Chapter 6
Benchmark Data: Integrating Biophysical and Economic Information in a Consistent Geospatial Dataset

Iman Haqiqi and Uris Lantz C. Baldos

Existing environmental databases often suffer from fragmentation, inconsistencies, and limited spatial representation. In addition, economic representation at the geospatial level is often limited. However, the SIMPLE-G dataset aims to overcome these limitations by providing a unified platform for exploring the interactions between land, water, and the environment. Each data point within the dataset is geographically referenced with details about economic values and biophysical variables, enabling researchers to conduct spatial analyses and uncover crucial insights into localized resource use, economic patterns, and sustainability challenges. Furthermore, the dataset's comprehensive scope, which encompasses both production and consumption aspects of the agri-food system, fosters a holistic understanding of resource flows and economic dependencies. This data resource empowers researchers to tackle complex environmental questions and propose data-driven solutions for promoting sustainable land and water management practices. While the chapter concentrates on the SIMPLE-G-US dataset (Baldos et al. 2020; Haqiqi et al. 2023), a similar approach is taken for other global and regional models (Liu et al. 2017; Haqiqi et al. 2022).

The SIMPLE-G model uses two types of information: benchmark data and behavioral parameters. Benchmark data include information describing the equilibrium conditions of input use, agricultural production, trade, and food consumption in the reference year. The values of the variables are expected to change and are updated to reflect the new economic conditions following a SIMPLE-G simulation. The benchmark data are collected from various sources, including international databases, satellite imagery, climate and hydrological data, and national censuses. These data must be processed and made consistent to produce gridded data that can be used in the model.

I. Haqiqi (✉) · U. L. C. Baldos
Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University, West Lafayette, IN, USA
e-mail: ihaqiqi@purdue.edu

I. Haqiqi, T. W. Hertel (eds.), *SIMPLE-G*,
https://doi.org/10.1007/978-3-031-68054-0_6

Behavioral parameters refer to information that describes how the model responds to changes in various factors (e.g., prices and income). These parameters are based on independent estimates of consumer and producer behavior. They govern consumers' and producers' responses to changes in the economic, climatic, and environmental conditions of agricultural production. By definition, these parameters are fixed over the course of the simulation.

Before describing the gridded data, we explain how the regional (non-gridded) SIMPLE database is constructed.

1 Non-gridded Data

The non-gridded SIMPLE database includes information on macroeconomic variables, commodity consumption, regional trade for crops, and regional production of crops, livestock, and processed food. Additional data on input and output carbon emissions for each commodity (Aguiar et al. 2019), land-use change emissions (West et al. 2010), and food security metrics (FAO 2020) are also available for model versions with environmental and food security modules. A publicly available version of the non-gridded database and compatible model version has detailed information for 150+ countries for the year 2017 (Baldos 2023).

1.1 Regional Production and Trade

The regional crop production database in SIMPLE-G is constructed using country-level economic and agricultural data taken from several sources. Crop production, producer crop price, and cropland area data are sourced from the United Nations Food and Agricultural Organization's FAOSTAT database (FAO 2020). Global prices are computed for each crop and are used to convert quantity of crop production into corn-equivalent production using the ratio of each crop's global price and the global price of corn. The value of crop production is also calculated using these data. Since we assume that the benchmark dataset represents a long-run equilibrium, we impose zero profit conditions in the crop sector. Therefore, the annual flow of land and nonland input costs in each region can be calculated using the value of crop production and the input cost shares from the GTAP database (Aguiar et al. 2019). The crop sale shares across the domestic and international markets are taken from the GTAP database (Aguiar et al. 2019).

1.2 Regional Consumption

On the demand side, the benchmark data include the value and quantity of crops used in direct food consumption, for feed in the livestock sectors, as raw inputs in the processed food industries, and as feedstocks in the biofuel sector. The amount of crop feedstock used by the biofuel sector in each region is calculated using the sales shares of the crop sector in the bioenergy sectors from the GTAP-BIO V.9 database (Taheripour et al. 2017). Remaining crop quantities are allocated to food, feed, and processed food input use using crop purchase shares in the local and global markets using share data from the GTAP database (Aguiar et al. 2019).

