as compared to other crops where corn is also grown. Cotton seems to function similarly to corn in this way. Animal feed instead is fairly uniform across all model types, likely due to the huge coverage of animal feed throughout the



**Figure 10.** Boxplots showing the mean and standard deviation across all counties where each crop is grown for four different model versions reliant on slightly different data inputs. Note that these plots are constructed using only non-zero values. The different versions are “NN” (new climate data and new irrigation data), “NO” (new climate data and old irrigation data), “ON” (old climate data and new irrigation data), and “OO” (old climate data and old irrigation data). Please note that “old” data refers to PCR-GLOBWB 2 global data inputs, while “new” data refers to PCR-CONUS data inputs.

country, probably resulting in model differences being compensated in different geographical areas. Soybeans and wheat are similarly relatively consistent across model types. Rice is interesting because both surface and (renewable) groundwater irrigation seem to be much more sensitive to changes in irrigation efficiency over climate data inputs. This has to do with the fact that rice is flood irrigated, consequently relying on large amounts of irrigation water regardless of climate (and particularly reliant on irrigation efficiency). Note that there is little spatial variability in crop fractional water use (see Figures S8–S10 in Supporting Information S1).

#### 4.4. Uncertainty

Estimation of crop-specific irrigation relies on various agricultural data sets as inputs. Each data set meets particular purpose necessitating its inclusion in this analysis. First, landcover information is required for assigning each grid-cell's crop coefficient across space and time. Crop coefficients are used together with crop calendars in calculations of evapotranspiration, which represent the crop water demands we are interested in. Climate and irrigation efficiencies of course also play a role in how much water is available and how much irrigation water must be applied to meet crop water demands, respectively. For irrigation efficiencies to be informative, we also need spatial irrigation data in the form of irrigated areas defined by irrigation type, which together create a whole picture of irrigation efficiencies by grid-cell.

Ideally, we would like to quantify the uncertainty introduced by these input variables to the PCR-CONUS model to better understand their influence on our estimates. As we have access only to the data sets in their published forms, we do not have the information required to calculate their uncertainties for them. However, many of these data sets are heavily validated, hopefully minimizing any influence of uncertainties on their published results.

CropScape landcover data are validated and accuracy assessed against ground truth data. The USDA also publishes accuracy data with CropScape, describing the accuracy of each landcover type by state and year. FAO crop coefficients are representative of general conditions gleaned from an extensive area of research. While localized FAO data would have been preferable, this data simply isn't available in any uniform way across the entirety of the CONUS, so we opted for consistency across the CONUS rather than gap-filling with more detailed values where available. The CCD notes limitations but does not quantify uncertainty propagated from the data sets upon which it relies, leaving us no way to carry uncertainties through to our model. GridMET climate data are heavily validated against an extensive network of weather stations but similarly do not publish uncertainty information for us to consider. Uncertainties of the irrigation efficiencies published by the FAO are also unavailable.

We rely on USGS irrigation areas as a model input, and USGS water withdrawals for correcting systematic bias by scaling our model estimates; this data is all contained in published USGS water use data tables. The USGS receives most of their water use data directly from the states, which do not have the capacity to perform quality control, let alone full uncertainty analyses. Additionally, the USGS uses a variety of methods for data collection that differ for each State, and their full suite of methods is not fully transparent: some States measure their water use information while others model it with a variety of input variables. For these reasons, any uncertainty analyses will differ markedly across States, and we therefore cannot quantify USGS uncertainties across the CONUS. This is a shortcoming that is recognized in the literature, leading to consistency and methodological transparency becoming two priorities for the USGS moving forward to improve their Water Use Database (Marston et al., 2022). Over the next several years, USGS will be developing capacity for nationally consistent modeling approaches to estimate water withdrawals, consumptive use, and associated uncertainty at the daily scale for 12-digit hydrologic unit code subwatersheds (Marston et al., 2022).

While each data set has built-in uncertainty, additional uncertainties are further introduced by extrapolations and other estimations necessitated by a lack of data in some cases. For example, crop coefficients are estimated across crop groups by using representative crops within each group; these of course do not perfectly capture each crop individually, but instead are approximated to represent the whole of the category. The crop coefficients described in Supporting Information S1 are also assumed to be uniform in time and space within the CONUS for each crop category, due to the FAO publishing extremely limited information within the CONUS regarding spatial variability and no information about the temporal variability of these variables. While crop calendars were primarily taken from the CCD, crop calendars for alfalfa were assigned based on acceptable climate boundaries rather than factual planting and harvesting dates, for lack of available data. Additionally, CCD growing periods are based on the year 2010 and do not vary in time, which informed our decision to treat alfalfa using 2010 climate information rather than computed temporally variant alfalfa crop coefficients, to maintain consistency with other crops. Finally, irrigated areas are extrapolated from the years 2010 and 2015 out to all other years in our data set, due simply to data only being collected every 5 years. These uncertainties cannot be calculated as uncertainties in the input data sets are not provided.

