SIMPLE-G Model and Database for US Groundwater

Local Groundwater Sustainability Policies and Global Spillovers

Chapter 12 of the SIMPLE-G Book

Iman Haqiqi[1]

October 12, 2024

Abstract

Recent computational advances have enabled economic analysis at fine spatial resolutions, known as gridded economic analysis. This approach incorporates numerous spatially heterogeneous locations, each with distinct economic and biophysical characteristics while accounting for interactions across input and output markets. This document presents the model, database, and other inputs required for SIMPLE-G-US-Allcrops, a multi-scale gridded economic model with gridded United States. This manuscript shows how to use GEMPACK to run the groundwater sustainability scenarios in SIMPLE-G. It also describes model codes, input data, parameters, and command files.

Keywords: computational economics, gridded economic model, sustainability, food security, agriculture, water resources, equilibrium, system of equations, policy.

Acknowledgments

The model introduced here is a product of long-term collaborations. The idea behind the SIMPLE-G model is from Tom Hertel and Uris Baldos. Mark Horridge created the first pilot code in TABLO. His suggestions and guidance were invaluable in the computation of the model. Uris Baldos, Jing Liu, and Iman Haqiqi initiated the first SIMPLE-G models and databases. Multiple grants supported the development of the model introduced here. I acknowledge support from the United States Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modelling, MultiSector Dynamics Contract DE-SC0022141; the National Science Foundation HDR award # 2118329: "NSF Institute for Geospatial Understanding through an Integrative Discovery Environment (I-GUIDE); and 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".

[1] Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University
Mailing address: 403 Mitch Daniels Blvd., West Lafayette, Indiana 47907-2056 USA
Email: ihaqiqi@purdue.edu

Code and data availability

The SIMPLE-G model introduced here is an open-source model used in Chapter 12 of the SIMPLE-G Book. The model code files, input datasets, parameters, and command files are available at: https://www.gtap.agecon.purdue.edu/simple-g/

Useful resources

- SIMPLE Book: https://doi.org/10.1007/978-3-319-22662-0
- SIMPLE-G Book: https://doi.org/10.1007/978-3-031-68054-0
- SIMPLE-G Webpage: https://www.gtap.agecon.purdue.edu/simple-g/
- SIMPLE-G YouTube Videos: https://tinyurl.com/SIMPLE-G-Book
- SIMPLE-G Short Course: https://mygeohub.org/courses/sustainability_shortcourse_2024
- SIMPLE-G visualization:
  - Analyze SIMPLE-G Results in Jupyter Notebook and R: https://tinyurl.com/simpleg-handout
  - SIMPLE-G Web Application: https://mygeohub.org/resources/simpleus
  - SIMPLE-G Online: Computation: https://mygeohub.org/resources/simplegus

Table of Contents

Abstract 1
Acknowledgments 1
Code and data availability 2
Useful resources 2
1 Introduction 5
1.1 Background 5
1.2 SIMPLE-G model dimension 5
1.2.1 Spatial resolution and scope 5
1.2.2 Time steps and reference year 5
1.2.3 Commodity aggregation 6
1.2.4 Land types 6
1.2.5 Water sources 6
1.3 Model specifications 6
1.3.1 Economic water supply 6
1.3.2 Economic water demand 6
1.3.3 Role of irrigation technology 6
1.3.4 Role of irrigation extent 6
1.3.5 Role of irrigation yield gap 6
1.4 Potential applications 7
1.4.1 SIMPLE-G for hydrology and water management 7
1.4.2 SIMPLE-G and agricultural sciences 7
1.4.3 SIMPLE-G and environmental sciences 7
1.4.4 Coupling human and environmental systems 7
1.5 Data access ..... **Error! Bookmark not defined.**
1.6 Software requirement 7
1.7 Operating systems 7
2 Model files for running simulations 8
2.1 TAB file for variables and equations 8
2.2 EXE file as pre-compiled executable 8
2.3 Other files 8
3 Data files 9
3.1 GRIDDATA includes the reference value of variables 9
3.2 GRIDPARM includes the value of behavioral parameters 10
3.3 GRIDSETS includes set elements and mappings 11
3.4 Model dimensions ..... **Error! Bookmark not defined.**
4 CMF command files 11
4.1 Overall options 11
4.2 Section 1: paths for old and new files 12
4.3 Section 2: model closure 12
4.4 Section 3: solution method 13
4.5 Section 4: experiment design and shocks 14
5 Run the SIMPLE-G model and view the results 14
5.1 Run the model with a batch file 14
5.2 View solution files: SL4 and TXT 15
5.3 Subtotals: projection example 16
5.4 Land conversion: conservation scenario 17
6 Discussion and conclusions 17
Reference 18

List of Boxes

Box 1. List of files in the main folder of SIMPLEG 8
Box 2. List of data files for SIMPLE-G-US-Allcrops 9

Box 3. List of command files for SIMPLE-G-US-Allcrops 11
Box 4. Options in the CMF file for SIMPLE-G-US-Allcrops 12
Box 5. List of options for CMF files 12
Box 6. Introducing files in the CMF file for SIMPLE-G-US-Allcrops 12
Box 7. Closure in the CMF file for SIMPLE-G-US-Allcrops 13
Box 8. Solution methods in the CMF file for SIMPLE-G-US-Allcrops 14
Box 9. Simulation scenario in the CMF file for SIMPLE-G-US-Allcrops 14
Box 10. List of batch files for SIMPLE-G-US-Allcrops 14
Box 11. Content of a batch file used for SIMPLE-G-US-Allcrops 15
Box 12. List of output files for test simulation for SIMPLE-G-US-Allcrops 15

List of Figures

Figure 1. Screenshot of the headers available in GRIDDATA.har 10
Figure 2. Screenshot of the headers available in GRIDPARM.har 10
Figure 3. Screenshot of home screen of SL4 file in ViewSOL 16
Figure 4. Projecting regional crop production and decomposing the drivers 17
Figure 5. The impacts of US groundwater sustainability on crop area and production. 17

1 Introduction

The SIMPLE Model, the Simplified International Model of agricultural Prices, Land use and the Environment, was developed for global analysis of sustainability and food security (Hertel and Baldos, 2016). Its gridded version, SIMPLE-G, is a model designed for multi-scale gridded economic analysis with focus on land, water, labor, and the environment. As detailed in the SIMPLE-G Book (Haqiqi and Hertel, 2024), the model captures interactions across regions, providing insights into resource use and policy impacts at fine spatial resolutions.