Consumer demand data in the model are mainly driven by population and per capita income. Population and real gross domestic product (GDP) data are taken from FAOSTAT (FAO 2020), and these are used to calculate the per capita income in each region. The food security module in SIMPLE includes food security metrics based on average food consumption in each region (Baldos and Hertel 2014). Following Neiken (2003), each region has a unique distribution of average caloric consumption and a minimum daily dietary energy intake which delineates the fraction of the distribution below this minimum threshold. The nutrition data include the prevalence of caloric undernutrition and the undernutrition headcount for each region. This distribution shifts depending on changes in average dietary energy consumed. The module also captures shifts in the caloric composition of food by linking food caloric content to per capita income levels (Baldos and Hertel 2013). Food security data from FAOSTAT (FAO 2020) are used to initialize the nutritional information for each region.

1.3 Food Processing and Livestock Production

The value of consumption for livestock and processed food commodities is equal to the total value of output for these sectors, given the assumption that total revenue is equal to total consumption expenditure for each commodity. Under zero-profit conditions, the total value of output is equal to the total production costs. The total cost of crop inputs in the livestock and processed food sectors is calculated using the crop prices and crop use quantities. The values of non-crop inputs are then computed from the total cost of crop inputs and input cost shares. The input cost shares for livestock and processed food commodities are taken from the GTAP database (Aguiar et al. 2019).

1.4 Agricultural Greenhouse Gas Emissions

Some model versions (e.g., Chap. 10) report changes in agricultural greenhouse gas (GHG) emissions, specifically emissions from crop, livestock, and processed food production as well as land-use change emissions due to cropland expansion. Data on GHG emissions from agricultural production are based on the GTAP database (Aguiar et al. 2019), which reports CO_2 emissions from fossil fuel combustion using detailed energy volume data from the International Energy Agency and combustion factors from the Revised 1996 *IPCC Guidelines for National Greenhouse Gas Inventories* (IPCC/OECD/IEA 1997). GHG emissions from methane, nitrous oxide and fluorinated, gases are based on the non-CO_2 GTAP database (Chepeliev 2023), which uses FAOSTAT data (FAO 2020) for agricultural emissions. Land-use change emissions from converting natural land into cropland rely on the global carbon stocks calculated by West et al. (2010). These carbon stocks are constructed using spatially explicit datasets on potential vegetation and soil carbon.

2 Benchmark Gridded Data: US Example

SIMPLE-G requires benchmark gridded data for key economic and biophysical variables describing the initial equilibrium of the crop economy. The preferred gridded data sources vary by region. Table 6.1 lists the variables and underlying data sources for the United States. Comparable data sources have been compiled for several other regions (see the applications in Part IV). Each model includes a gridded dataset that includes cropland use, crop production, nitrate leaching, and water use.

Given the centrality of the gridded database to the use and credibility of SIMPLE-G, this section provides a brief overview of how to construct a SIMPLE-G dataset. Figure 6.1 describes the workflow for constructing the benchmark data and behavioral parameters for a US-focused version of SIMPLE-G, utilizing gridded

Table 6.1 Gridded data for SIMPLE-G-US

Data	Source	Description of process
Cropland area	USDA CDL, USGS MIrAD	The share of cropland area in each grid cell is calculated from 30 m resolution. The share of irrigated area from cropland is calculated from 25 m resolution
Value of land	USDA Census of Agriculture	Land area is multiplied by rent in USD per ha assuming uniform rent within a county
Value of crops produced	USDA Census of Agriculture, GCWM by crop	The value of crops sold by county is distributed to grid cells within a county based on gridded pattern from GCWM and irrigated yield gap
Quantity index of crops	Estimated	The quantity index follows the value of crops sold divided by reference crop price to create a price adjusted corn-equivalent index
Value of irrigation	USGS and USDA	Irrigation cost shares are estimated based on county-specific data on share of sprinkler, drip irrigation, surface water, and groundwater
Volume of water	USGS	Water volume follows land multiplied by water intensity in m^3/ha
Quantity of N fertilizer	Cao et al. (2018), Agro-IBIS	Total fertilizer application follows land multiplied by kg/ha
Value of N fertilizer	Estimated	Value of fertilizer follows quantities multiplied by price USD per kg

