New data for representing irrigated agriculture in economy-wide models

Kirby Ledvina, Niven Winchester, Kenneth Strzepek and John M. Reilly

MIT Joint Program on the Science and Policy of Global Change combines cutting-edge scientific research with independent policy analysis to provide a solid foundation for the public and private decisions needed to mitigate and adapt to unavoidable global environmental changes. Being data-driven, the Joint Program uses extensive Earth system and economic data and models to produce quantitative analysis and predictions of the risks of climate change and the challenges of limiting human influence on the environment— essential knowledge for the international dialogue toward a global response to climate change.

To this end, the Joint Program brings together an interdisciplinary group from two established MIT research centers: the **Center for Global Change Science (CGCS)** and the **Center for Energy and Environmental Policy Research (CEEPR)**. These two centers—along with collaborators from the Marine Biology Laboratory (MBL) at Woods Hole and short- and long-term visitors—provide the united vision needed to solve global challenges.

At the heart of much of the program's work lies MIT's Integrated Global System Model. Through this integrated model, the program seeks to discover new interactions among natural and human climate system components; objectively assess uncertainty in economic and climate projections; critically and quantitatively analyze environmental management and policy proposals; understand complex connections among the many forces that will shape our future; and improve methods to model, monitor and verify greenhouse gas emissions and climatic impacts.

This reprint is intended to communicate research results and improve public understanding of global environment and energy challenges, thereby contributing to informed debate about climate change and the economic and social implications of policy alternatives.

—Ronald G. Prinn and John M. Reilly,
Joint Program Co-Directors

New data for representing irrigated agriculture in economy-wide models

Kirby Ledvina, Niven Winchester, Kenneth Strzepek and John M. Reilly

Abstract: We develop a framework to represent the value of irrigated crop production and the expansion potential of irrigated land within economy-wide models, providing integrated assessment capabilities for energy, land, and water interactions. Specifically, we compute the value of production on irrigated and rainfed cropland at both a 5 arcminute by 5 arcminute level (about 10 square kilometers) and for the 140 regions and eight crop sectors in Version 9 of the Global Trade Analysis Project (GTAP) Data Base. For each crop category, we estimate the shares of production on irrigated and rainfed land using estimates of production quantities and prices, compared to approximations based on output volumes used in the GTAP-Water Data Base. We construct a global dataset of evaluation metrics to identify region-crop combinations where there are large differences in irrigated production value shares based on direct calculation and approximated by output volumes. The scope to expand the amount of irrigated land and the cost of doing so is quantified through irrigable land supply curves for 126 water regions globally, based on water availability and the costs of irrigation infrastructure. We also make available our adaptable work stream to calculate crop production values and to estimate irrigable land supply elasticities for use in economy-wide models. Altogether, this work can enhance integrated assessment and economy-wide modeling by more accurately capturing the value of crop production and facilitating the representation of endogenous investment in irrigation infrastructure in response to changing water availability. These data and modeling contributions allow for a more rigorous exploration of the regional and global impacts of water availability on land use, energy production, and economic activity.

1. Introduction

An expanding world population and global economy is expected to increase food demand and place pressure on current food crop production (Gurgel *et al.*, 2007; OECD/ FAO, 2017). Additionally, amidst a changing climate, new energy and climate policies may be proposed to support bioenergy production as an alternative to conventional fossil-based methods, placing food and energy production in direct competition for land resources (and Azar, 2007; Wise *et al.*, 2014; Winchester and Reilly, 2015). One way to accommodate a growing demand for both food and bioenergy is to intensify existing crop land by increasing crop yields through investments in irrigation technology (Taheripour *et al.* 2016). However, to explore the potential impacts of intensification on food prices and bioenergy production, we need to understand the physical and cost constraints on irrigable land expansion. How much additional land can be irrigated, in which parts of the world, and at what cost? While we can use applied general equilibrium (AGE) modeling techniques to investigate these questions, we first need the ability to explicitly represent irrigated land and its expansion potential within economy-wide models, a capability that is the main focus of this paper.

Early work in the area includes Taheripour *et al.* (2013a), Taheripour *et al.* (2013b), Liu *et al.* (2014), and Liu *et al.* (2016). Subsequently, Haqiqi *et al.* (2016) enhanced the Global Trade Analysis Project (GTAP) Power Data Base (Peters, 2016)—an augmentation of Version 9 of the GTAP Data Base (Aguiar *et al.*, 2016) that represents electricity generation in detail—to separately represent rainfed and irrigated agricultural production and form the GTAP-Water Data Base. To disaggregate crop production value into rainfed and irrigated shares, they use data on irrigated land's contribution to output quantity under the simplifying assumption that, for each GTAP crop category, the share of total production (in tonnes) on irrigated land is equal to the share of total output value on irrigated land. While this assumption seems reasonable for GTAP crop categories that consist of a single crop, it may not be valid for crop categories that include heterogeneous crops. Specifically, for heterogeneous crop categories, the production shares on irrigated and rainfed land may differ from output value shares if high-value crops are grown on irrigated land and low-value crops on rainfed land, or *vice versa*. Though Haqiqi *et al.* (2016) note this possible issue, they do not explore the specific crops and regions for which differences may arise.

Winchester *et al.* (2016) build on Haqiqi *et al.*'s (2016) work by using irrigated and rainfed value shares calculated with a global price dataset rather than output shares. While the output shares used by Haqiqi *et al.* (2016) are more attainable on a global scale, the value shares esti-mated by Winchester *et al.* (2016) bypass the implicit assumption that, in each GTAP crop category, the proportional cultivation of constituent crops on irrigated land is equal to that on rainfed land. Beyond estimating the current extent of production, Winchester *et al.* (2016) advance current modeling practices by enabling the endogenous expansion of irrigated land within the Economic Projection and Policy Analysis (EPPA) model (Chen *et al.*, 2017) through water region-specific irrigable land supply curves, which quantify the amount of irrigated land gained from investments in irrigation systems and water storage. Prior to Winchester *et al.* (2016), endogenous responses in irrigation infrastructure had not been considered in the AGE literature. The development of irrigable land supply curves to incorporate these responses within AGE models provides a mechanism for regions to expand production on irrigated land, subject to water constraints and infrastructure costs, and allows for more robust exploration of the effects of a carbon policy and water constraints on economic performance, biomass production, land use, and greenhouse gas emissions.

While Haqiqi *et al.* (2016) and Winchester *et al.* (2016) explain their findings in detail, the purpose of this paper is to elucidate and advance the development of the irrigated land framework itself. Specifically, we (1) assist the representation of irrigated land in the GTAP Data Base by improving estimates of production values shares on rainfed and irrigated land, and (2) provide estimates of the scope to increase irrigated land at a finer level of spatial aggregation than Winchester *et al.* (2016). Both items are tools that the modeling community may find useful, and we make available as supplementary materials our complete work stream. Specific materials include data on (1) irrigated and rainfed land area by GTAP region and crop sector, (2) the directly calculated value of irrigated and rainfed production by GTAP region and crop sector, (3) irrigable land supply functions for 126 water regions, (4) irrigable land supply curves for the irrigation response units (IRUs) used in Winchester *et al.* (2016), and (5) source code for all aggregation routines.[1] The supplementary files allow aggregation of rainfed and irrigated land area, output volume, and production value from a finer level of spatial resolution to user-defined regions other than the GTAP regions. We also analyze the approaches of Winchester *et al.* (2016) and Haqiqi *et al.* (2016) in estimating production value, and we flag regions and crop sectors where it may not be suitable to equate production shares with value shares.

---

1   Haqiqi *et al.* (2016) also provide irrigated and rainfed crop areas by GTAP region and crop, as well as the GAMS code to generate irrigated and rainfed production values calculated from output shares.

The remainder of this paper describes the irrigated land framework and tools used. Section 2 discusses the valuation of irrigated and rainfed crop production and evaluates the methodologies of Winchester *et al.* (2016) and Haqiqi *et al.* (2016). Section 3 details the construction of irrigable land supply curves. Section 4 summarizes the code and other tools we make available. Section 5 concludes.

2. Production on Irrigated Land and Rainfed Land

As AGE models are calibrated using transaction values, a key requirement in separately representing irrigated and rainfed crop production is the disaggregation of aggregate crop production into production values on irrigated land and rainfed land. As noted above, Haqiqi *et al.* (2016) have incorporated an irrigated land framework into the GTAP-Power Data Base to form the GTAP-Water Data Base. They estimate irrigated and rainfed production value by first identifying the volume of production on irrigated and rainfed land and splitting the aggregate value of production in a GTAP crop sector according to each land type's share of crop output. Winchester *et al.* (2016) extend the methodology used by Haqiqi *et al.* (2016) by applying crop and region-specific prices to rainfed and irrigated crop production. To evaluate the substitutability of the two methods to estimate production value shares, we follow Winchester *et al.* (2016) in incorporating crop prices, and compare the resulting production values to estimates based on output shares. Within a region, production value shares generally equal the output shares for individual crop sectors (eg. wheat) but can diverge

for composite crop sectors (eg. coarse grains), because of differences in individual crop prices and the mix of composite crops grown on the two land types. We make available rainfed and irrigated production values for 26 disaggregated crops and 282 river basins; as well as for the 140 GTAP regions, 19 aggregated regions, and eight crop sectors considered in the GTAP-Water Data Base.