1.1 Background

The model described here is an extension of the SIMPLE-G1 (Liu et al., 2017), SIMPLE-G-US (Baldos et al., 2020), and SIMPLE-G Global (Haqiqi et al., 2022). This manuscript describes the model codes and data files for SIMPLE-G-US Allcrops version used in Chapter 12 of the SIMPLE-G book. The model follows full production structure as introduced in Chapter 5 of the SIMPLE-G book without explicit modeling of the labor market.

Table 1. SIMPLE-G-US-Allcrops Model specifications

Model name	simpleg_us_allcrops
Grids	US 5 arc-min
Terrestrial coverage	All agricultural land
Year	Around 2010
Version	v21p1
Crop details	Allcrops composite
Production technology	Irrigated and rainfed
Water module	Groundwater, surface water, irrigation equipment
Labor module	Not available
Land module	Quantity preserving constant elasticity of transformation (QCET)
Fertilizer module	Regional market
Product differentiation	Sub-regional Armington

1.2 SIMPLE-G model dimension

This section describes the spatial resolution, temporal horizon, and commodity aggregation details of the SIMPLE-G-US-Allcrops. For this SIMPLE-G-US Allcrops 2010 version, we have:

- REG: 16 regions.
- GID: 75,666 grid cells
- SMKT: 24 (15 regions + 9 US Farm Resource Regions) sub-regions.

1.2.1 Spatial resolution and scope

SIMPLE-G operates on a global scale, employing a spatial resolution that varies according to the specific application. For this particular application, a US gridded approach is applied, dividing the Continental United States into 5 arc-minute cells. Each 5 arc-minute cell is approximately 6 to 8 kilometers on each side. Other world regions are aggregated into 15 global regions.

1.2.2 Time steps and reference year

The model is comparative static, illustrating the change from one economic equilibrium state to a new economic equilibrium state. It can simulate both historical and future

scenarios. The model uses 2010 as the reference baseline for simulations.

1.2.3 Commodity aggregation

SIMPLE-G includes markets for crops, livestock, and processed food. Commodities are often aggregated into groups based on their production and consumption patterns. This aggregation level can vary depending on the specific analysis. For this version, all crops are aggregated into one composite crop, which varies by grid cell. Composite crops are differentiated by the production location and irrigation vs rainfed technology.

1.2.4 Land types

The model distinguishes between different land types, such as irrigated and rainfed cropland. This allows for the analysis of cropland conversion and its impact on agriculture and the environment.

1.2.5 Water sources

SIMPLE-G incorporates the decisions about water resources, considering both surface water and groundwater in combination with an index for irrigation technology.

1.3 Model specifications

As the model is used for groundwater sustainability policy assessment, it is important to look at how water demand and supply is considered in the model. Note that economic water supply is linked to hydrological supply, but they are different concepts. Similarly, economic water demand is different from agronomic crop water demand.

1.3.1 Economic water supply

Economic water supply shows how the supply of water can change in response to change in the value of water for each grid cell. This can reflect the costs of water extraction and transfer.

1.3.2 Economic water demand

Economic water demand shows how the demand for irrigation water changes in response to change in value of water relative to other production inputs. Specifically, this is important to consider relative costs of different water sources as well as cropland rents and return.

1.3.3 Role of irrigation technology

In case of increase in pumping costs or reducing water availability, an improvement in irrigation technology can buffer the negative impacts. However, this require considering the current conditions of irrigation technology for each location.

1.3.4 Role of irrigation extent

Irrigated area has a direct relationship with water demand. However, there is a substitution possibility between water and land.

1.3.5 Role of irrigation yield gap

The value of water depends on irrigation yield gap and average price of crops produced. Comparing with the irrigation costs, farmers may decide to increase or decrease the irrigation practice.

1.4 Potential applications

While the SIMPLE-G model is an economic model, it integrate data and insights with various scientific disciplines. By incorporating biophysical inputs and generating economic decision outputs that are relevant to other sciences, SIMPLE-G offers a framework for analyzing interactions between economic and environmental systems.

1.4.1 SIMPLE-G for hydrology and water management

For instance, the model can assess the hydrological implications of economic policies and environmental changes. By simulating changes in water use and irrigation patterns, SIMPLE-G provides insights about change in cropland area and irrigation technology that can be used for hydrological modeling.

1.4.2 SIMPLE-G and agricultural sciences

Similarly, in the of agricultural sciences, the model can evaluate the impact of policies and climate change on crop yields, land use, and fertilizer application. These outputs can be used to assess potential impacts on water quality, soil health, and environmental pollutions (Liu et al., 2023).

1.4.3 SIMPLE-G and environmental sciences

In the field of environmental science, SIMPLE-G can help quantify the environmental consequences of economic decisions. By simulating changes in land use and water use, the model can help assess the impact on biodiversity, ecosystem services, and other environmental indicators. This information can inform sustainable land-use planning and policy development.

1.4.4 Coupling human and environmental systems

The integration of human and environmental systems is critical for study of sustainable development. SIMPLE-G can be coupled with other models, such as Water Balance Model, to explore the bidirectional feedback between human activities and the environment (Woo et al., 2022).

1.5 Software requirement

The model is solved in GEMPACK, the General Equilibrium Modelling PACKage (Horridge et al., 2018). The input data is prepared in a format that is efficient for use in GEMPACK. For simulating SIMPLE-G, you need to have access to a machine with the software or install the GEMPACK software on your computer. A free-trail version of GEMPACK is available at https://www.copsmodels.com/gpeidl.htm

1.6 Operating systems

The model is tested on the Windows systems. For MacOS, users may try to install GEMPACK as a container or with a virtual machine with several third-party options:

- CrossOver (https://www.codeweavers.com/crossover/)
- Virtual Box (https://www.virtualbox.org/wiki/Downloads)
- Parallels Desktop (http://www.parallels.com/products/desktop/)
- VMWare Fusion (http://www.vmware.com/products/fusion.html)

For the latest updates on the Linux version, please look at:

- The Centre of Policy Studies: https://www.copsmodels.com/,
- Installation and Use on Linux PCs: https://www.copsmodels.com/gpdocl.htm,
- Demonstration Version for Linux: https://www.copsmodels.com/gpdemol.htm

The remaining of this document contains three sections.

- Section 2 provides an overview of files for the SIMPLE-G model.
- Section 3 introduces the data and parameters for gridded US and Allcrops.
- Section 4 describes the command files and model closures.
- Section 5 described the solution files and results.

2 Model files for running simulations

The "simpleg_us_allcrops.zip" file includes model data files and codes. Under the main folder of SIMPLE-G, users will find the following files needed to run the executable "simpleg.exe" program.