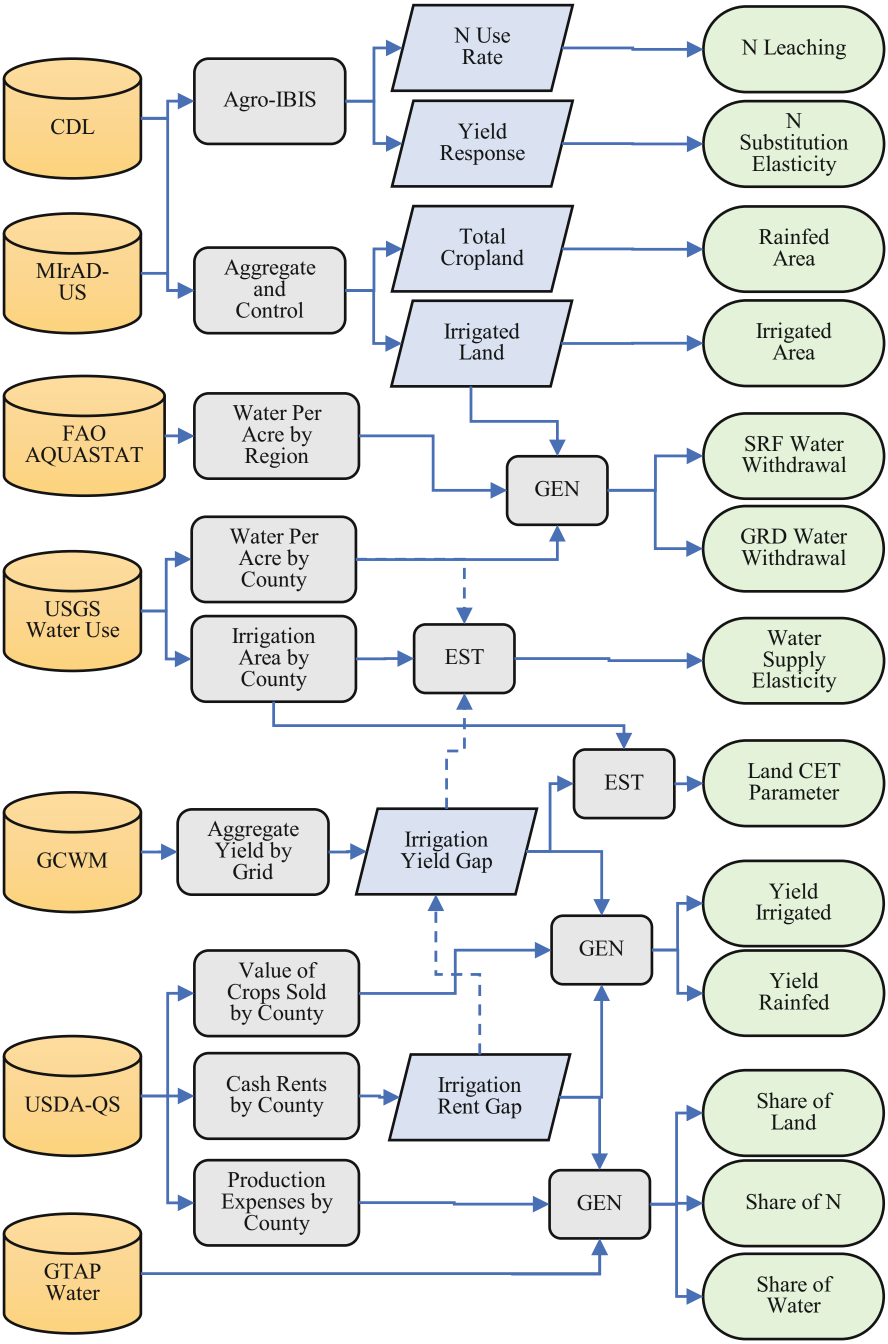

Fig. 6.1 Overview of main data and parameter processing method for SIMPLE-G-US. "GEN" represents a computation that does not involve statistical regressions; "EST" is a process that includes statistical estimation. Other versions, (gridded World, gridded China, and gridded Brazil) follow similar flows but employ rich global and national data sources.

data for the United States and employing regional information for other parts of the world. Other versions (e.g., SIMPLE-G-China, SIMPLE-G-Brazil) follow similar flows but employ the best available national data sources from those countries. In addition to the primary source data, there are two types of methodologies for generating the gridded data: "GEN" represents a computation that does not involve statistical regressions, and "EST" is a process that includes statistical estimation.