2.1 Direct Calculation of Production Value

Following Haqiqi *et al.* (2016), we use spatial datasets on harvested areas and crop yields to calculate production volumes, which we combine with country-level data on crop prices to estimate crop production value on rainfed and irrigated land, as in Winchester *et al.* (2016). We obtain harvested areas for rainfed and irrigated land from the Monthly Irrigated and Rainfed Crop Areas (MIRCA2000) dataset (Portmann *et al.*, 2010). Areas are reported for the year 2000 and are available globally for 26 crops and 2 land types at a 5 arc-minute by 5 arc-minute spatial resolution (about 0.083 square degrees, or 10 square km at the equator).[2] **Figure 1** provides a snapshot of the grid-level MIRCA2000 dataset for barley with relative amounts of irrigated and rainfed crop areas overlaid on a map of GTAP 9 regions.

The high resolution of the MIRCA2000 dataset allows for flexibility in how the harvested areas are aggregated, whether regionally or globally. Because the GTAP databases are ubiquitous in economic modeling, for illustration purposes, we calculate harvested area for the 140 regions and 8 crop sectors represented in the GTAP 9 Data

---

2   The full spatial dataset of harvested areas is included in the **1_SpatialData** subfolder of the supplementary materials.
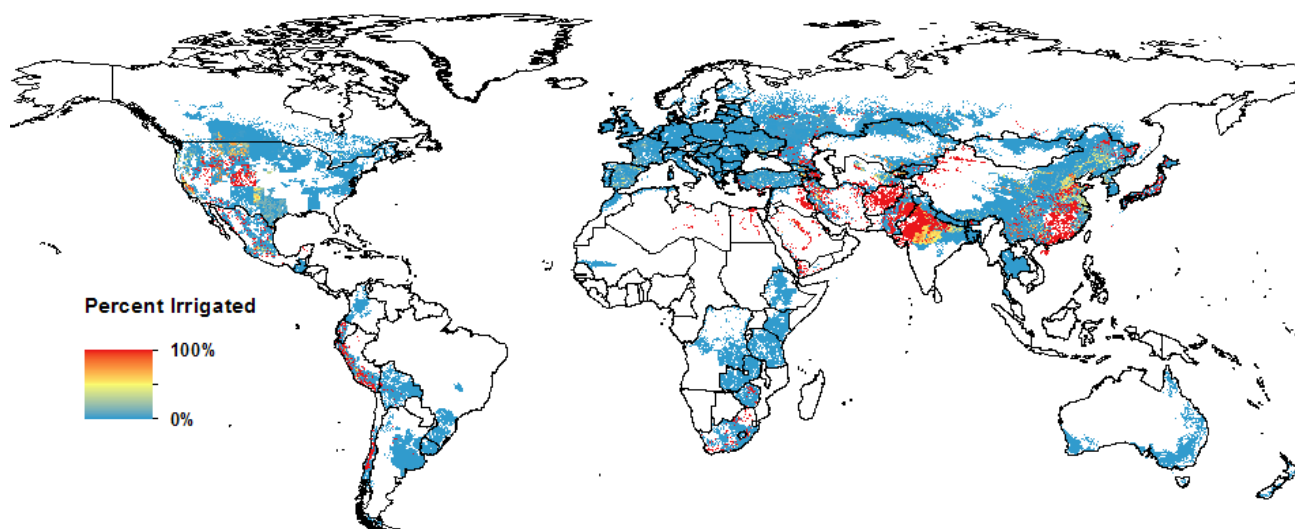

Figure 1. Percentage of harvested barley in each grid cell coming from irrigated production. White areas have no barley production. Source: Authors' aggregation of data from Portmann *et al.* (2010).

Base. **Figure 2** summarizes the global harvested area of each GTAP crop by land type. Globally, seventy-six percent of cropland is rainfed, with paddy rice (*pdr*) the only crop primarily grown on irrigated land.

Data on rainfed and irrigated crop yields are sourced from Siebert and Döll (2010) and are available at a 5 arc-minute by 5 arc-minute spatial resolution for 29 crops.[3] These 29 crops include the original 26 crops covered by the MIRCA2000 dataset but with (1) three types of animal feed rather than a single fodder category, and (2) the addition of pasture land, which we exclude from our analysis. We take a simple average of fodder from maize, fodder from barley, and fodder from wheat yields to approximate a single fodder yield at the grid cell-level for each land type. Yields of 0 tonnes per hectare (t/ha) for a specific fodder crop are reclassified as missing data and excluded from the average for that particular grid cell. We calculate tonnes of production by land type and crop as the product of the harvested area and yield in each grid cell.

To determine the dollar value of rainfed and irrigated crop production, we follow Winchester *et al.* (2016) by applying crop prices obtained from a Food and Agricultural Organization (FAO) database (FAO, 2015c) to

the spatial production data. The FAO dataset consists of country-specific prices from the year 2000 for 215 crops and does not differentiate between crops grown on different land types.[4] In other words, crops grown on rainfed land and irrigated land within the same country are assumed to sell for the same price.

We create a raster dataset of prices that can be applied to production output at the grid cell level by (1) consolidating the 215 FAO crops into the 26 MIRCA2000 crop categories, (2) interpolating missing country-level price data, and then (3) pixelating the country-level data to create a raster dataset.[5] We complete the first adjustment by assigning each FAO crop to one of the 26 MIRCA2000 crop categories and using FAO production data (FAO, 2015a) to compute a production-weighted price within each crop category.[6] The specific crop mapping employed is specified in **Table A1**. We use a listing generated by the Natural Resources Conservation Service (NRCS) of the U.S. Department of Agriculture (USDA) to identify

---

3   The full spatial dataset of yields is included in the **1_SpatialData** subfolder of the supplementary materials.

4   Price data is provided in FAOprices.xlsx in the **1_SpatialData\Prices** subfolder of the supplementary materials.

5   See the **1_SpatialData\Prices** subfolder in the supplementary materials for the spatial price dataset. Instructions to update and regenerate the dataset are provided in the README file of the supplementary materials.

6   Production-weighted prices are calculated in **Prices.gms** in the **1_SpatialData\Prices** subfolder of the supplementary materials.
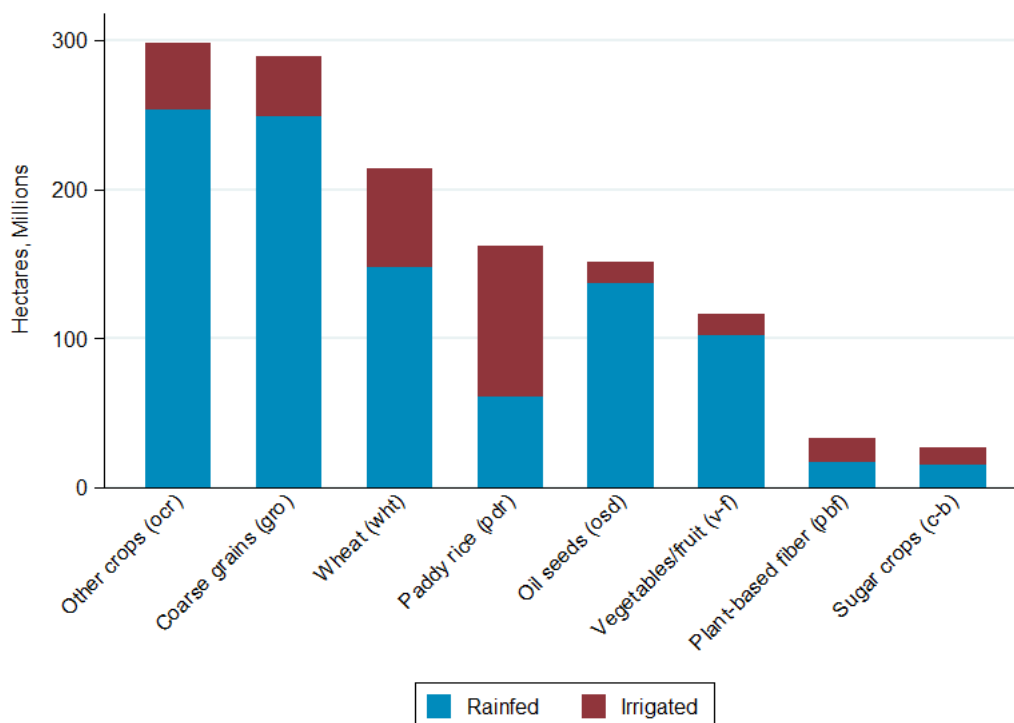

Figure 2. Global rainfed and irrigated harvested area by GTAP crop. Source: Authors' aggregation using data from Portmann *et al.* (2010). Crop mapping from Haqiqi *et al.* (2016).