Box 1. List of files in the main folder of SIMPLEG

```
simpleg.tab
simpleg.exe
simpleg.axs
simpleg.axt
simpleg.inf
simpleg.LOG
simpleg.min
simpleg.cls
simpleg-ceqexplen.har
gfortran_link_ltg.fig
LTG.LOG
parcomp.log
```

2.1 TAB file for variables and equations

The file "simpleg.tab" provides the original script containing variable names, parameter labels, relevant headers in SIMPLE-G data source, and equation definitions in TABLO language for use in GEMPACK.

2.2 EXE file as pre-compiled executable

The "simpleg.exe" file is the pre-compiled executable model for running simulations. Modification of the equations and variables in the TABLO code requires compiling a new model version.

2.3 Other files

All other files are automatically generated when the model is compiled using GFortran.

3 Data

The files for running the simulations with SIMPLE-G-US-Allcrops are available in "simpleg_us_allcrops.zip" archive on the SIMPLE-G webpage. Some data layers are available for general users in txt and raster formats. The gridded quanity of the composte allcrops is a price-adjusted quantity index based on corn-equivalent calculations (Haqiqi et al., 2023a). The value of water for US grid cells is estimated based on irrigation yield gap and available on MyGeoHub (Haqiqi, 2023a). The data on land supply elasticity is empirically estimated using county level information in the US (Villoria et al., 2023). The data needed to define groundwater sustainability policy is obtained using the withdrawal to recharge ratio (Haqiqi, 2023b).

3.1 Files

The SIMPLE-G model data inputs are organized in three main files. Each simulation may contain additional files depending on the model extensions or scenario experiments. The folder "in" contains the input data, stored as header array files (*.har). As their names suggest, GRIDDATA, GRIDPARM, and GRIDSET contain the data, parameters and sets of the model, respectively. These files provide the key information for conducting simulations with SIMPLE-G.

The folder "shf" contains the shock files as "USERDATA.har" which includes data for determining the experiment and scenarios. And they can be used in experiments with the command file determining what variables to be mounted.

Box 2. List of data files for SIMPLE-G-US-Allcrops

```
in\
       GRIDDATA.har
       GRIDPARM.har
       GRIDSETS.har
shf\
       USERDATA.har
```

If GEMAPCK is installed, users can open and view these "*.HAR" files using "viewHAR.exe" which is one of the utilities available in GEMPACK. The data is stored in headers. In addition to main values, each header includes the following metadata:

- Header: 4-character label
- Type: real, integer, string
- Dimension: either in the form of numbers or embedded set labels
- Coefficient: the long coefficient label of the parameter, variable, or set
- Name: description of the header
- Total, min, max, etc.: Numerical summaries of all data in each header

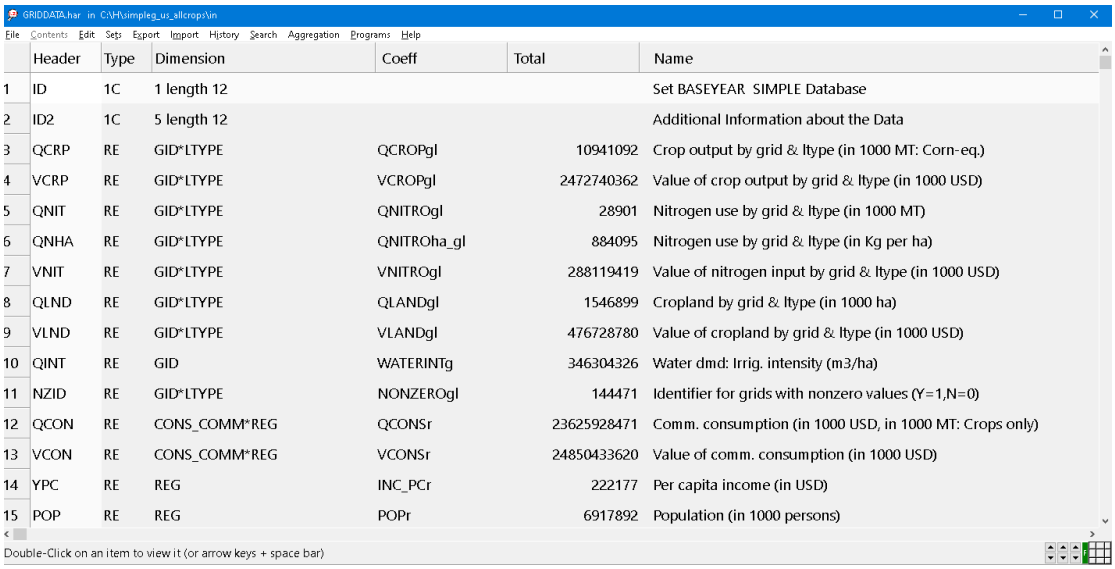
3.2 GRIDDATA includes the reference value of variables

The "GRIDDATA.har" file includes regional and gridded information to show the value of the variables for the reference year (2010 for this version). These variables are related to production, demand, and supply of agricultural commodities as well as other relevant information such as consumption, exports, imports, income, population, etc.

Figure 1 is the screenshot showing some of the headers in this file. Here, ID includes the reference year, while ID2 includes information about the module name and versions. More details about the dimensions of these variables (GID, TYPE, REG, etc.) are available in the "GRIDSETS.har".

3.3 GRIDPARM includes the value of behavioral parameters

The "GRIDPARM.har" file includes the value of the parameters of the model. These are mostly elasticity parameters. The parameters are those components of the model that are not allowed to change during the simulation, contrary to variables that are allowed to change.



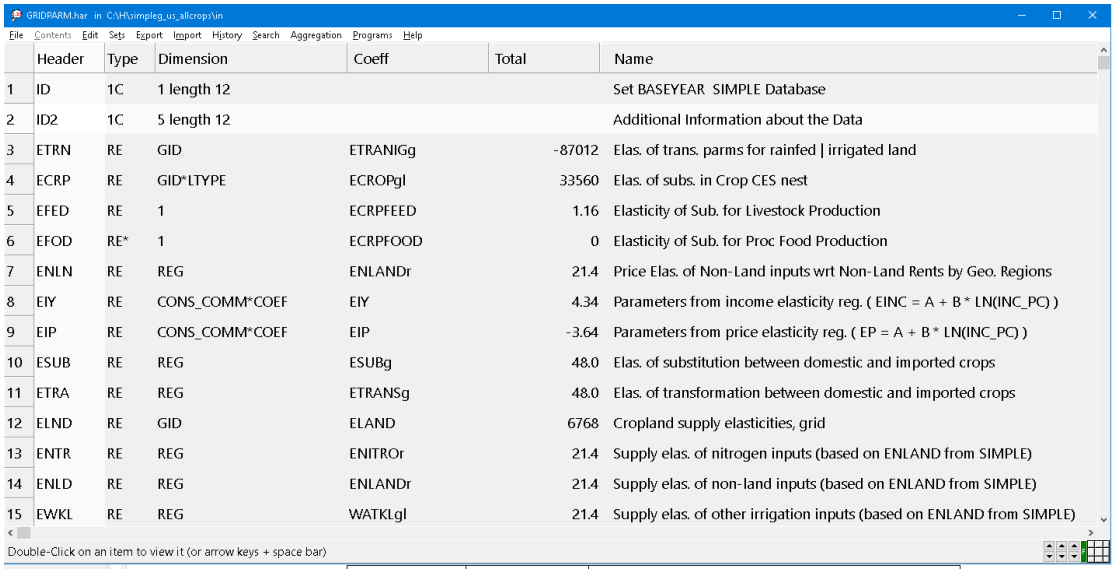
GRIDDATA.har in C:\H\simpleg_us_allcrops\in