Although resolution is not a constraint for SIMPLE-G (see computational section for examples with millions of grid cells), current versions work with crop production at the level of georeferenced grid-cell units at 5 arcmin resolution (squares with sides of 9.26 km at the equator). Most of our collaborators in the biophysical sciences use this resolution. We have added gridded information for US crop production covering both irrigated and nonirrigated practices and including the value and quantity of crop output, land use, N fertilizer input, water, and other aggregated inputs.

2.1 Cropland Area: Rainfed and Irrigated

Cropland area data were obtained from the USDA Cropland Data Layer (Han et al. 2012) at a resolution of 30 meters and aggregated to 5 arcmin. Irrigated cropland data were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset for the United States (MIrAD-US) provided by the US Geologic Survey (USGS) at a resolution of 250 m (Brown and Pervez 2014), aggregated to 5 arcmin. These data determine the distribution of irrigated and nonirrigated cropland in the United States.

2.2 Crop Production: Price-Adjusted Quantity Index

Aggregated output for each grid cell is recorded as the corn-equivalent total crop output, calculated as the sum of the value the crops sold divided by the price of corn in the base year, yielding "corn-equivalent tons of output." We take the value of crops sold per acre from USDA National Agricultural Statistics Service by county (USDA-NASS 2019) and use simulated yields from the Global Crop Water Model (GCWM) (Siebert and Döll 2010) to generate gridded yield for the grid cells in each county. The GCWM data are aggregated over all crops using USDA/FAO actual prices to calculate the corn-equivalent crop output for each grid cell.

We split the base data into irrigated and rainfed crop production using various satellite datasets and county-level information from the USDA and USGS, as well as the simulated yield of irrigated and nonirrigated crops. For yield estimation, we assume that grid cell aggregated yield per hectare is equal to the county average in which the grid cell is located. We split the gridded total production into irrigated and nonirrigated components using total, irrigated, and nonirrigated land as obtained from MIrAD-US and the cropland data layer (CDL) and the ratio of rainfed to irrigated yields in a given grid cell as estimated by Siebert and Döll (2010) for 29 crop categories and aggregated to all crops according to production value weights. Total cropland area from the CDL is matched with USDA county-level cropland to ensure consistency of yield and area at the county level.

2.3 Nitrogen Fertilizer Applications

Nitrogen (N) fertilizer application rates per hectare per year for each grid cell were obtained from Cao et al. (2018) for all crops at 5 km and from Agro-IBIS (the agricultural version of the Integrated Biosphere Simulator) for rainfed and irrigated production of major crops (Lark et al. 2022). This product provided high-resolution (8 arcsecond) land cover and nutrient application maps across the continental United States for the period of 1750 to 2017, accounting for the nutrient legacies of historical land use/cover. The Agro-IBIS land cover categories are based on vegetation type simulations. This product is also consistent with our landcover data, which is based on several gridded land cover datasets and historical county-level USDA Census of Agriculture data. Irrigation maps from Agro-IBIS were created using the MIrAD-US and historical data from the USDA Census of Agriculture. With its rich set of information on gridded economic activity, the agroecological Agro-IBIS model helped us estimate irrigated versus nonirrigated fertilizer application rates.

2.4 Volume of Water Withdrawals

Irrigation water withdrawal rates are estimated using county-level USGS water use data (Maupin et al. 2014). We calculate total water withdrawal per irrigated hectare and allocate it to groundwater and surface water using county-level USGS water use data by source. For some version of SIMPLE-G a quantity index is introduced for water which requires conversion to m3 to show the level of variables.