FAO crops classified as 'other annual' or 'other perennial' (NRCS, 2014).

We interpolate missing prices by separately handling (1) countries with no crop price data and (2) nations with some but not complete crop price data. The FAO provides no price information for 32 countries, so we substitute known prices from a geographically proximate country. **Table A2** lists the pairings of countries missing price data and those selected to provide proxy prices. Country pairing assignments are determined on geographic basis (e.g. Dominican Republic and Haiti, and Sudan and South Sudan). Although alternative assignments are unlikely to have a large impact at the aggregate level,[7] production values for countries with missing price data should be used with caution.

For cases where FAO (2015c) is missing prices for a subset of crops in a country, we estimate global prices of the 26 MIRCA2000 crops, with the understanding that several countries may have no production of the specific crop and will not affect aggregate value. We develop these missing country-level prices for each MIRCA2000 "target" crop (the crop missing price data) by using price ratios based on data from countries with known prices. Specifically, we use the ratio of the target crop's price to the price of several candidate "guide" crops—barley, citrus, maize, wheat, rice, potatoes, and groundnuts, or some combination of these crops—in a country with known prices for both the target and guide crops. The specific crop combination selected is based on the ratios with the lowest variance across countries.

For example, the FAO reports that Paraguay produced 82 thousand tonnes of sunflower seed in the year 2000, but FAO (2015c) does not report a producer price for this crop in 2000. To estimate the price of sunflower in Paraguay, for the countries with available price data, we compute a ratio of sunflower price to each guide crop's price and calculate the mean and variance for each ratio across countries. Because of their low variances, the groundnut, rice, maize, and potato ratios are used to generate four country-level sunflower price estimates,[8] which are averaged to produce a single price estimate. **Table 1** presents the ratio statistics for each guide crop and the calculations leading to a final sunflower price estimate for Paraguay of $249.10/t, which can be compared to $257.92/t, the simple average of known FAO sunflower prices across countries. An Excel workbook with this methodology and calculations for all crops and countries is included in the supplementary materials.[9]

Finally, we generate a complete spatial dataset of prices for GTAP regions by (1) assigning each GTAP region an average price weighted by country area, and then (2) pixelating the regional price data to form grid-level spatial data. We calculate production value at the grid cell level by multiplying output by crop prices.[10] Determining production values at the grid cell level allows for flexibility in selecting a regional aggregation. For illustration purposes, **Figure 3** depicts our estimates of global production value by land type and crop when the estimates are aggregated to the eight GTAP crops. Despite having approximately three times the harvested area as irrigated crops, rainfed crops generate only twice the value.

---

7   Countries with missing prices provide 3.6% of global crop production, according to FAO (2015a) data.

8   For each country, an estimate is generated only if a price for the guide crop exists.

9   See **CropPriceInterpolation.xlsx** in the **1_SpatialData\Prices** subfolder of the supplementary materials.

10   See **GenerateRasters.py** in the **1_SpatialData** subfolder of the supplementary materials for the code to create the price, output, and value spatial datasets. See **Aggregate.py** in **2_ProductionValue** for the code to aggregate the spatial data.

Table 1. Sunflower Price Estimation in Paraguay

| Guide Crop	(1) Mean of Country Ratios	(2) Variance of Country Ratios	(3) Price in Paraguay *USD per t*	Sunflower Price Estimate (1)×(3) *USD per t*
Groundnuts/peanuts	0.485	*0.036*	487.60	236.68
Rice	1.161	*0.265*	112.40	130.49
Maize	1.705	*0.310*	137.70	234.83
Potatoes	1.375	*0.485*	286.80	394.43
Citrus	1.338	*0.956*	104.37	139.64
Wheat	1.731	*1.134*	121.90	211.02
Barley	1.988	*1.226*	No Data	No Data
Averaged Estimate (Outlined Cells Only)				**249.10**
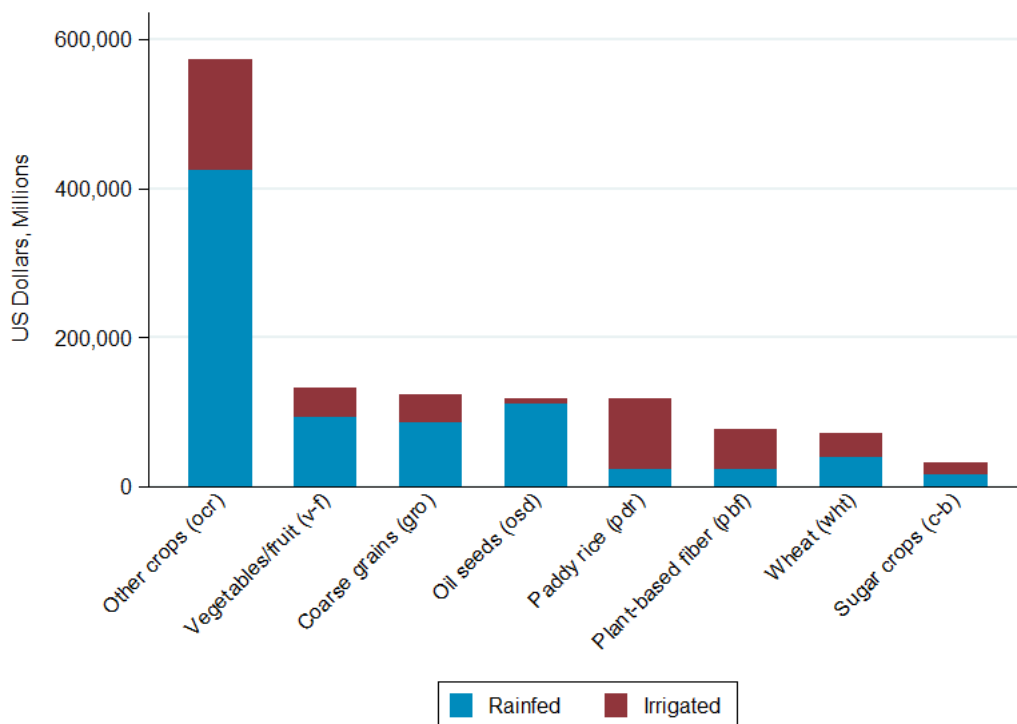
Figure 3. Global production value on rainfed and irrigated land by GTAP crop. Source: Authors' calculation using data from Portmann *et al.* (2010), Siebert and Döll (2010), and FAO (2015a, 2015c). Crop mapping from Haqiqi *et al.* (2016).

Table 2. Production values under alternative methods are distinguished by color. Source: Authors' own calculations.

	Crop Sector	(1) From Output Share		(2) Direct Calculation		Average Ratio of Values (1)/(2)
		Irrigated Value *Million USD*	Irrigated Value Share	Irrigated Value *Million USD*	Irrigated Value Share	
osd	Oilseeds	7,098	6.0%	7,724	6.5%	1.10
v-f	Vegetables & Fruit	42,355	32.0%	40,164	30.4%	1.09
gro	Coarse Grains	37,504	30.2%	38,119	30.7%	1.02
ocr	Other crops	130,265	22.8%	147,761	25.8%	1.01
wht	Wheat	30,908	43.9%	30,915	43.9%	1.01
c-b	Sugar crops	16,334	52.2%	16,336	52.2%	1.00
pdr	Paddy rice	95,626	81.0%	95,613	81.0%	0.99
pfb	Plant-based Fiber	52,878	69.6%	52,904	69.7%	0.98
	ALL CROPS	**412,967**	**33.2%**	**429,536**	**34.6%**	**1.02**

2.2 Comparison of Production Methodologies

Using the estimates detailed in Section 2.1., we compare production value shares calculated by multiplying output volumes by prices, as in Winchester *et al.* (2016), and by assuming that production value shares are equal to output volume shares, as in Haqiqi *et al.* (2016). The goal of this comparison is to gain insight into where and for which crops the simplifying assumption used by Haqiqi *et al.* (2016) may or may not be valid.

Table 2 summarizes global production values calculated under the two methods. While results for total crop production are comparable, estimates from the two methods vary greatly when assessed by crop sector. This is because value shares estimated from output shares may not capture differences between production quantities and values within composite GTAP crop sectors, i.e. GTAP sectors containing multiple MIRCA2000 crops—oilseeds (*osd*), vegetables and fruit (*v-f*), coarse grains (*gro*), sugar crops (*c-b*), and other crops (*ocr*). However, three of the

GTAP crop sectors—wheat (*wht*), paddy rice (*pdr*), and plant-based fiber (*pfb*)—consist of a single MIRCA2000 crop—wheat, rice, and cotton, respectively—and therefore should have identical estimates from output shares and from direct calculation of production values (prices do not differentiate between rainfed and irrigated crops). Deviations from a value ratio of 1.0 in these homogenous crop sectors can be attributed to shapefile variations in the price and production datasets.