File Contents Edit Sets Export Import History Search Aggregation Programs Help

	Header	Type	Dimension	Coeff	Total	Name
1	ID	1C	1 length 12			Set BASEYEAR SIMPLE Database
2	ID2	1C	5 length 12			Additional Information about the Data
3	QCRP	RE	GID*LTYPE	QCROPgl	10941092	Crop output by grid & ltype (in 1000 MT: Corn-eq.)
4	VCRP	RE	GID*LTYPE	VCROPgl	2472740362	Value of crop output by grid & ltype (in 1000 USD)
5	QNIT	RE	GID*LTYPE	QNITROgl	28901	Nitrogen use by grid & ltype (in 1000 MT)
6	QNHA	RE	GID*LTYPE	QNITROha_gl	884095	Nitrogen use by grid & ltype (in Kg per ha)
7	VNIT	RE	GID*LTYPE	VNITROgl	288119419	Value of nitrogen input by grid & ltype (in 1000 USD)
8	QLND	RE	GID*LTYPE	QLANDgl	1546899	Cropland by grid & ltype (in 1000 ha)
9	VLND	RE	GID*LTYPE	VLANDgl	476728780	Value of cropland by grid & ltype (in 1000 USD)
10	QINT	RE	GID	WATERINTg	346304326	Water dmd: Irrig. intensity (m3/ha)
11	NZID	RE	GID*LTYPE	NONZEROgl	144471	Identifier for grids with nonzero values (Y=1,N=0)
12	QCON	RE	CONS_COMM*REG	QCONSr	23625928471	Comm. consumption (in 1000 USD, in 1000 MT: Crops only)
13	VCON	RE	CONS_COMM*REG	VCONSr	24850433620	Value of comm. consumption (in 1000 USD)
14	YPC	RE	REG	INC_PCr	222177	Per capita income (in USD)
15	POP	RE	REG	POPr	6917892	Population (in 1000 persons)

Double-Click on an item to view it (or arrow keys + space bar)

Figure 1. Screenshot of the headers available in GRIDDATA.har



GRIDPARM.har in C:\H\simpleg_us_allcrops\in

File Contents Edit Sets Export Import History Search Aggregation Programs Help

	Header	Type	Dimension	Coeff	Total	Name
1	ID	1C	1 length 12			Set BASEYEAR SIMPLE Database
2	ID2	1C	5 length 12			Additional Information about the Data
3	ETRN	RE	GID	ETRANIGg	-87012	Elas. of trans. parms for rainfed | irrigated land
4	ECRP	RE	GID*LTYPE	ECROPgl	33560	Elas. of subs. in Crop CES nest
5	EFED	RE	1	ECRPFEED	1.16	Elasticity of Sub. for Livestock Production
6	EFOD	RE*	1	ECRPFOOD	0	Elasticity of Sub. for Proc Food Production
7	ENLN	RE	REG	ENLANDr	21.4	Price Elas. of Non-Land inputs wrt Non-Land Rents by Geo. Regions
8	EIY	RE	CONS_COMM*COEF	EIY	4.34	Parameters from income elasticity reg. (EINC = A + B * LN(INC_PC))
9	EIP	RE	CONS_COMM*COEF	EIP	-3.64	Parameters from price elasticity reg. (EP = A + B * LN(INC_PC))
10	ESUB	RE	REG	ESUBg	48.0	Elas. of substitution between domestic and imported crops
11	ETRA	RE	REG	ETRANSg	48.0	Elas. of transformation between domestic and imported crops
12	ELND	RE	GID	ELAND	6768	Cropland supply elasticities, grid
13	ENTR	RE	REG	ENITROr	21.4	Supply elas. of nitrogen inputs (based on ENLAND from SIMPLE)
14	ENLD	RE	REG	ENLANDr	21.4	Supply elas. of non-land inputs (based on ENLAND from SIMPLE)
15	EWKL	RE	REG	WATKLgl	21.4	Supply elas. of other irrigation inputs (based on ENLAND from SIMPLE)

Double-Click on an item to view it (or arrow keys + space bar)

Figure 2. Screenshot of the headers available in GRIDPARM.har

3.4 GRIDSETS includes set elements and mappings

The "GRIDSETS.har" file includes the metadata and elements of the sets that are used in the model. For example, the set "REG" defines a domain of the regions included in the model as shown in Table 2.

In addition, this file includes "mapping" information that shows how a subset of elements from one set are linked to an element of another set. For example, mapping MAP1 maps set GID to REG. Each row of MAP1 represents the corresponding grid cell in GID, and it records the region (by its index) to which the grid cell is mapped.

Note that LON and LAT show the coordinates 120 x longitude or 120 x latitude for the center of grid cells at 5 arc-min size following "coord. ref. lon/lat WGS 84".

Table 2. Regions in SIMPLE-G

No.	REG	Description
1	E_Euro	Eastern Europe
2	N_Afr	North Africa
3	SSA	Sub Saharan Africa
4	S_Amer	South America
5	AUS_NZ	Oceania
6	EU	European Union
7	S_Asia	South Asia
8	CC_Amer	Central America
9	S_Afr	Southern Africa
10	SE_Asia	Southeast Asia
11	CAN	Canada
12	US	USA
13	CHN_MNG	China and Mongolia
14	M_East	Middle East
15	JPN_KR	Japan / Korea
16	C_Asia	Central Asia

4 CMF command files

The folder "CMF" contains command files (*.cmf). They contain important information for simulation, including the path of input and output files, parameters for model solving, closure of the current model, and shock (policy levers) to be changed in the experiment. Each experiment is stored in a separate command file. For this exercise, there are two main CMF files:

Box 3. List of command files for SIMPLE-G-US-Allcrops

```
cmf\
       simpleg_water_ssp2_erl23.cmf
       simpleg_water_gsp4_erl23.cmf
```

4.1 Overall options

The CMF file usually starts with some comments followed by general options. For

example, the command below will set the initial memory allocation, which can vary by model version, and users can update it after changing the model.