2.5 Value of Crop Production

Calculating the value of crops produced at the grid-cell level requires several steps. First, satellite data with sufficient resolution are acquired to identify cropland within each grid cell. These data are combined with gridded maps of irrigated areas to separate irrigated and rainfed areas in each grid cell. To avoid the computational issues that arise when extremely small pieces of cropland are modeled, we assign a minimum area to each grid cell in the model if it is cultivated. Next, utilizing the agronomic GCWM data, crop yields and areas are estimated by crop at the grid-cell level. Agronomic models incorporate factors such as weather, soil conditions, and crop management practices to simulate crop growth and predict yields by irrigation technology. The combination of land area and yield provides information on production by crop. Next, information on the average prices of crops at the national level is collected from FAOSTAT. For each grid cell, crop-specific values are calculated from the agronomic model by multiplying production by the corresponding crop price, considering the crop area data layer derived from the satellite data. This will provide an estimate of the total value of crops produced within that grid cell by irrigation technology. Since our goal is to match the USDA census data by county, we apply the gridded patterns of production value to the USDA county-level data on the value of crops produced in each county to obtain census-consistent, gridded estimates of the value of crop production.

This approach provides a comprehensive and spatially explicit assessment of crop production values, providing granular insights into agricultural productivity and economic contributions.

2.6 Value of Land

The annual payment to land, also known as land rent, is a crucial component of the economic model. It represents the compensation that landowners receive for the use of their land in agricultural production. We assume that these same landowners consider alternative uses, where feasible; therefore, each grid cell has an upward-sloping land supply curve, reflecting the opportunity cost of crop land. We consider explicit payments from rental lands to landowners as well as implicit self-payment of landowner farmers. The value of water is calculated based on irrigated and rainfed cash rents by county in the United States.

In SIMPLE-G, total annual payments to land play a key role in constructing the database. At the grid-cell level, the calculation of land rent typically involves two key factors: land area and rental rates by irrigation type. While land area information can be obtained from various sources (e.g., satellite imagery), information about land rent requires surveys or census data. Land rents represent the per unit payment, typically expressed in terms of dollars per hectare (or per acre) per year, and vary depending on several factors, including soil quality, proximity to markets, access to irrigation, crop mix, and overall demand for land in the county. Understanding these factors and their influence on land rents is crucial for understanding spatial heterogeneity in cost structures across a region.

To calculate the total payment to land in a grid cell, the gridded land area is multiplied by the average county-level rental rate, as reported by the USDA. This calculation provides an estimate of the overall compensation that landowners receive, implicitly and explicitly, for the use of their land for agricultural purposes. It is important to note that this calculation represents an average and may not accurately reflect the specific rental payments in a given grid cell.

2.7 Value of Irrigation

The central challenge with regard to valuing water is that there is usually no market for water. Only a small portion of irrigation water is traded in markets. Even where they exist, market prices for water rarely reflect the true value added of irrigation water in agriculture. Thus, we developed an alternative approach to estimate the value of water. We start from the land rent. The cost share (θ) for cropland (L) for each grid-cell (g) is defined as

$$\theta^{L}_{g,l} = \mathrm{QL}_{g,l}\mathrm{PL}_{g,l}/\mathrm{QY}_{g,l}\mathrm{PY}_{g,l} \tag{6.1}$$

where QY and QL are production level and land area, and PY and PL are price index and land rents, respectively. As noted, much of the cropland is owned by the enterprise or individual farming the land. Thus, the true value added of land is not observed. Therefore, we infer rental rates from average cash rents reported by other farms in the neighborhood. In combination of CDL cropland data, we then calculate the total rental costs land. We have constructed an empirical model to estimate the cost shares of water inputs for groundwater, surface water, and irrigation input. We assume that the distribution of cost shares is driven by the volume share of water extracted from groundwater (XGW) and surface water (XSW), extent of area irrigated by each source (LGW and LSW, respectively), and shares of area equipped with higher efficiency irrigation technology such as microdrip irrigation (XMIC) and sprinklers (XSPR). The cost shares for groundwater, surface water, and irrigation equipment inputs are estimated using the following equations from Haqiqi (2023):

$$\begin{aligned}
\theta^{\mathrm{GW}}_{l,g} &= \alpha^{G}_{0} + \alpha^{G}_{1}\mathrm{XGW}_{l,g} + \alpha^{G}_{1}\mathrm{LGW}_{l,g} \\
\theta^{\mathrm{SW}}_{l,g} &= \alpha^{S}_{0} + \alpha^{S}_{1}\mathrm{XSW}_{l,g} + \alpha^{S}_{1}\mathrm{LSW}_{l,g} \\
\theta^{\mathrm{KW}}_{l,g} &= \alpha^{K}_{0} + \alpha^{K}_{1}\mathrm{XMIC}_{g} + \alpha^{K}_{2}\mathrm{XSPR}_{g}
\end{aligned} \tag{6.2}$$