We compare production value estimates by GTAP region and crop sector to flag cases where the two methods yield significantly different results. Containing separate graphs for the five composite GTAP crop sectors, **Figure 4** compares irrigated land's production value share with its
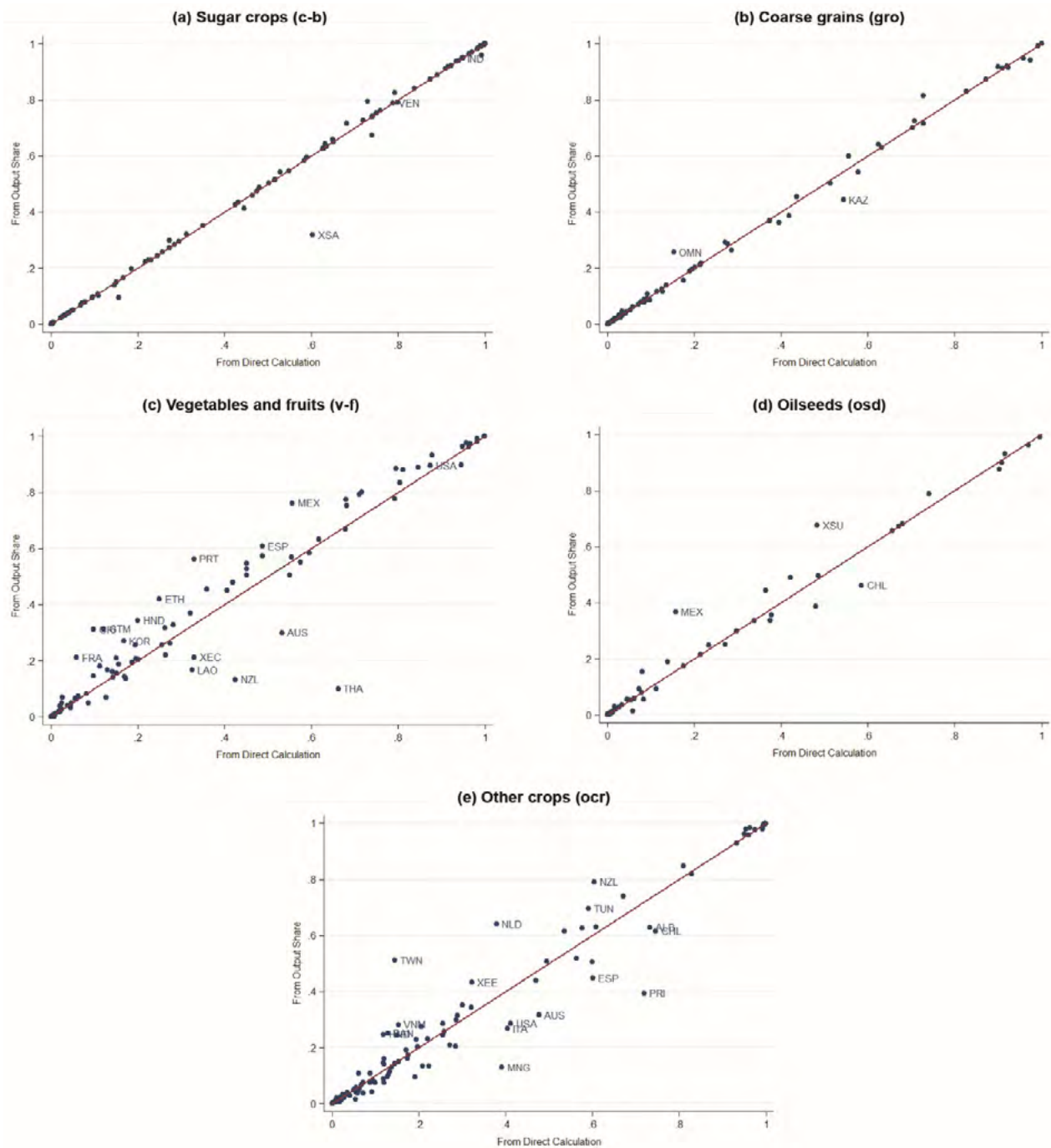
Figure 4. Irrigated share of production value estimated from the output share (y-axis) versus direct calculation (x-axis) for each GTAP crop sector. Points represent GTAP regions. The red line marks 45 degrees. Source: Authors' calculation using data from Portmann *et al.* (2010), Siebert and Döll (2010), and FAO (2015a, 2015c). Crop mapping from Haqiqi *et al.* (2016).

output value share for each GTAP region.[11] For a given crop in a particular region, the production value share on irrigated land will equal the corresponding output value share if (1) the production share of each MIRCA2000 crop within a composite GTAP sector on irrigated land is equal to that on rainfed land, as noted by Haqiqi *et al.* (2016); or/and (2) the prices per tonne for each MIRCA2000 crop within a composite GTAP sector are equal.

The two methods produce comparable estimates within the sugar crop sector (*c-b*). The *c-b* sector contains sugarcane and sugar beet, with Venezuela and India being among the countries with greater than 75% irrigated sugar crop production. In Venezuela, both sugarcane and sugar beet sell for about $223/t, so the crops' irrigated output shares are approximately equal to their irrigated value shares. In India, nearly all of the *c-b* production is sugarcane rather than sugar beet, so sugarcane's contributions to rainfed and irrigated *c-b* production approach 100% while sugar beet's contributions to irrigated and rainfed production approach 0%. A heavy focus on one sugar crop or the other is characteristic of many of the GTAP regions, and therefore the output share often provides a reasonable estimate for value share.

The vegetables and fruit sector (*v-f*) shows more regional variation in the value comparison. The United States, the largest irrigated *v-f* producer by both output and value, falls along the 45-degree line, indicating similar value estimates from the two calculation methods. In this region, the most produced *v-f* crops, citrus and potatoes, share a similar price—$108/t for citrus and $112/t for potatoes—while the third most produced crop, grapes/vines, has a higher price but contributes in similar proportions to USA irrigated and rainfed *v-f* production. In contrast, Thailand, which produces 90% of its *v-f* output on rainfed land, features a higher irrigated *v-f* value from direct calculation than from an output share estimation. This difference in value arises because rainfed cassava contributes to 88% of the Thailand's total *v-f* production but, with a price of $16/t, makes up just 22% of total value. In contrast, irrigated citrus contributes only 10% of the total *v-f* output in Thailand yet generates 66% of *v-f* value with its price of $419/t. Because highly cultivated rainfed crops generate less value than the irrigated crops, dividing total value into irrigated and rainfed parts based on output share is not a suitable method for the *v-f* sector in Thailand.

To more formally evaluate the appropriateness of using output shares to approximate value shares, we calculate a

relative error metric, $k_{c,r}$, for each GTAP crop category, c, and region, r, given by

$$k_{c,r} = \frac{p_{c,r} - v_{c,r}}{0.5 * (p_{c,r} + v_{c,r})} \qquad (1)$$

where $p_{c,r}$ and $v_{c,r}$ are, respectively, the production value computed from output share and from direct calculation for crop c on irrigated land in region r. Relative error values for all GTAP crop-region combinations are provided in the supplementary materials included with this paper.[12] To focus on crop-region combinations where the assumption that the production share equals the value shares is most likely to be violated, **Table 3** lists the region-crop combinations with the highest absolute relative errors. Across all region-crop combinations, relative error values range from -1.48 (*v-f* in Thailand) to 1.13 (*v-f* in France) and yield an average absolute relative error of 0.09. Of the region-crop combinations with a non-zero relative error, fifty-four percent have a positive relative error, suggesting that use of the production share simplification tends to overestimate the value of production on irrigated land. For crop-country combinations with high absolute relative errors, we encourage the use of the tools and data detailed in this paper to adjust the irrigated land-augmented version of the GTAP-Power Data Base provided by Haqiqi *et al.* (2016).

3. Representation of Irrigable Land Supply Curves

The development of irrigable land supply curves enables regions to adapt to changes in water resources and agri-

---

12   See **EvaluationMetric.xlsx** in the **2_ProductionValue** subfolder. Related calculations are in **GTAPreporting.gms**.

Table 3. Largest 10 Irrigated Relative Errors

	Region	Crop	Value share	Absolute relative error
THA	Thailand	*v-f*	10.7%	1.48
MOZ	Mozambique	*v-f*	28.7%	1.44
IDN	Indonesia	*v-f*	2.4%	1.42
AZE	Azerbaijan	*osd*	0.3%	1.24
FRA	France	*v-f*	27.7%	1.13
GBR	United Kingdom	*ocr*	69.0%	1.13
TWN	Taiwan	*ocr*	71.0%	1.12
IDN	Indonesia	*osd*	72.7%	1.10
MYS	Malaysia	*v-f*	0.3%	1.09
NZL	New Zealand	*v-f*	2.7%	1.05

ᵃ Crop sector's share of the region's total crop value

---

11   Output value shares and production value shares for each GTAP crop-region combination are provided in **ValueSummary.xlsx** in **2_ProductionValue** subfolder of the supplementary materials.

culture demand by investing in irrigation infrastructure and intensifying crop production. To account for variations in water resources across river basins, we use 282 river basins defined by the Integrated Global Assessment Model—Water Resource System (IGSM-WRS) framework (Strzepek *et al.* 2012). Because river basins are in part delineated by political borders, we define 126 water regions as aggregations of adjacent river basins that can cross country lines to better capture the transnational aspect of water resource systems. **Figure 5** maps the river basins and their water region aggregations.