Box 4. Options in the CMF file for SIMPLE-G-US-Allcrops

```
! Options
! -----------
 start with mmnz = 2429034;
! -------------------
```

There are other options that users may decide to use. For a list of "XXX=yes|no" commands, look at GEMPACK manual.

Box 5. List of options for CMF files

```
CPU = yes|NO ; ! yes means report CPU times
EAA = yes|no ; ! yes means echo activity to screen
NRP = yes|NO ; ! yes means do not re-use pivots
NEQ = yes|NO ; ! yes means do no equations, only formulas
```

4.2 Section 1: paths for old and new files

The CMF file determines the name and location of the input files (data, sets, parameters). It also determines the name and location that output files should be saved. The output files are mainly LOG file, SL4 file, and updated GRIDDATA.

Box 6. Introducing files in the CMF file for SIMPLE-G-US-Allcrops

```
! Input files
! -----------
 File GRIDSETS = in\GRIDSETS.har ; !  sets and mapping information
 File GRIDDATA = in\GRIDDATA.har ; !  base data: gridded and regional
 File GRIDPARM = in\GRIDPARM.har; ! parameter data

! Output files
! -----------
 LOG file      = out\simpleg_water_ssp2_erl23.LOG;
 Solution file = out\simpleg_water_ssp2_erl23;
 Updated file GRIDDATA = out\simpleg_water_ssp2_erl23.upd ;
 condensation information file = out\simpleg_water_ssp2_erl23-cond.txt
```

4.3 Section 2: model closure

Model closure refers to the process of determining which variables in the model are exogenous and which are endogenous in response to shocks or changes in policy. This defines the boundaries of changes and the scope of its analysis. Exogenous variables are variables that are assumed to be fixed or predetermined and do not change in response to changes in other variables within the model. Examples include policies, technological progress, and global population growth. Endogenous variables are variables that are allowed to vary freely within the model in response to changes in exogenous variables. Examples include prices, production levels, consumption patterns, and trade flows. Closure rules also include specific assumptions that determine which model setup to be used. Common closure rules in SIMPLE-G are:

- Integrated market closure.
- Mini model closure to fix crop prices.
- Input spatial mobility closure to select input supply equations.
- Input local allocation closure to determine the allocation to different practices.

Different closure rules can lead to different model outcomes.

4.4 Section 3: solution method

Users can determine the solution methods, steps, and sub-intervals in the CMF file. The solution methods include:

- Johanson Method: It is used for solving linearized versions of a model, where changes are calculated around a known equilibrium.
- Midpoint Method: Involves solving at the midpoint of each step for higher accuracy in approximating changes.
- Euler Method: A straightforward numerical approach that solves models by approximating derivatives through finite steps.
- Gragg's Method: An advanced, extrapolation-based approach that improves the accuracy of solutions by using Richardson extrapolation. It is often employed for more precise results.

Box 7. Closure in the CMF file for SIMPLE-G-US-Allcrops

```
! Closure
! -------
!                    KEY EXOGENOUS VARIABLES
Exogenous   p_INC_PC     ;!per capita income by inc. reg.
Exogenous   p_POP        ;!population by inc. reg.
Exogenous   p_QCRPBIOF   ;!crop demand for global biofuel use
Exogenous   s_QCRPBIOFr  ;!crop demand for biofuel use | alternative

Exogenous   p_P(NFOOD_COMM,REG) ;!price of non-food commodities by inc.
Exogenous   p_PNCRPFEED ;!price of non-feed inputs used in livestock pro
Exogenous   p_PNCRPFOOD ;!price of non-crop inputs used in proc. food pr

!                     Livestock Sector
Exogenous   p_AOCRPFEED    ;!hicks-neutral eff. index in livestock prod.
Exogenous   p_AFCRPFEEDW   ;!sub-component of feed eff. index by inc. reg
Exogenous   p_AFCRPFEEDr   ;!sub-component of feed eff. index by inc. reg
Exogenous   p_AFNCRPFEEDW ;!sub-component of non-feed eff. index by inc.
Exogenous   p_AFNCRPFEEDr ;!sub-component of non-feed eff. index by inc.

!                     Processed Food Sector
Exogenous   p_AOCRPFOOD    ;!hicks-neutral eff. index in proc. food prod.
Exogenous   p_AFNCRPFOODW ;!sub-component of eff. index of non-crop inpu
Exogenous   p_AFCRPFOODW   ;!sub-component of eff. index of crops in proc
Exogenous   p_AFCRPFOODr   ;!sub-component of eff. index of crops in proc
Exogenous   p_AFNCRPFOODr ;!sub-component of eff. index of non-crop inpu
```

Box 8. Solution methods in the CMF file for SIMPLE-G-US-Allcrops

```
 ! Verbal description and solution method
 ! ------------------
Verbal Description = GRDwaterProjection;
Method = Gragg;
Steps = 2 4 6;
subintervals = 3;

 automatic accuracy = NO;
 accuracy figures = 2;
 accuracy percent = 80;
 minimum subinterval length =  0.001;
 minimum subinterval fails = stop;
 accuracy criterion = both;
```

4.5 Section 4: experiment design and shocks

Experiment design involves introducing changes to "exogeneous variables" in the model as determined by the closure. Users may apply multiple changes at a time. Figure xxx shows a scenario that involves change in population (p_POP), income per capita (p_INC_PC), technological improvement in crop production (p_AOCROPr), biofuel mandate (p_QCROPBIOF) and other change.

Box 9. Simulation scenario in the CMF file for SIMPLE-G-US-Allcrops

```
! ************** !
! Shocks Section !
! ************** !


Shock p_POP = file shf\shocks_ERL.har header "POP"; !  from GEC paper
Shock p_INC_PC = file shf\shocks_ERL.har header "INC"; !  from GEC paper
Shock p_AOCROPr = file shf\shocks_ERL.har header "TFPC"; !  from GEC paper
Shock p_QCRPBIOF = file shf\shocks_ERL.har header "QBIO"; !
Shock p_AOCRPFEED = file shf\shocks_ERL.har header "TFPL"; ! Ludena et al.
```

5 Run the SIMPLE-G model and view the results

It is possible to run the model by opening the TAB file in "TABmate.exe" and clicking on "RUN CMF". This requires the CMF and TAB files to be in the same directory.