Equation (6.2) considers the associated costs of irrigation, including the labor, fuel, and other inputs required for water withdrawal, transport, and application. Therefore, we can empirically estimate the irrigation costs using the observed variables at the county level. In a statistical framework, we partially attribute the spatial differences in fuel, labor, and other nonland material to irrigation. Specifically, county-level nonland costs are estimated as a function of irrigated areas in acres, surface water withdrawals in gallons per acre, groundwater withdrawals in gallons per acre, share of area with sprinkler irrigation, and share of area with microdrip irrigation. We differentiate the marginal costs of groundwater and surface water irrigation. County-level expense data are obtained from USDA Census of Agriculture for total costs and input costs, including fuel, labor, fertilizer, seeds, chemicals, and other inputs for 2002, 2007, 2012, and 2017. Information on physical area and volume of water (IR-WGWFr, IR-WSWFr, IR-IrSpr, and IR-IrMic) is obtained from USGS Estimated Use of Water in the United States County-Level Data for 2005, 2010, and 2015.

3 Summary

The SIMPLE-G benchmark data include a consistent geospatial dataset that provides information on both economic values and biophysical variables for a reference year. At the regional level, this dataset includes exports, imports, income, and consumer expenditures on different food and non-food categories. At the grid-cell level, this includes the value of crop sales and production costs such as land, water, and labor expenses. The biophysical information includes cropland area, volume of water withdrawals, fertilizer applications, crop yields, and different land uses. These rich data enable researchers to conduct in-depth geospatial economic and sustainability analyses related to agriculture, land, water, and environment.

Acknowledgment and Competing Interests The authors acknowledge support from the National Science Foundation award #2118329: "NSF Institute for Geospatial Understanding through an Integrative Discovery Environment (I-GUIDE)," the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modeling, MultiSector Dynamics under Cooperative Agreement DE-SC0022141; the United States Department of Agriculture AFRI grant #2019-67023-29679, "Economic Foundations of Long Run Agricultural Sustainability," and the National Science Foundation INFEWS award #1855937, "Identifying Sustainability Solutions Through Global-Local-Global Analysis of a Coupled Water-Agriculture-Bioenergy System."

The findings and conclusions presented in this chapter are those of the authors and should not be construed to represent any official determination or policy of the US Department of Agriculture (USDA), the US government, the Department of Energy (DOE), or the National Science Foundation (NSF). Furthermore, we declare that there is no conflict of interest related to this work.

References

Aguiar, Angel, Maksym Chepeliev, Erwin L. Corong, Robert McDougall, and Dominique van der Mensbrugghe. 2019. The GTAP data base: Version 10. Journal of global. *Economic Analysis* 4: 1–27. https://doi.org/10.21642/JGEA.040101AF.

Baldos, Uris Lantz C. 2023. SIMPLE database and model for base year 2017. https://doi.org/10.13019/RPZW-BX12.

Baldos, Uris Lantz C., and Thomas W. Hertel. 2013. Looking back to move forward on model validation: Insights from a global model of agricultural land use. *Environmental Research Letters* 8: 034024. https://doi.org/10.1088/1748-9326/8/3/034024.

———. 2014. Global food security in 2050: The role of agricultural productivity and climate change. *Australian Journal of Agricultural and Resource Economics* 58: 554–570. https://doi.org/10.1111/1467-8489.12048.

Baldos, Uris Lantz C., Iman Haqiqi, Thomas W. Hertel, Mark Horridge, and J. Liu. 2020. SIMPLE-G: A multiscale framework for integration of economic and biophysical determinants of sustainability. *Environmental Modelling & Software* 133. Elsevier: 104805. https://doi.org/10.1016/j.envsoft.2020.104805.

Brown, Jesslyn F., and Md Shahriar Pervez. 2014. Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. *Agricultural Systems* 127: 28–40. https://doi.org/10.1016/j.agsy.2014.01.004.