3.1 Irrigation Efficiency

Winchester *et al.* (2016) use irrigation system efficiency (SEF) values from the International Food Policy Research Institute's (IFPRI) International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT) model (Rosegrant *et al.* 2012) to characterize the current extent of irrigation in each of the 282 river basins. The SEFs of the 126 water regions are the average of constituent basin efficiencies, weighted by the area of irrigated land. Based on FAO irrigation and drainage data (FAO, 2015b) and expert knowledge, Winchester *et al.* (2016) split the water region SEF values into two separate efficiency metrics for the base year: conveyance efficiency and field efficiency. The conveyance efficiency is determined by the amount of water lost to seepage and/or evaporation within a system of canals. Field efficiency—referring to the portion of the water released on the fields that ultimately waters the crops—depends on the

type of irrigation system and increases with more targeted methods.

To allow for irrigation upgrades, we consider improvements in conveyance efficiency through the addition of canal lining. Conveyance efficiency for unlined canals is 75% (i.e., 75% of water released from dams reaches the field) and increases to 95% after the addition of lining. Additionally, each water region can improve its field efficiency beyond its base year efficiency through four possible system upgrades at increasing costs: flood, furrow, low-efficiency sprinkler, and high-efficiency sprinkler. The "representative" irrigation technology currently in use is selected according to field efficiency in the base year, as summarized in **Table 4**. In each water region, additional investments in field irrigation systems can be used to improve the field efficiency of irrigation systems. **Table 5** lists the efficiency values associated with each possible irrigation upgrade, as well as the new SEF incorporating the updated field efficiency.

To determine the quantity of irrigated land gained with each system upgrade, Winchester *et al.* (2016) consider the sector water requirement (SWR) of irrigation, i.e. the water withdrawal required to meet irrigation demands of the current crop mix in each water region. A water region's SWR for irrigation is defined as irrigated crop consumptive use across the water region divided by the region's SEF—this relationship highlights that because of transport inefficiencies, the amount of water allocated to irrigation exceeds the amount consumed by irrigated
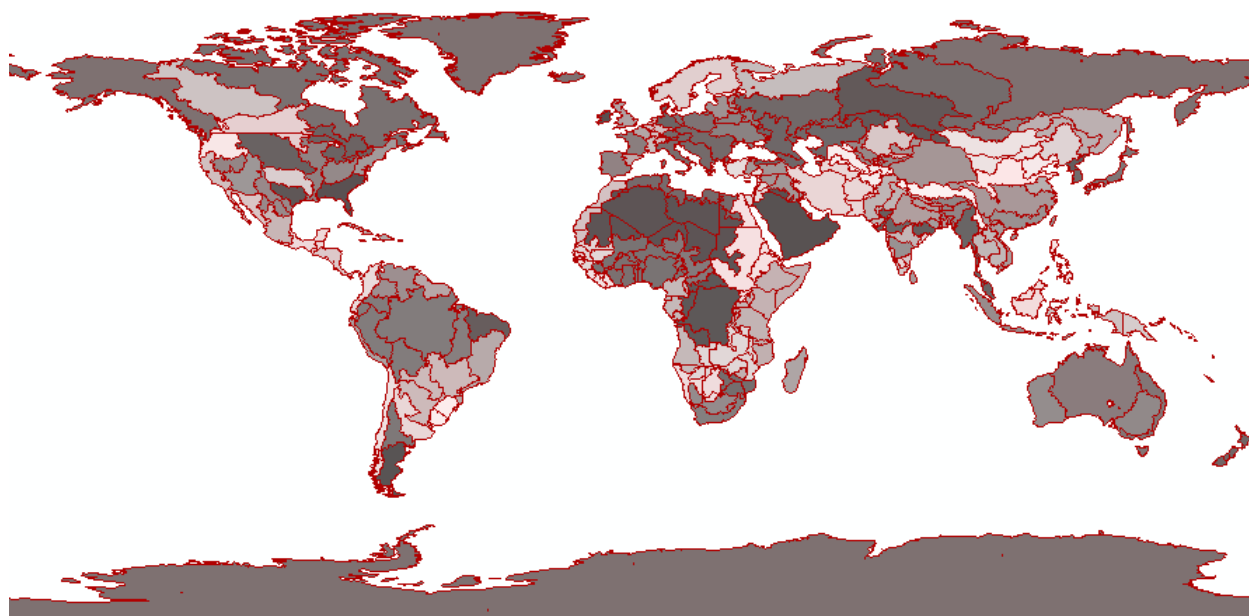

Figure 5. River basins (outlined) and their water region aggregations (shaded). Shapefile is provided in the **1_SpatialData\SHP** subfolder of the supplementary materials. Source: International Food Policy Research Institute (IFPRI) IMPACT Model (Rosegrant *et al.* 2012)

Table 4. Field Efficiencies and Corresponding Irrigation Technology in Base Year

Field Efficiency Range[a]		Corresponding Technology
Low	High	
0	0.35	None
0.35	0.55	Flood
0.55	0.75	Furrow
0.75	0.85	Low-eff. sprinkler
0.85	0.90	High-eff. sprinkler

[a] Field efficiency is backed out from values in IFPRI's IMPACT model (Rosegrant *et al.* 2012)

Table 5. Efficiency Values from Irrigation Upgrades

Upgraded Field Technology	(1) Field Efficiency	(2) Conveyance Efficiency[a]	SEF (1)*(2)
Flood	0.45	0.95	0.43
Furrow	0.65	0.95	0.62
Low-eff. sprinkler	0.80	0.95	0.76
High-eff. sprinkler	0.88	0.95	0.83

[a] Conveyance efficiency assumes lined canals. Efficiency values are from IFPRI's IMPACT model (Rosegrant *et al.* 2012)

crops. Crop consumptive use is estimated for each water region using CliCrop (Fant *et al.*, 2012)—a biophysical crop model that considers temperature, precipitation, and potential evapotranspiration—integrated with IG-SM-WRS. All together, we use the data on irrigation efficiency to calculate the water saved from an irrigation upgrade, and then, knowing the amount of water needed by the current crop mix, we determine how much additional land can be irrigated from the surplus water.

Winchester *et al.* (2016) use IGSM-WRS to compute the additional areas that can be irrigated from investments in irrigation systems. Of the possible improvements to conveyance and field efficiency, the addition of canal lining allows for the greatest increase in irrigable land at the lowest cost per hectare and is therefore the first irrigation system upgrade for all water regions, followed by upgrades in field technology in order of increasing field efficiency. Some irrigation upgrades in some water regions, typically high-efficiency sprinkler, are assumed to be infeasible because the resulting irrigated land expansion would exceed the quantity of available rainfed land.

Some water regions require manual adjustments to the estimates of additional hectares that can be irrigated. Specifically, the IGSM-WRS model overstates the irrigated land gained from the addition of canal lining in large rice-producing water regions in China. This overestimate is a result of rice cultivation methods. Because rice paddies are grown in flooded fields, water that leaks out of unlined canals contributes to crop irrigation. Consequentially, the addition of lining does not substantially improve irrigation efficiency, and the amount of irrigable land gained according to IGSM-WRS is overstated. Winchester *et al.* (2016) address this issue by setting the amount of additional land that can be irrigated due to the addition of canal lining equal to one-tenth of the estimated amount in eight water regions in China.[13]

---

13   These regions are SE_Asia_Coast, Chang_Jiang, Hail_He, Hual_ He Langcang_Jiang, Songhua, Yili_He, and Zhu_Jiang water regions.

3.2 Increases in Water Storage

Beyond improvements in irrigation efficiency, irrigable land in a water region can expand through investments in water storage. Each region's water storage capacity is modeled through a storage-yield curve relating available water supply to the quantity of storage. A region's storage-yield curve extends from no storage, which creates no additional irrigated land, to the amount needed to accommodate the region's mean annual runoff. Mean annual runoff is based on estimates of surface and subsurface runoff generated by the Community Land Model (CLM) (Bonan *et al.*, 2002) within the IGSM framework (Sokolov *et al.*, 2005). In modeling surface runoff, CLM considers the effect of soil infiltration limits, runoff from saturated surface conditions, frozen soil, and root density on soil hydraulic conductivity. For subsurface runoff, CLM employs within each water region a representation of an unconfined aquifer, as opposed to the alternative artesian aquifer, which restricts water from entering through the top and bottom of the reservoir. The unconfined system is employed because it is less complex to model and because there is not yet academic consensus on the extent of each type of aquifer storage. The full methodology in modeling storage capacity is detailed by Strzepek *et al.* (2013).