5.1 Run the model with a batch file

The "do_simpleg_*.bat" files are batch files that automatically run the model and specific simulations by double-clicking them, running in the command line, or using in other programs like R.

Box 10. List of batch files for SIMPLE-G-US-Allcrops

```
do_simpleg.bat
do_simpleg_gsp4.bat
do_simpleg_ssp2.bat
```

The batch file manages model simulations. This file can be opened with any text editor. The content of this file is similar to this Box:

Box 11. Content of a batch file used for SIMPLE-G-US-Allcrops

```
call simpleg -cmf simpleg_water_gsp4_erl23.cmf
pause
```

This command means:

- `call`: the command to run the compiled model as executable file (simpleg.exe)
- `simpleg`: the name of the executable file under the same folder, here it refers to "simpleg.exe"
- `-cmf`: it indicates the model should be simulated with command files
- `cmf\simpleg_water_gsp4_erl23.cmf`: this is the relative path of the command file.

The last line "pause" asks the console window to remain open after all simulations, for the convenience of users to check if simulations are conducted successfully or not. Users can also get the same report from log files (*.log) in the folder "out". For multiple simulations, it is possible to use separate batch files for each simulation or a combined call in one batch file.

5.2 View solution files: SL4 and TXT

After running the simulations, the out folder includes output files, log files, and updated data files for each simulation.

Box 12. List of output files for test simulation for SIMPLE-G-US-Allcrops

```
out\
     simpleg.log
     simpleg.sl4
     simpleg.slc
     simpleg.upd
     simpleg-cond.txt
```

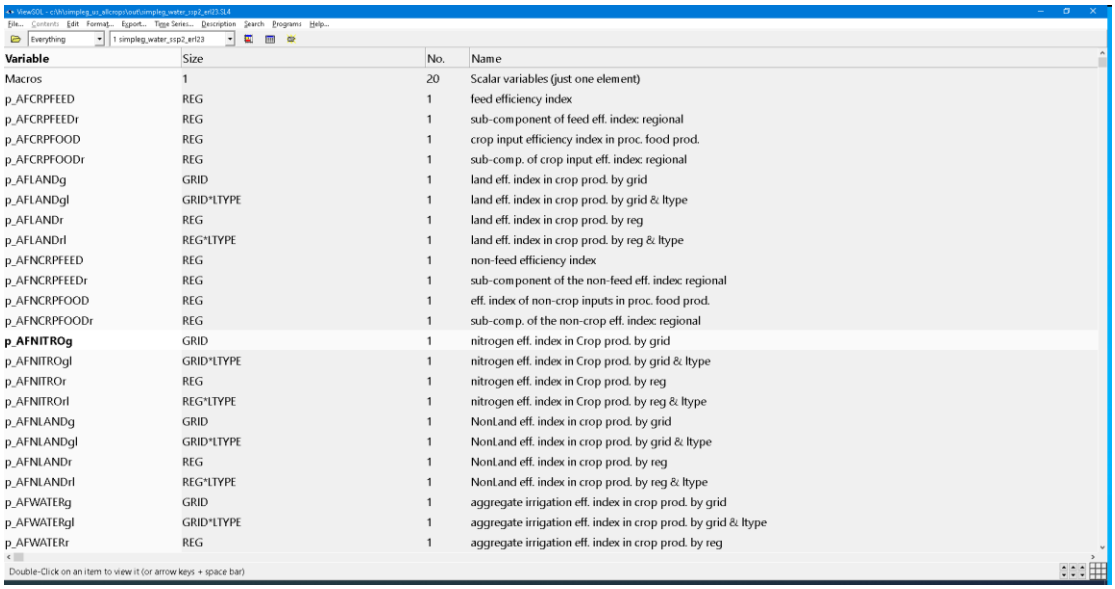
The typical output files for each simulation are:

- LOG file: Records the sequence of commands, errors, warnings, and reports from running the simulation. It provides a useful trail for debugging.
- SL4 file: Contains detailed solution results, including percentage changes in variables, pre- and post-simulation values, and subtotals if available.
- SLC/UDC file: Contains the values of the model's coefficients at the pre/post simulation levels.
- UPD file: Updated data file.
- TXT files: User-specified output files or condensation file. The condensation process reduces the size of the original model equations by substituting or eliminating redundant variables and equations (Baldos et al., 2020).

There are two general ways to see the results:

- Double-click to open the solution file with ViewSOL.exe (a tool installed together with GEMPACK), which will list variables and their descriptions for further checking.
- Access the solution file in R using the "HARr" package (Ivanic, 2023), which allows users to save simulation results as data frames with scripts for saving or further analysis.

For the results of each variable in the percentage form (starts with "p_"), the solution file will report the percentage of that variable in the first column (here 10 means 10% increase). If the variable also contains the level form (the quantity at base), the solution file will also report its value before the shock (the second column), after the shock (the third column), and the difference in net values, not percentage change (the fourth column). The easiest way of exporting simulation results is to "Export – Copy" in ViewSOL, which can be pasted to spreadsheets and other editors.