Cao, Peiyu, Lu Chaoqun, and Yu. Zhen. 2018. Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: Application rate, timing, and fertilizer types. *Earth System Science Data* 10: 969–984. https://doi.org/10.5194/essd-10-969-2018.

Chepeliev, Maksym. 2023. GTAP-power data base: Version 11. *Journal of Global Economic Analysis* 8: 101–135. https://doi.org/10.21642/JGEA.080203AF.

FAO. 2020. *FAOSTAT: Food and agriculture data*. Rome: Food and Agriculture Organization of the United Nations. http://faostat.fao.org/. Accessed 1 Dec 2020.

Han, Weiguo, Zhengwei Yang, Liping Di, and Richard Mueller. 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. https://doi.org/10.1016/j.compag.2012.03.005.

Haqiqi, Iman. 2023. The value of water in US agriculture: Integrating spatially and temporally heterogeneous hydroclimatic and economic data (version 1.0). MyGeoHUB. https://doi.org/10.13019/9MXE-T280.

Haqiqi, Iman, Chris J. Perry, and Thomas W. Hertel. 2022. When the virtual water runs out: Local and global responses to addressing unsustainable groundwater consumption. *Water International* 47: 1060–1084. https://doi.org/10.1080/02508060.2023.2131272.

Haqiqi, Iman, Laura Bowling, Sadia Jame, Uris Baldos, Jing Liu, and Thomas Hertel. 2023. Global drivers of local water stresses and global responses to local water policies in the United States. *Environmental Research Letters* 18: 065007. https://doi.org/10.1088/1748-9326/acd269.

IPCC/OECD/IEA. 1997. *1996 IPCC guidelines for national greenhouse gas inventories*. Bracknell: IPCC/OECD/IEA UK Meteorological Office.

Lark, Tyler J., Nathan P. Hendricks, Aaron Smith, Nicholas Pates, Seth A. Spawn-Lee, Matthew Bougie, Eric G. Booth, Christopher J. Kucharik, and Holly K. Gibbs. 2022. Environmental outcomes of the US Renewable Fuel Standard. *Proceedings of the National Academy of Sciences* 119. Proceedings of the National Academy of Sciences: e2101084119. https://doi.org/10.1073/pnas.2101084119.

Liu, Jing, Thomas W. Hertel, Richard B. Lammers, Alexander Prusevich, Lantz C. Uris, Danielle S. Baldos, and Grogan, and Steve Frolking. 2017. Achieving sustainable irrigation water withdrawals: Global impacts on food security and land use. *Environmental Research Letters* 12: 104009. https://doi.org/10.1088/1748-9326/aa88db.

Maupin, Molly A., Joan F. Kenny, Susan S. Hutson, John K. Lovelace, Nancy L. Barber, and Kristin S. Linsey. 2014. *Estimated use of water in the United States in 2010*. Reston, VA: US Geological Survey. https://doi.org/10.3133/cir1405.

Neiken, Loganaden. 2003. FAO methodology for estimating the prevalence of undernourishment. In *International Scientific Symposium on Measurement and Assessment of Food Deprivation and Undernutrition*. Rome: Food and Agriculture Organization of the United Nations. http://www.fao.org/3/Y4249E/y4249e06.htm.

Siebert, Stefan, and Petra 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: 198–217. https://doi.org/10.1016/j.jhydrol.2009.07.031.

Taheripour, Farzad, Luis Pena-Levano, and Wally Tyner. 2017. *Introducing first and second generation biofuels into GTAP 9 Data Base*. Research Memorandum 29. Department of Agricultural Economics, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). 10.21642/GTAP.RM29.

USDA-NASS. 2019. *Value of crop sold per acre*. Washington, DC: US Department of Agriculture, National Agricultural Statistics Service. https://data.nal.usda.gov/dataset/nass-quick-stats. Accessed 15 Dec 2019.

West, Paul C., Holly K. Gibbs, Chad Monfreda, John Wagner, Carol C. Barford, Stephen R. Carpenter, and Jonathan A. Foley. 2010. Trading carbon for food: Global comparison of carbon stocks vs crop yields on agricultural land. *Proceedings of the National Academy of Sciences* 107: 19645–19648. https://doi.org/10.1073/pnas.1011078107.