Winchester *et al.* (2016) divide each water region's storage-yield curve into ten possible upgrades of equal capacity but increasing marginal cost, with the assumption that lower cost upgrades are the first adopted. In each region, the storage-yield curve is combined with an estimate of current storage to calculate the scope and cost of increasing annual water yields beyond its current level. Estimates of current storage in each region are sourced from the IMPACT model, which uses data from an online, global database of large dams managed by the International Commission on Large Dams (ICOLD). To determine the cost of additional storage, Winchester *et al.* (2016) use IGSM-WRS, which takes water runoff data as an input, to determine both the marginal cost and increase in yield
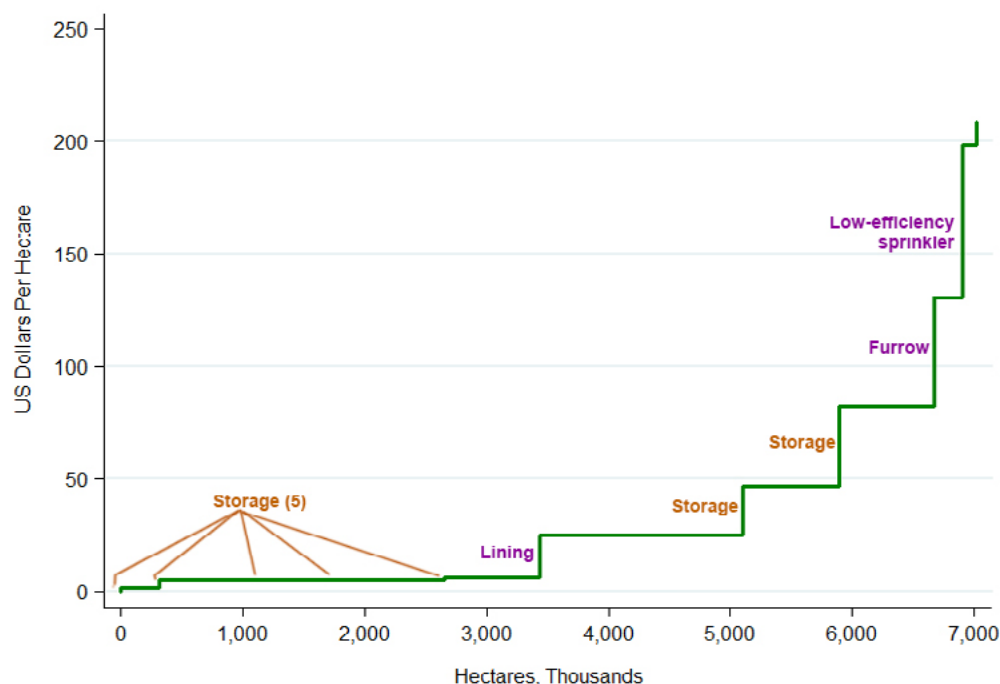
Figure 6. Irrigable land supply curve for Mississippi River (MIS) water region. Source: Authors' own calculations.

Table 6. Additional land that can be irrigated from upgrades in Mississippi River (MIS) water region

Upgrade number	1	2	3	4	5	6	7	8	9	10
Upgrade type	Storage	Storage	Storage	Storage	Storage	Lining	Storage	Storage	Furrow	Low-Eff. Sprinkler
Upgrade cost *USD/ha*	1.62	5.30	5.30	5.30	6.46	25.36	46.73	82.67	130.85	198.75
Add. irrig. land from upgrade *1,000 ha*	313	781	781	781	781	1,666	781	781	234	113
Total add. irrig. land *1,000 ha*	313	1,095	1,876	2,658	3,439	5,106	5,887	6,669	6,902	7,015

Note: Water supply limitations preclude investment beyond the tenth upgrade.

volume from the capacity upgrade. They assume that additional storage upgrades are not adopted once capacity can fully accommodate the mean annual runoff.

3.3 Constructing Supply Curves for Additional Irrigated Land

Winchester *et al*. (2016) arrange the five possible irrigation efficiency upgrades and ten possible storage upgrades into supply curves for additional irrigable land for each water region. The supply curves convey the number of additional hectares that could be irrigated and the one-time cost per hectare of the infrastructure upgrades. As an example, the irrigable land supply step function for the Mississippi River (MIS) water region is illustrated in **Figure 6**, with corresponding supply data provided in **Table 6**. In this example, the irrigable land supply function includes seven incremental storage upgrades

before maximum storage is reached, which indicates that existing storage is equivalent to the sum of the first three incremental storage upgrades. As with all of the water regions, the addition of lining is the lowest cost efficiency upgrade, and because the next irrigation system upgrade is to a furrow system, it can be inferred that flood irrigation is used as the representative system for this water region. In total, investment in irrigation and storage infrastructure in the MIS water region could irrigate an additional 7.0 million ha beyond the current 3.0 million ha of irrigated land. Data on upgrade types, costs, and expansion possibilities for all water regions are provided as supplementary materials.[14]

---

14   See **WaterRegionSupply.xlsx** in the **3_SupplyCurves** subfolder. A key to upgrade types is included in the main README file.
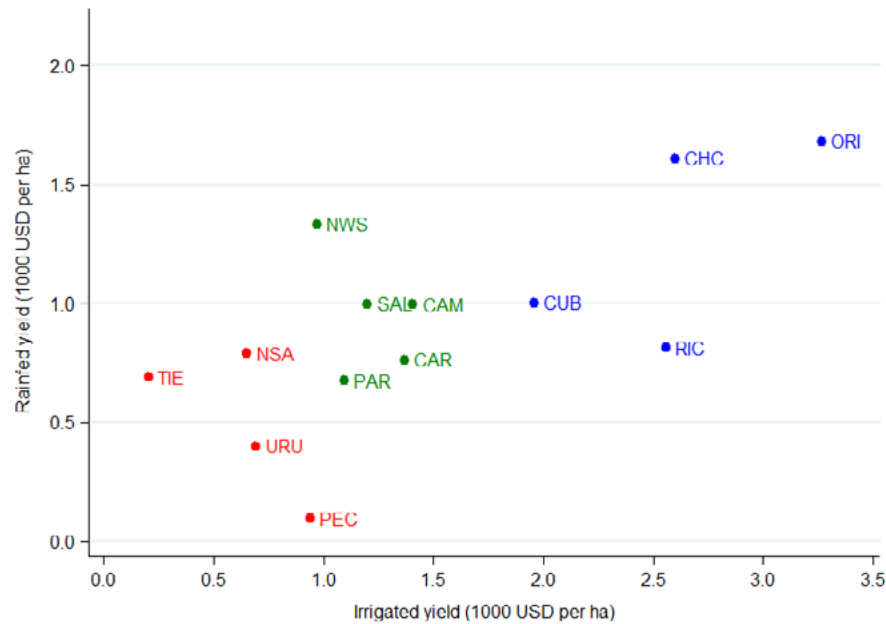
Figure 7. Cluster analysis of water regions to form irrigation response units (IRUs) within the Latin America (LAM) EPPA region.

3.4 Including Irrigable Land Supply Curves in an Economy-Wide Model

Winchester *et al.* (2016) include supply curves for additional irrigable land in the EPPA model, an economy-wide model that represents 16 global regions. For modeling tractability, they define irrigation response units (IRUs) as groups of water regions with similar crop yields within an EPPA region. Following Winchester *et al.* (2016), we determine IRU membership using a k-means cluster analysis of water regions based on their rainfed and irrigated yields. Each EPPA regions contains between one and four IRUs, yielding 46 IRUs globally. The IRU assignments for all regions are summarized in **Table A3**. The Stata code for the cluster analysis and a collection of figures illustrating the results in each region are included in the supplementary materials.[15] As an example, **Figure 7** depicts the cluster analysis results in the Latin America (LAM) region.

We aggregate the irrigable land supply curves for constituent water regions to form IRU supply curves for additional irrigable land. To approximate step irrigable land supply curves as 'smooth' functions, we econometrically estimate a supply elasticity parameter of the form $q = \beta p^\lambda$ to fit each IRU step function, with quantity in thousands of ha as q and price in USD per ha as p.[16] An exponential form is adopted so that the supply elasticity λ will remain constant for an IRU at all quantities of irrigated land supply curves for additional irrigable land

can be modeled using constant elasticity of substitution functions following the procedure outlined by Rutherford (2002). In the base year in EPPA, each IRU's quantity of irrigated land is set proportional to its contribution to aggregate production value within the EPPA region, with total land endowment remaining fixed from 2005 to 2050, the timespan modeled. Winchester *et al.*'s (2016) augmentation of the EPPA model allows additional irrigated land to be produced by combining rainfed land with capital representing investment in irrigation infrastructure. Expansion of irrigated land is determined endogenously in the model at an increasing cost until the difference in revenues from irrigated land versus rainfed land no longer cover the cost of investment in irrigation infrastructure. There is also a maximum limit on the amount of land that can be irrigated in each region.[17] Although we model irrigable land expansion at the IRU-level, irrigable land supply curves could be specified for each water region, or a user-defined aggregation of water regions, in other AGE models.