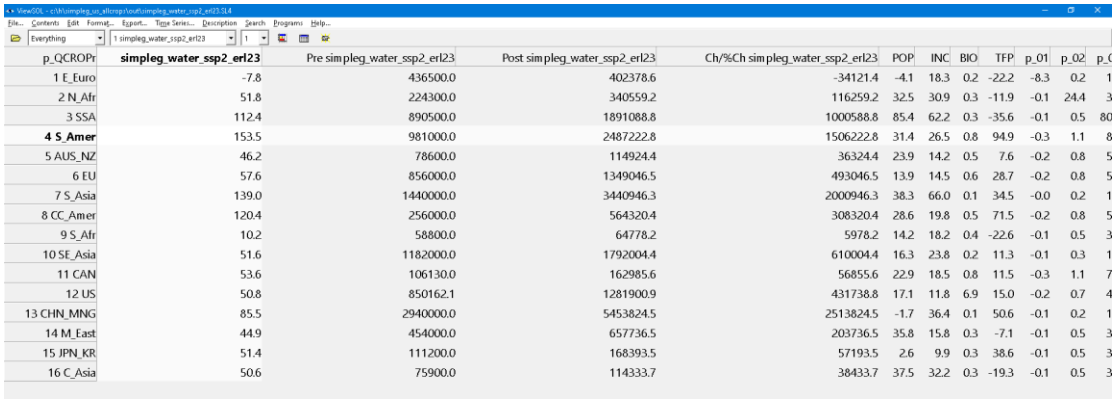


Variable	Size	No.	Name
Macros	1	20	Scalar variables (just one element)
p_AFCRPFEED	REG	1	feed efficiency index
p_AFCRPFEEDr	REG	1	sub-component of feed eff. indec regional
p_AFCRPFOOD	REG	1	crop input efficiency index in proc. food prod.
p_AFCRPFOODr	REG	1	sub-comp. of crop input eff. indec regional
p_AFLANDg	GRID	1	land eff. index in crop prod. by grid
p_AFLANDgl	GRID*LTYPE	1	land eff. index in crop prod. by grid & ltype
p_AFLANDr	REG	1	land eff. index in crop prod. by reg
p_AFLANDrl	REG*LTYPE	1	land eff. index in crop prod. by reg & ltype
p_AFNCRPFEED	REG	1	non-feed efficiency index
p_AFNCRPFEEDr	REG	1	sub-component of the non-feed eff. indec regional
p_AFNCRPFOOD	REG	1	eff. index of non-crop inputs in proc. food prod.
p_AFNCRPFOODr	REG	1	sub-comp. of the non-crop eff. indec regional
p_AFNITROg	GRID	1	nitrogen eff. index in Crop prod. by grid
p_AFNITROgl	GRID*LTYPE	1	nitrogen eff. index in Crop prod. by grid & ltype
p_AFNITROr	REG	1	nitrogen eff. index in Crop prod. by reg
p_AFNITRORl	REG*LTYPE	1	nitrogen eff. index in Crop prod. by reg & ltype
p_AFNLANDg	GRID	1	NonLand eff. index in crop prod. by grid
p_AFNLANDgl	GRID*LTYPE	1	NonLand eff. index in crop prod. by grid & ltype
p_AFNLANDr	REG	1	NonLand eff. index in crop prod. by reg
p_AFNLANDrl	REG*LTYPE	1	NonLand eff. index in crop prod. by reg & ltype
p_AFWATERg	GRID	1	aggregate irrigation eff. index in crop prod. by grid
p_AFWATERgl	GRID*LTYPE	1	aggregate irrigation eff. index in crop prod. by grid & ltype
p_AFWATERr	REG	1	aggregate irrigation eff. index in crop prod. by reg

Figure 3. Screenshot of home screen of SL4 file in ViewSOL

5.3 Subtotals: projection example

Figure 4 shows the results of the 2010-2050 growth scenario on % change in crop production. This data was used to generate Fig. 12.2 of SIMPLE-G Book. The POP, INC, BIO, and TFP are the subtotals showing the contribution of each driver in the final change of the variable following the decomposition method in GEMPACK (Harrison et al., 2000).

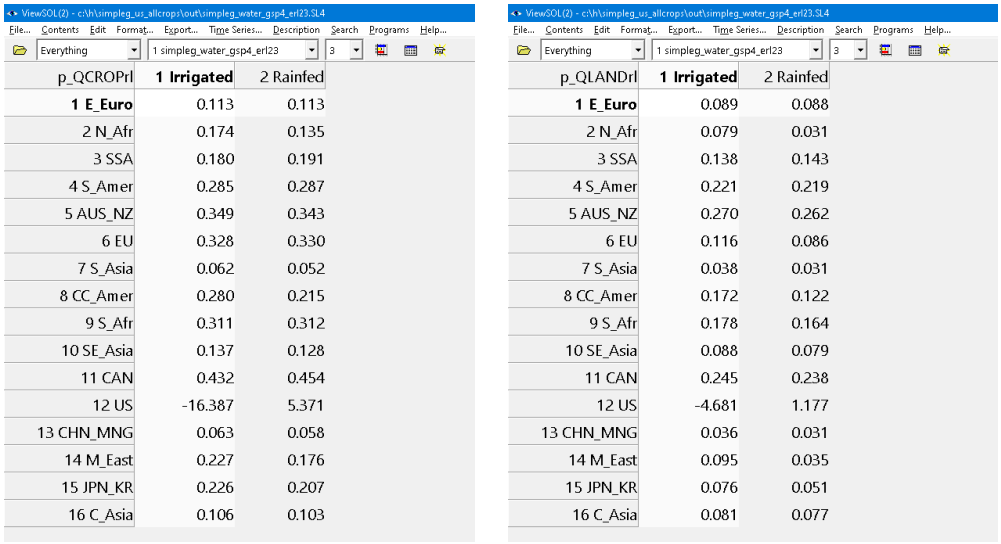


p_QCROPr	simpleg_water_ssp2_erl23	Pre simpleg_water_ssp2_erl23	Post simpleg_water_ssp2_erl23	Ch/%Ch simpleg_water_ssp2_erl23	POP	INC	BIO	TFP	p_01	p_02	p_0
1 E_Euro	-7.8	436500.0	402378.6	-34121.4	-4.1	18.3	0.2	-22.2	-8.3	0.2	1
2 N_Afr	51.8	224300.0	340559.2	116259.2	32.5	30.9	0.3	-11.9	-0.1	24.4	3
3 SSA	112.4	890500.0	1891088.8	1000588.8	85.4	62.2	0.3	-35.6	-0.1	0.5	80
4 S_Amer	153.5	981000.0	2487222.8	1506222.8	31.4	26.5	0.8	94.9	-0.3	1.1	8
5 AUS_NZ	46.2	78600.0	114924.4	36324.4	23.9	14.2	0.5	7.6	-0.2	0.8	5
6 EU	57.6	856000.0	1349046.5	493046.5	13.9	14.5	0.6	28.7	-0.2	0.8	5
7 S_Asia	139.0	1440000.0	3440946.3	2000946.3	38.3	66.0	0.1	34.5	-0.0	0.2	1
8 CC_Amer	120.4	256000.0	564320.4	308320.4	28.6	19.8	0.5	71.5	-0.2	0.8	5
9 S_Afr	10.2	58800.0	64778.2	5978.2	14.2	18.2	0.4	-22.6	-0.1	0.5	3
10 SE_Asia	51.6	1182000.0	1792004.4	610004.4	16.3	23.8	0.2	11.3	-0.1	0.3	1
11 CAN	53.6	106130.0	162985.6	56855.6	22.9	18.5	0.8	11.5	-0.3	1.1	7
12 US	50.8	850162.1	1281900.9	431738.8	17.1	11.8	6.9	15.0	-0.2	0.7	4
13 CHN_MNG	85.5	2940000.0	5453824.5	2513824.5	-1.7	36.4	0.1	50.6	-0.1	0.2	1
14 M_East	44.9	454000.0	657736.5	203736.5	35.8	15.8	0.3	-7.1	-0.1	0.5	3
15 JPN_KR	51.4	111200.0	168393.5	57193.5	2.6	9.9	0.3	38.6	-0.1	0.5	3
16 C_Asia	50.6	75900.0	114333.7	38433.7	37.5	32.2	0.3	-19.3	-0.1	0.5	3