#### 4.5. Implications for Future Work

Estimates of crop-specific IWU creates opportunities for future research and decision-making. While crop-specific water footprints have been widely studied (Hoekstra et al., 2011), differentiation between surface water and



groundwater sources has become increasingly important. Additionally, the most widely used crop-specific water footprint data set is representative of averaged climate from 1995 to 2005 (Mekonnen & Hoekstra, 2011), whereas our new data set provides annual values over a 13-year period, enabling time series analyses. We also create a path forward for calculating crop-specific water footprints in terms of mass, cost, and calorie output, and understanding how these variables may change in time.

Our work is limited to the CONUS region but outlines a new methodology for estimating agricultural water use applicable to any region, simultaneously motivating the creation of important, high-resolution, local input data sets of variables such as crop locations (Jackson et al., 2019) and irrigation information (Xie et al., 2021). By exploring this area of research, we also provide some insights as to what may be some of the more important data sets needed for modeling an area of comparable size to the CONUS. For example, the unique trends in soybean GWA may point to the need for more detailed information for this particular crop. Additionally, while we have elected to use crop coefficients which are solely spatially variant and not time-variant as a result of available data, future work may further explore the reasonableness of this assumption.

By improving IWU data, we also enable future water footprint studies in the CONUS. This extends to research in virtual water trade (Dalin et al., 2012; Gumidyala et al., 2020), flows (Dang et al., 2015), and storage (Ruess & Konar, 2019), which can all benefit from our new data set. Water footprint studies would be able to build on our estimates, which use the same commodity coding system as other national databases (ORNL, 2022; USDOT, 2022). Additionally, recent work downscaling this trade data to the county scale (Karakoc et al., 2022; Lin et al., 2019) combined with our work would enable highly resolved virtual water flow studies.

## 5. Conclusions

In this study, we estimated crop-specific IWU throughout the CONUS. Specifically, we quantify how SWW, total GWW, and GWD is allocated to each of 20 crops in each county and year from 2008 to 2020. To do this, we integrated statistical information with an existing global hydrological model. Specifically, we employed the PCR-GLOBWB 2 hydrological model with higher-resolution and localized inputs for the CONUS area. Then, we scaled our model output volumes to ensure matching with the total surface and groundwater irrigation volumes provided in the established USGS Water Use database.

Across all crops we see an overall 20% decrease in SWW and increases in both GWW (3.21% increase) and GWD (2.82% increase). “Other animal feed” is the largest user of all three sources of irrigation water in the CONUS in 2020: 33 surface water, 13 groundwater, and 10 km<sup>3</sup> GWD. Conversely, sweet potatoes use the smallest amount of water across all sources. Soybeans are interesting because, while they use a significant amount of groundwater for irrigation, most of this groundwater use does not contribute to depletion. “Other produce” uses less surface water over the 13-year study period, accompanied by increases in groundwater and GWD, particularly in the Central Valley of California. Contrastingly, other notable crops, such as rice and wheat, decreased their overall total IWU across all three water sources.

Future research could improve upon our work. This study highlights future opportunities to improve agricultural water use modeling across the CONUS. Future efforts to collect data on crop-specific irrigation allocation would be particularly helpful in validating our work. Our study complements the USGS Water Use Database by breaking it down by crop, which can enable future research on water use in agriculture and inform decision-making. We make the complete data set on crop-specific IWU available with the paper, including information for 3,142 counties, 13 years, 20 crops, and 3 unique water sources (nearly 2.5 million data points).

## Data Availability Statement

The data that support the findings of this study are openly available:

- Repository Name: Illinois Data Bank
- DOI: [https://doi.org/10.13012/B2IDB-4607538\\_V1](https://doi.org/10.13012/B2IDB-4607538_V1)
- Access Conditions: Open
- Licensing/Permissions: Creative Commons (CC0)

## Acknowledgments

This material is based upon work supported by the National Science Foundation Grant CBET-1844773 (“CAREER: A National Strategy for a Resilient Food Supply Chain”), DEB-1924309 (“CNH2-L: Feedbacks between Urban Food Security and Rural Agricultural Systems”), BCS-2032065 (“RAPID: Spatial Resilience of Food Production, Supply Chains, and Security to COVID-19”), and CBET-2115405 (“SRS RN: Multiscale RECIPES (Resilient, Equitable, and Circular Innovations with Partnership and Education Synergies) for Sustainable Food Systems”). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or Department of Agriculture. P. J. Ruess acknowledges the Sloan Minority Ph.D. Program for their financial and structural support, as well as our collaborators and others at Utrecht University who facilitated conducting this research on their national supercomputer Cartesius with the help of SURFsara Amsterdam. All data sources are detailed in Table 1 and are publicly available. We gratefully acknowledge these sources, without which this work would not be possible. We appreciate the constructive feedback from two anonymous reviewers that strengthened this paper.