4. Supplementary Materials

We provide our datasets and code to allow other researchers to replicate and build upon the production value analysis and supply curve construction. We hope these resources will support future work to improve the representation of agricultural productivity in AGE models, whether at a national or global level. **Table 7** summarizes the available supplementary materials.

---

15   See **ClusterAnalysis.do** in the **3_SupplyCurves** subfolder.

16   See **SupplyCurves.gms** for the script to aggregate and econometrically estimate IRU supply curves

17   Winchester *et al.* (2016) provide a detailed explanation of this implementation within the EPPA model.

Table 7. Important supplementary resources

Folder	Contents	File
1_SpatialData	Spatial data for crop areas, yields, output prices, and production value	
	Shapefiles for aggregation to water regions, countries, or GTAP regions	
	Code (**A**) and datasets (including **B**) to calculate prices for MIRCA2000 crops	**A. Prices.gms** **B. Prices\CropPriceInterpolation.xlsx**
	Code to regenerate price, output, and value spatial data	**GenerateRasters.py**
2_ProductionValue	Code to aggregate area, output, and value spatial data to desired regions (**A**) and resulting summary files (**B**, etc.)	**A. Aggregate.py** **B. Compiledarea_GTAP.csv**
	Code to create alternative value estimates by GTAP region & crop	**GTAPreporting.gms**
	Summary file with production values by GTAP region & crop	**ValueSummary.xlsx**
	Summary file with relative errors by GTAP region & crop	**EvaluationMetric.xlsx**
	Summary file with disaggregated crop prices and production shares by MIRCA2000 crop and GTAP region	**CropDisaggregation.xlsx**
3_SupplyCurves	Data for water region supply step functions	**WaterRegionSupply.xlsx**
	Code to perform water region cluster analysis	**ClusterData.gms**, **ClusterAnalysis.do**
	Scatterplots illustrating current cluster results	**Graphs**
	Code to create IRU irrigable land supply curves	**SupplyCurves.gms**

5. Conclusion

The explicit representation of irrigated land within integrated assessment and economy-wide models can provide insights into how regions will balance the competing demands for land amidst the changing availability of regional water resources, and is an important step for looking at energy, water, and land interactions. To develop this irrigated land framework, modelers need to identify the current scope of irrigated land and define its potential for expansion. Previously, Haqiqi *et al.* (2016) formed the GTAP-Water Data Base to separate irrigated and rainfed land within an economy-wide model. However, to divide the production value for each crop into irrigated and rainfed components, they use estimates based on production volumes rather than a direct calculation of production values. While this approach is sufficient for some GTAP regions and crop sectors, irrigated production value shares based on output volumes differ from the directly calculated values in composite GTAP crop sectors within several regions because the mix of crops grown on each land type differs. Additionally, existing AGE analyses do not consider endogenous changes in irrigation infrastructure from changing water resources and food demand.

Winchester *et al.* (2016) address both of these short-comings. First, Winchester *et al.* (2016) use production value shares computed from a global crop price and output databases to divide rainfed and irrigated land within the EPPA model. Second, they derive irrigable land supply curves for 126 global water regions using cost and water yield estimates for water storage and irrigation system improvements. However, they do not generate value estimates compatible with GTAP regions and crop sectors, nor do they provide the data and work stream to develop irrigable land supply curves.

Thus, we extended the work of Winchester *et al.* (2016) in two ways. First, we provide estimates of irrigated and rainfed production value at a finer spatial resolution, as well as aggregated to the GTAP 9 regions and crop sectors. We also compute a relative error value for specific GTAP regions and crops to guide researchers and AGE modelers in adjusting production value estimates based on output shares. Second, we provide supply curves to expand irrigated land at a disaggregated, water-region level in addition to the aggregated IRU-level considered by Winchester *et al.* (2016). Finally, we make available the complete set of data and code needed to directly quantify the current production value of irrigated land, evaluate differences in the validity of estimating production values based on output volumes, and to aggregate the 126 water-region irrigable land supply curves. We hope this access will provide useful and useable data and tools for the integrated assessment community and for those working on natural resource links in economy-wide models.

Acknowledgments

6. References

Aguiar, A., B. Narayanan & R. McDougall (2016): An Overview of the GTAP 9 Data Base. *Journal of Global Economic Analysis*, 1(1) (doi:10.21642/JGEA.010103AF).

Bonan, G.B., K.W. Oleson, M. Vertenstein, S. Lewis, X. Zeng, Y. Dai, R.E. Dickinson & Z.-L. Yang (2002): The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model. *Journal of Climate*, **15**: 3123–3149 (doi:10.1175/1520-0442(2002)015<3123:TLSCOT>2.0.CO;2).

Chen, Y.-H.H., S. Paltsev, J. Reilly, J. Morris, V. Karplus, A. Gurgel, N. Winchester, P. Kishimoto, É. Blanc & M. Babiker (2017): *The MIT Economic Projection and Policy Analysis (EPPA) Model: Version 5*. MIT Joint Program **Technical Note 16**, 32 p. (http://globalchange.mit.edu/publication/16620).

Fant, C., A. Gueneau, K. Strzepek, S. Awadalla, W. Farmer, É. Blanc & C.A. Schlosser (2012): *CliCrop: a Crop Water-Stress and Irrigation Demand Model for an Integrated Global Assessment Modeling Approach*. MIT Joint Program **Report 214**, April, 26 p. (http://globalchange.mit.edu/publication/15732).

Food and Agriculture Organization of the United Nations (FAO) (2015a): *Crops*. FAOSTAT. (http://www.fao.org/faostat/en/#data/QC).

Food and Agriculture Organization of the United Nations (FAO) (2015b): *Irrigation and drainage*. AQUASTAT. (http://www.fao.org/nr/water/aquastat/irrigationdrainage/index.stm).

Food and Agriculture Organization of the United Nations (FAO) (2015c): *Producer Prices - Annual*. FAOSTAT. (http://www.fao.org/faostat/en/#data/PP).

Gurgel, A., J.M. Reilly & S. Paltsev (2007): Potential land use implications of a global biofuels industry. *Journal of Agricultural & Food Industrial Organization*, 5(2): Article 9 (doi:10.2202/1542-0485.1202)

Haqiqi, I., F. Taheripour, J. Liu & D. van der Mensbrugghe (2016): Introducing Irrigation Water into GTAP Data Base Version 9. *Journal of Global Economic Analysis*, **1**(2): (doi:10.21642/JGEA.010203AF).

Johansson, D.J.A. & C. Azar (2007): A scenario based analysis of land competition between food and bioenergy production in the US. *Climatic Change*, **82**(3–4): 267–291. (doi:10.1007/s10584-006-9208-1).

Liu, J., T. Hertel & F. Taheripour (2016): Analyzing Future Water Scarcity in Computable General Equilibrium Models. *Water Economics and Policy*, **2**(4): 30 p. (doi:10.1142/S2382624X16500065).

Liu, J., T. Hertel, F. Taheripour, T. Zhu & C. Ringler (2014): International trade buffers the impact of future irrigation shortfalls. *Global Environmental Change*, **29**: 22–31. (doi:10.1016/j.gloenvcha.2014.07.010).

NRCS [Natural Resources Conservation Service] (2014): *Annual and Perennial Crop List*. United States Department of Agriculture.

OECD/FAO (2017), OECD-FAO Agricultural Outlook 2017-2026, OECD Publishing, Paris. (doi:10.1787/agr_outlook-2017-en)

Peters, J.C. (2016): The GTAP-Power Data Base: Disaggregating the Electricity Sector in the GTAP Data Base. *Journal of Global Economic Analysis*, **1**(1), 209–250. (doi:10.21642/JGEA.010104AF).

Portmann, F.T., S. Siebert & 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 Biogeochemical Cycles*, **24**(1) (doi:10.1029/2008GB003435).

Rosegrant, M.W. & the IMPACT Development Team (2012): *International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): Model Description*. International Food Policy Research Institute (IFPRI), Washington, D.C. (https://www.researchgate.net/publication/275582667_IMPACT_Technical_Description).

Rutherford, T.F. (2002): *Lecture notes on constant elasticity functions*. University of Colorado. (http://www.gamsworld.org/mpsge/debreu/ces.pdf).

Siebert, S. & 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): 198–217 (doi:10.1016/j.jhydrol.2009.07.031).

Sokolov, A.P., C.A. Schlosser, S. Dutkiewicz, S. Paltsev, D.W. Kicklighter, H.D. Jacoby, R.G. Prinn, C.E. Forest, J.M. Reilly, C. Wang, B. Felzer, M.C. Sarofim, J. Scott, P.H. Stone, J.M. Melillo & J. Cohen (2005): *The MIT Integrated Global System Model (IGSM) Version 2: Model Description and Baseline Evaluation*. MIT Joint Program **Report 124**, 40 p. (http://globalchange.mit.edu/publication/14579)

Strzepek, K., C.A. Schlosser, A. Gueneau, X. Gao, E. Blanc, C. Fant, B. Rasheed & H.D. Jacoby (2012): *Modeling Water Resource Systems under Climate Change: IGSM-WRS*. MIT Joint Program **Report 236**, 54 p. (http://globalchange.mit.edu/publication/15904).