Figure 4. Projecting regional crop production and decomposing the drivers

5.4 Land conversion: conservation scenario

Figure 5 shows the results of the groundwater sustainability scenario on land area and crop production by irrigated and rainfed practices. This information was used to create Fig. 12.6 of the SIMPLE-G Book.



p_QCROPrl	1 Irrigated	2 Rainfed
1 E_Euro	0.113	0.113
2 N_Afr	0.174	0.135
3 SSA	0.180	0.191
4 S_Amer	0.285	0.287
5 AUS_NZ	0.349	0.343
6 EU	0.328	0.330
7 S_Asia	0.062	0.052
8 CC_Amer	0.280	0.215
9 S_Afr	0.311	0.312
10 SE_Asia	0.137	0.128
11 CAN	0.432	0.454
12 US	-16.387	5.371
13 CHN_MNG	0.063	0.058
14 M_East	0.227	0.176
15 JPN_KR	0.226	0.207
16 C_Asia	0.106	0.103

p_QLANDrl	1 Irrigated	2 Rainfed
1 E_Euro	0.089	0.088
2 N_Afr	0.079	0.031
3 SSA	0.138	0.143
4 S_Amer	0.221	0.219
5 AUS_NZ	0.270	0.262
6 EU	0.116	0.086
7 S_Asia	0.038	0.031
8 CC_Amer	0.172	0.122
9 S_Afr	0.178	0.164
10 SE_Asia	0.088	0.079
11 CAN	0.245	0.238
12 US	-4.681	1.177
13 CHN_MNG	0.036	0.031
14 M_East	0.095	0.035
15 JPN_KR	0.076	0.051
16 C_Asia	0.081	0.077

Figure 5. The impacts of US groundwater sustainability on crop area and production.

6 Discussion and conclusions

This document provides an overview of the SIMPLE-G-US-Allcrops model and its associated databases for the 2010 reference year. It is a global, gridded economic framework designed to analyze water and land use, enabling “what-if” policy analyses and assessments of environmental and climate impacts. This manuscript is intended to serve as an introductory user guide for researchers seeking to employ the SIMPLE-G model in their studies.

Current model provides important insights about the impacts of environmental changes and polices on food production and prices. It also used for projecting the future stress on land and water resources. However, we recognize that future research would greatly benefit from modeling of labor wages and employment (Ray et al., 2023) that can affect

the effectiveness of conservation policies and food security. Furthermore, a more comprehensive food security analysis is necessary incorporating malnutrition metrics for offering more comprehensive health impact assessments and policy recommendations (Haqiqi et al., 2023b).

The SIMPLE-G model is instrumental to the Global-to-Local-to-Global framework, which highlights the importance of market-mediated effects on sustainability and food security—an area that deserves greater attention in global studies.

Reference

Baldos, U.L.C., Haqiqi, I., Hertel, T.W., Horridge, M., Liu, J., 2020. SIMPLE-G: A multiscale framework for integration of economic and biophysical determinants of sustainability. Environmental Modelling & Software 133, 104805. https://doi.org/10.1016/j.envsoft.2020.104805

Haqiqi, I., 2023a. The Value of Water in US Agriculture: integrating spatially and temporally heterogeneous hydroclimatic and economic data. https://doi.org/10.13019/9MXE-T280

Haqiqi, I., 2023b. A Gridded Dataset for Groundwater Sustainability Restriction Policy Scenarios for the Contiguous U.S. https://doi.org/10.13019/AHZR-4843

Haqiqi, I., Bowling, L., Jame, S., Baldos, U., Liu, J., Hertel, T., 2023a. A Gridded Price-Adjusted Quantity Index for Total Production of Crops for Irrigated and Rainfed Farms for the Contiguous U.S. https://doi.org/10.13019/RQ0D-JH17

Haqiqi, I., Grogan, D.S., Bahalou Horeh, M., Liu, J., Baldos, U.L.C., Lammers, R., Hertel, T.W., 2023b. Local, regional, and global adaptations to a compound pandemic-weather stress event. Environ. Res. Lett. 18, 035005. https://doi.org/10.1088/1748-9326/acbbe3

Haqiqi, I., Hertel, T.W. (Eds.), 2024. SIMPLE-G: A Gridded Economic Approach to Sustainability Analysis of the Earth's Land and Water Resources, 2025th edition. ed. Springer.

Haqiqi, I., Perry, C.J., Hertel, T.W., 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

Harrison, W.J., Horridge, J.M., Pearson, K.R., 2000. Decomposing Simulation Results with Respect to Exogenous Shocks. Computational Economics, Computational Economics 15, 227–249. https://doi.org/10.1023/A:1008739609685

Hertel, T.W., Baldos, U.L.C., 2016. Global Change and the Challenges of Sustainably Feeding a Growing Planet. Springer International Publishing, New York.

Horridge, J.M., Jerie, M., Mustakinov, D., Schiffmann, F., 2018. GEMPACK manual. Victoria University, Centre of Policy Studies/IMPACT Centre.

Ivanic, M., 2023. GEMPACK simulations in R: Ademonstration of running the GTAP modeland processing its results entirely in R using packages HARr and tabloToR. JGEA 8, 1–20. https://doi.org/10.21642/JGEA.080101AF

Liu, J., Bowling, L., Kucharik, C., Jame, S., Baldos, U.L.C., Jarvis, L., Ramankutty, N., Hertel, T.W., 2023. Tackling policy leakage and targeting hotspots could be key to addressing the 'wicked' challenge of nutrient pollution from corn production in the U.S. Environmental Research Letters 18, 105002. https://doi.org/10.1088/1748-9326/acf727

Liu, J., Hertel, T.W., Lammers, R.B., Prusevich, A., Baldos, U.L.C., Grogan, D.S., Frolking, S., 2017. Achieving sustainable irrigation water withdrawals: global

impacts on food security and land use. Environ. Res. Lett. 12, 104009. https://doi.org/10.1088/1748-9326/aa88db

Villoria, N., Liu, J., Haqiqi, I., Chakravarty, S., Delgado, M., Cisneros-Pineda, A., Hertel, T., 2023. Gridded Cropland Supply Elasticity for the Continental United States (5 arc-min Spatial Resolution). https://doi.org/10.13019/37T9-0E88

Woo, J., Zhao, L., Grogan, D.S., Haqiqi, I., Lammers, R., Song, C.X., 2022. C3F: Collaborative container-based model coupling framework, in: PEARC '22: Practice and Experience in Advanced Research Computing. Association for Computing Machinery, Boston, MA. https://doi.org/10.1145/3491418.3530298