## References

- Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(1), 121–131.
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration—Guidelines for computing crop water requirements—FAO Irrigation and drainage paper 56. Chapter 6: Etc – Single Crop Coefficient (Kc). *FAO, Rome*, 300(9), D05109. Retrieved from <http://www.fao.org/3/x0490e/x0490e00.htm>
- Brouwer, C., Prins, K., & Heibloem, M. (1989). *Irrigation water management: Irrigation scheduling. Annex I: Irrigation efficiencies*. FAO. Retrieved from <http://www.fao.org/3/t7202e/t7202e08.htm>
- Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., et al. (2020). WFD5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, 12, 2097–2120.
- Dalin, C., Konar, M., Hanasaki, N., Rinaldo, A., & Rodriguez-Iturbe, I. (2012). Evolution of the global virtual water trade network. *Proceedings of the National Academy of Sciences of the United States of America*, 109(16), 5989–5994.
- Dalin, C., Wada, Y., Kastner, T., & Puma, M. J. (2017). Groundwater depletion embedded in international food trade. *Nature*, 543(7647), 700–704.
- Dang, Q., Lin, X., & Konar, M. (2015). Agricultural virtual water flows within the United States. *Water Resources Research*, 51(2), 973–986.
- Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., et al. (2018). *Estimated use of water in the United States in 2015* (No. 14411). US Department of the Interior, Geological Survey.
- Famiglietti, J. S., & Rodell, M. (2013). Water in the balance. *Science*, 340(6138), 1300–1301.
- FAO. (2021). FAOSTAT: Food and agriculture association of the United Nations statistical database. Retrieved from <https://www.fao.org/faostat/en/%23data/QCL>
- Gleeson, T., Wada, Y., Bierkens, M. F., & Van Beek, L. P. (2012). Water balance of global aquifers revealed by groundwater footprint. *Nature*, 488(7410), 197–200.
- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., et al. (2010). Food security: The challenge of feeding 9 billion people. *Science*, 327(5967), 812–818.
- Gumdiyala, S., Ruess, P. J., Konar, M., Marston, L., Dalin, C., & Wada, Y. (2020). Groundwater depletion embedded in domestic transfers and international exports of the United States. *Water Resources Research*, 56(2), e2019WR024986.
- Han, W., Yang, Z., Di, L., & Mueller, R. (2012). CropScape: A web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, 84, 111–123.
- Hoekstra, A. Y., Chapagain, A. K., Mekonnen, M. M., & Aldaya, M. M. (2011). *The water footprint assessment manual: Setting the global standard*. Routledge.
- Hutson, S. S., Barber, N. L., Kenny, J. F., Linsey, K. S., Lumia, D. S., & Maupin, M. A. (2004). *Estimated use of water in the United States in 2000* (No. 12683). US Department of the Interior, Geological Survey.
- Jackson, N. D., Konar, M., Debaere, P., & Estes, L. (2019). Probabilistic global maps of crop-specific areas from 1961 to 2014. *Environmental Research Letters*, 14(9), 094023.
- Karakoc, D. B., Wang, J., & Konar, M. (2022). Food flows between counties in the United States from 2007 to 2017. *Environmental Research Letters*, 17(3), 034035.
- Kenny, J. F., Barber, N. L., Hutson, S. S., Linsey, K. S., Lovelace, J. K., & Maupin, M. A. (2009). *Estimated use of water in the United States in 2005* (No. 13444). US Department of the Interior, Geological Survey.
- Lehrsch, G. A., Lentz, R. D., Bjorneberg, D. L., & Sojka, R. E. (2014). Erosion: Irrigation-induced. In *Reference module in earth systems and environmental sciences (online database)* (pp. 1–11).
- Lin, X., Ruess, P. J., Marston, L., & Konar, M. (2019). Food flows between counties in the United States. *Environmental Research Letters*, 14(8), 084011.
- MacKichan, K. A. (1951). *Estimated Use of Water in the United States – 1950* (No. 115). US Department of the Interior, Geological Survey.
- MacKichan, K. A. (1957). *Estimated use of water in the United States, 1955* (No. 398). US Department of the Interior, Geological Survey.
- MacKichan, K. A. (1961). *Estimated use of water in the United States, 1960* (No. 456). US Department of the Interior, Geological Survey.
- Marston, L. T., Abdallah, A. M., Bagstad, K. J., Dickson, K., Glynn, P., Larsen, S. G., et al. (2022). *Water-use data in the United States: Challenges and future directions*. JAWRA Journal of the American Water Resources Association.
- Maupin, M. A., Kenny, J. F., Hutson, S. S., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2014). *Estimated use of water in the United States in 2010* (No. 1405). US Department of the Interior, Geological Survey.
- Mekonnen, M. M., & Hoekstra, A. Y. (2011). The green, blue and grey water footprint of crops and derived crop products. *Hydrology and Earth System Sciences*, 15(5), 1577–1600.
- Murray, C. R. (1968). *Estimated use of water in the United States, 1965* (No. 556). US Department of the Interior, Geological Survey.
- Murray, C. R., & Reeves, E. B. (1972). *Estimated use of water in the United States in 1970* (No. 676). US Department of the Interior, Geological Survey.
- Murray, C. R., & Reeves, E. B. (1977). *Estimated use of water in the United States in 1975* (No. 765). US Department of the Interior, Geological Survey.
- ORNL. (2022). Oak Ridge National Laboratory freight analysis framework. Retrieved from <https://faf.ornl.gov/faf5/>
- Portmann, F. T., Siebert, S., & Döll, P. (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 Biogeochemical Cycles*, 24(1), GB1011.
- Rohwer, J., Gerten, D., & Lucht, W. (2007). *Development of functional irrigation types for improved global crop modelling*. Web.
- Ruess, P. J., & Konar, M. (2019). Grain and virtual water storage capacity in the United States. *Water Resources Research*, 55(5), 3960–3975.
- Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: An analysis of global patterns. *Global Ecology and Biogeography*, 19(5), 607–620.
- Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24), 9320–9325.
- Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Schmied, H. M., Van Beek, L. P., et al. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences of the United States of America*, 115(6), E1080–E1089.
- Siebert, S., Henrich, V., Frenken, K., & Burke, J. (2013). *Update of the digital global map of irrigation areas to version 5*. Rheinische Friedrich-Wilhelms-Universität, Food and Agriculture Organization of the United Nations.

- Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., & Scanlon, B. R. (2015). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, *19*(3), 1521–1545.
- Solley, W. B., Pierce, R. R., & Perlman, H. A. (1993). *Estimated use of water in the United States in 1990* (No. (1081)). US Department of the Interior, Geological Survey.
- Solley, W. B., Pierce, R. R., & Perlman, H. A. (1998). *Estimated use of water in the United States in 1995* (No. (1200)). US Department of the Interior, Geological Survey.
- Solley, W. R., Chase, E. B., & Mann IV, W. B. (1983). *Estimated use of water in the United States in 1980* (No. (1001)). US Department of the Interior, Geological Survey.
- Solley, W. R., Merk, C. F., & Pierce, R. R. (1988). *Estimated use of water in the United States in 1985* (No. (1004)). US Department of the Interior, Geological Survey.
- Sutanudjaja, E. H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H., Drost, N., et al. (2018). PCR-GLOBWB 2: A 5 arcmin global hydrological and water resources model. *Geoscientific Model Development*, *11*(6), 2429–2453.
- USDOT. (2022). *2017 commodity flow Survey standard classification of transported goods (SCTG): SCTG commodity codes*. United States Department of Transportation. Retrieved from: [https://www2.census.gov/programs-surveys/cfs/technical-documentation/code-list/CFS-1200\\_17.pdf](https://www2.census.gov/programs-surveys/cfs/technical-documentation/code-list/CFS-1200_17.pdf)
- USGS. (2021). Water-use data available from USGS. Retrieved from <https://water.usgs.gov/watuse/data/index.html>
- Van Beek, L. P. H., Wada, Y., & Bierkens, M. F. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, *47*(7), W07517.
- Wada, Y., Van Beek, L. P., & Bierkens, M. F. (2012). Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resources Research*, *48*(6), W00L06.
- Wada, Y., Van Beek, L. P. H., & Bierkens, M. F. (2011). Modelling global water stress of the recent past: On the relative importance of trends in water demand and climate variability. *Hydrology and Earth System Sciences*, *15*(12), 3785–3808.
- Wada, Y., Wisser, D., & Bierkens, M. F. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, *5*(1), 15–40.
- Xie, Y., Gibbs, H. K., & Lark, T. J. (2021). Landsat-based irrigation dataset (LANID): 30 m resolution maps of irrigation distribution, frequency, and change for the US, 1997–2017. *Earth System Science Data*, *13*(12), 5689–5710.