Strzepek, K., A. Schlosser, A. Gueneau, X. Gao, É. Blanc, C. Fant, B. Rasheed & H.D. Jacoby (2013): Modeling water resource systems within the framework of the MIT Integrated Global System Model: IGSM-WRS. *Journal of Advances in Modeling Earth Systems*, **5**(3), 638–653: (doi:10.1002/jame.20044)

Taheripour, F., T.W. Hertel & J. Liu (2013a): Introducing water by river basin into the GTAP-BIO model: *GTAP-BIO-W. GTAP Working Paper No. 77*. (https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=4304).

Taheripour, F., T.W. Hertel & J. Liu (2013b): The role of irrigation in determining the global land use impacts of biofuels. *Energy, Sustainability and Society*, **3**:4, 18 p. (doi:10.1186/2192-0567-3-4).

Taheripour, F., T.W. Hertel, B. Narayanan & S. Sahin (2016): Economic and Land Use Impacts of Improving Water Use Efficiency in Irrigation in South Asia. *Journal of Environmental Protection*, **7**(11): 1571–1591 (doi:10.4236/jep.2016.711130).

Winchester, N., K. Ledvina, K. Strzepek & J.M. Reilly (2016): *The Impact of Water Scarcity on Food, Bioenergy and Deforestation*. MIT Joint Program **Report 300**, 20 p. (http://globalchange.mit.edu/publication/16280).

Winchester, N. & J.M. Reilly (2015): The feasibility, costs, and environmental implications of large-scale biomass energy. *Energy Economics*, **51**: 188–203 (doi:10.1016/j.eneco.2015.06.016).

Wise, M., J. Dooley, P. Luckow, K. Calvin & P. Kyle (2014): Agriculture, land use, energy and carbon emission impacts of global biofuel mandates to mid-century. *Applied Energy*, **114**: 763–773 (doi:10.1016/j.apenergy.2013.08.042).

7. Appendix

Table A1. Mapping of FAO crops to the 26 MIRCA crops

MIRCA2000[a]	FAO[b]				
Barley	• Barley				
Cassava	• Cassava				
Citrus	• Fruit, citrus, nes	• Grapefruit (inc. pomelos)	• Lemons & limes • Oranges	• Tangerines, mandarins, clementines, satsumas	
Cocoa	• Cocoa, beans				
Coffee	• Coffee, green				
Cotton	• Cotton lint				
Date palm	• Dates				
Fodder	• Maize	• Rye	• Wheat		
Grapes/vine	• Grapes				
Groundnuts/ peanuts	• Groundnuts, with shell				
Maize	• Maize				
Millet	• Millet				
Oil palm	• Oil, palm fruit				
Others Annual	• Anise, badian, fennel, coriander • Beans, green • Buckwheat • Cabbages & other brassicas • Canary seed • Carrots and turnips • Cauliflowers and broccoli • Cereals, nes • Chillies & peppers, dry • Chillies & peppers, green	• Cucumbers & gherkins • Eggplants (aubergines) • Fibre crops, nes • Flax fibre & tow • Fonio • Garlic • Ginger • Grain, mixed • Hemp tow waste • Hempseed • Jute	• Leeks, other alliaceous vegetables • Lettuce & chicory • Linseed • Lupins • Melons, other (inc. cantaloupes) • Melonseed • Mushrooms & truffles • Mustard seed • Oats • Oil, stillingia • Oilseeds, nes • Okra	• Onions, dry • Onions, shallots, green • Peas, green • Pineapples • Popcorn • Poppy seed • Pumpkins, squash & gourds • Quinoa • Roots & tubers, nes • Safflower seed • Sesame seed • Spinach	• Strawberries • String beans • Sugar crops, nes • Sweet potatoes • Taro (cocoyam) • Tobacco, unmanufactured • Tomatoes • Triticale • Vegetables, fresh, nes • Watermelons • Yams • Yautia (cocoyam)
Others Perennial	• Agave fibres, nes • Almonds, with shell • Apples • Apricots • Areca nuts • Artichokes • Asparagus • Avocados • Bananas • Berries nes • Blueberries • Brazil nuts, with shell • Carobs • Cashew nuts, w/ shell	• Cashewapple • Castor oil seed • Cherries • Cherries, sour • Chestnut • Chicory roots • Cinnamon (canella) • Cloves • Cocoa, beans • Coconuts • Cranberries • Currants • Dates • Figs	• Fruit, fresh nes • Gooseberries • Gums, natural • Hazelnuts, with shell • Hops • Jojoba seed • Kapok fruit • Karite nuts (sheanuts) • Kiwi fruit • Kola nuts • Mangoes, mango-steens, guavas • Manila fibre (abaca) • Maté	• Nutmeg, mace & cardamoms • Nuts, nes • Olives • Papayas • Peaches & nectarines • Pears • Pepper (piper spp • Peppermint • Persimmons • Pistachios • Plantains • Plums & sloes	• Pyrethrum, dried • Quinces • Ramie • Raspberries • Rubber, natural • Sisal • Spices, nes • Tallowtree seed • Tea • Tea, nes • Tung nuts • Vanilla • Walnuts, with shell
Potatoes	Potatoes				
Pulses	• Bambara beans • Beans, dry	• Broad beans, horse beans, dry	• Chick peas • Cow peas, dry	• Lentils • Lupins	• Peas, dry • Pigeon peas	• Pulses, nes • Vetches
Rapeseed/ canola	• Rapeseed				
Rice	• Rice, paddy				
Rye	• Rye				
Sorghum	• Sorghum				
Soybeans	• Soybeans				
Sugar beet	• Sugar beet				
Sugarcane	• Sugar cane				
Sunflower	• Sunflower seed				
Wheat	• Wheat				

[a]Crops included in the MIRCA2000 dataset (Portmann *et al.,* 2010).   [b]Crops included in the FAO's price dataset (FAO, 2015c)

Table A2. Pairings between countries missing prices and countries providing proxy prices

Country Missing Prices	Country Providing Proxy Prices	Country Missing Prices	Country Providing Proxy Prices
Afghanistan	Pakistan	Mauritania	Mali
Angola	Namibia	Montenegro	Bosnia & Herzegovina
Benin	Togo	Myanmar	Thailand
Central African Republic	Cameroon	North Korea	South Korea
Chad	Niger	Oman	Yemen
Cuba	Dominican Republic	Papua New Guinea	Indonesia
Democratic Republic of the Congo	Congo	Sierra Leon	Guinea
Djibouti	Eritrea	Somalia	Ethiopia
Gabon	Cameroon	South Sudan	Sudan
Haiti	Dominican Republic	Swaziland	South Africa
Iraq	Iran	Syria	Lebanon
Kuwait	Qatar	Uganda	Rwanda
Lesotho	South Africa	United Arab Emirates	Qatar
Liberia	Ivory Coast	Uzbekistan	Kazakhstan
Libya	Algeria	Zambia	Botswana
Liechtenstein	Switzerland	Zimbabwe	Botswana

Table A3. Water regions within each EPPA region and Irrigation Response Unit (IRU)

EPPA Region	Description	IRU-1	IRU-2	IRU-3	IRU-4
AFR	Africa	CAF, EAC, KAL, MAD, NIG, SEN, WAC	NAC, NLE, SAC	HOA, LCB, LIM, NWA, ORA, SAF, SAH, VOT	CON
ANZ	Australia-New Zealand	EAU, WAU	CAU, MAU	NZE	PAO
ASI	Dynamic Asia	BOR, INW	SKP	INE, MEK, PHI, TMM	--
BRA	Brazil	TOC	AMA, SAN	NEB	--
CAN	Canada	CCA	RWI	CAN, GLA	--
CHN	China	CHJ, HAI, HUN, SON,	LAJ, LMO, SEA	HUL, YHE, ZHJ	--
EUR	European Union	BRI, IRE, RHI	ELB, SCA	IEM, IWA	ITA, LBO, RHO, SEI
IND	India	CHO	BRR, CHO, EGH, GAN, GOD, IEC, KRI, LUN, MAT, SAY	--	--
JPN	Japan	JAP	--	--	--
LAM	Other Latin America	CHC, CUB, ORI, RIC	CAM, CAR, NWS, PAR, SAL	NSA, PEC, TIE, URU	--
MES	Middle East	EME, TIG, WAI	ARA	--	--
MEX	Mexico	YUC	MIM, UME	--	--
REA	Rest of East Asia	BRT, IND, SRL	NKP, ROW	--	--
ROE	Rest of Europe and Central Asia	BAL, DAN	AMD, LBA, SYD	BLA, DNI, ODE	--
RUS	Russia	NER, VOG	OB, UMO, URA	AMR, YEN	--
USA	United States	CAL, COL, GBA	RIG, SEU, WGM	ARK, MIS, MOU, OHI	COB, USN

Joint Program Report Series - Recent Articles

For limited quantities, Joint Program Reports are available free of charge. Contact the Joint Program Office to order.
Complete list: http://globalchange.mit.edu/publications