



NEXOGENESIS
STREAMLINING WATER RELATED POLICIES

Deliverable 2.4

Socioeconomic data at grid level

Lead : Wageningen Research, WR

Date : 31/12/2023



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Project Deliverable

Project Number 101003881	Project Acronym NEXOGENESIS	Project Title Facilitating the next generation of effective and intelligent water-related policies, utilizing artificial intelligence and reinforcement learning to assess the water-energy-food-ecosystem (WEFE) nexus
Instrument: H2020RIA		Thematic Priority LC-CLA-14-2020
Title Socioeconomic data at grid level		
Contractual Delivery Date 31/12/2023		Actual Delivery Date 31/12/2023
Start Date of the project 01 September 2021		Duration 48 months
Organisation name of lead contractor for this deliverable WR		Document version 1
Dissemination level Internal		Deliverable Type Report with Excel files in data repository
Authors (organisations) Walter Rossi Cervi (WR), Diti Oudendag (WR), Wil Hennen (WR), Vincent Linderhof (WR)		
Reviewers (organisations) Janez Susnik (IHE), Antonio Trabucco (CMCC)		



Abstract

In this deliverable, we described the results of land use indicators projected at national level by G-RDEM that were downscaled to grid and river basin level using the framework of MagnetGrid. This document contains a technical description on how macro-economic indicators are connected with MagnetGrid and how MagnetGrid relate the macro-economic indicators with land specific biophysical indicators, which are the key proxies for simulating the land use dynamic patterns. The results of the downscaling framework are presented, as well as the implications on the SDMs.

Keywords

macro-economic modelling, land use, downscaling

Disclaimer



1. Introduction

Computable General Equilibrium (CGE) models are global macro-economic tools based on national accounts that are capable of assessing ex-ante impacts of global scenarios on climate and demographic changes, as well as WEFE (Water-Energy-Food-Ecosystem) policies at multiple scales. These models can have a large contribution to *Nexogenesis* as they combine various scenarios providing broader comprehension on the socio-economic impacts of climate change mitigation and adaptation measures on the different WEFE sectors.

In *Nexogenesis*, G-RDEM [1] is the CGE model employed to assess several WEFE related socio-economic indicators under different global scenarios of climatic and demographic changes. As a downside, G-RDEM results are generated at administrative level (NUTS0, but capable of producing results at NUTS2) for the European case studies and at national level for the South-African case study. As *Nexogenesis* has its case studies at river basin level, there are spatial mismatches between the results from G-RDEM and the actual study areas. Moreover, the simulation results of the several socio-economic indicators from G-RDEM are only available in percentage changes over time. At one hand, this gives more flexibility for assessing and integrating local socio-economic data provided by the case studies, but on the other hand provides little up to no interaction with indicators that contain high spatial variability (e.g. land use, land productivity, water use).

In this deliverable, we adapted a CGE downscaling method to G-RDEM in order to provide spatially explicit results on land use related indicators for the river basin case studies. This may provide more detailed input data for the System Dynamic Models (SDMs) used in this project. Therefore, the objective of this deliverable is to describe the results of the downscaling of land use related indicators derived from G-RDEM. To do so, we make use of MagnetGrid downscaling concept (described in section 2) [2], which is responsible to downscale land-economic indicators from MAGNET CGE model [3]. The workflow of activities for downscaling G-RDEM results is described in section 3, and the spatially explicit results are presented in section 4. In section 5, we discuss potential improvements and what can still be done before connecting with case studies SDMs. The first 3 Chapters were derived from the milestone MS16 of *Nexogenesis*,



2. MagnetGrid: a GTAP based land-economic downscaling tool

MagnetGrid is a model framework that simulates the spatial patterns of agricultural land use resulting from economic decisions on the use of land. It does so by combining future scenario-based projections on the supply, demand, prices and production costs of different agricultural commodities (as simulated by equilibrium models, such as MAGNET and G-RDEM) with spatially-explicit projections on the biophysical suitability (as simulated, for example, with gridded crop growth models such as LPJmL [4]) for agricultural production. Hence, MagnetGrid allows to project and visualize future agricultural land-use change patterns that emerge from climatic and socio-economic developments under a set of conditions that are specified in scenarios. It is able to explicitly simulate the effects of discontinuities such as the emergence of new land-use types (e.g. 2nd generation biofuel crops), the effects of policies affecting the economic performance of production systems (e.g. subsidy schemes, tax reductions/exemptions, removal of trade barriers), and the economic decisions leading to the adoption of innovative agricultural practices.

In its current configuration, MagnetGrid is able to downscale GTAP (Global Trade Analysis Project)-based [5] regional projections on the use of land for the production of agricultural commodities, and provide scenario-based map projections of agricultural land-use change, both at the global level and for dedicated case studies at the regional/country level. MagnetGrid applies a probabilistic allocation algorithm, according to which each unit of land (e.g. a regular grid cell) within a region is allocated to a percentage for each simulated land-use type (indicating the share of total area of the grid cell that is used by that land-use type), so that the scenario projections for total aggregated land claims in a region (e.g. as projected by MAGNET) are simultaneously fulfilled for all simulated land-use types. The configuration of the model is based on flexible templates, which allows for different scenario alternatives and configurations (e.g. combination of crop types into sectors, aggregation of countries into simulation regions) that can be seamlessly and efficiently accommodated. The original documentation of the model is published by Diogo et al. [2] (see Diogo et al. [6] for a detailed description of the theory), but an updated and improved global demo version of MagnetGrid is under preparation and it will be disclosed in a R package in 2023.

3. Downscaling G-RDEM results

3.1 Pre-processing raw G-RDEM results

The G-RDEM results files were provided by CAFoscari through the datafile *GRDEMOutput.gdx*, containing all the raw simulation results for the SSP4 scenario. This file can only be read and operated in GAMS (General Algebraic Modeling System) modelling platform and contains all the key variables encoded in the latest release GTAP version 7. As mentioned before, no major resources were needed to adapt and calculate the needed variables for land use downscaling.

For the demand of land and the volume production as a percentual change over time, compared to the baseline year of 2014, we need to calculate supply of and demand for commodities first from the G-RDEM results. The results are calculated for activities, regions and years which are reflected by the indices a , r and t respectively.

As we do not need all activities for downscaling, the index i reflects the selection of activities that are used in the downscaling exercise with MagnetGrid. The G-RDEM results are presented at the national level, so that the case studies with a river basin in one country have one region, and case studies with cross country river basins will have two or more regions.

There are >100 activities distinguished in G-RDEM.

The first indicators are the primary factor purchases by firms at basic prices (*EVFB*). It varies with region r , type of factor f , type of activity a , and year t . It is determined by prices of factor ($PF.l$) and volumes of factors ($XF.l$), see Eq. (1), where $PF.l$ and $XF.l$, are G-RDEM results. This is reflected by the “. l ” extension of the indicator names. The second indicator (*VDFB*) is the domestic purchases by firms at basic prices for activity i from the activity a , which vary per region r , and year t . It is determined by prices of domestically purchased factor ($PD.l$) and volumes of domestically purchased factors ($XD.l$), see Eq. (2). The third indicator (*VMFB*) is the imported purchases by firms at basic prices for activity i from the activity a , which vary per region r , and year t . It is determined by prices of imported factor ($PMT.l$) and imported volumes of factors ($XM.l$), see Eq. (3).

$$EVFB(r, f, a, t) = PF.l(r, f, a, t) * XF.l(r, f, a, t) \quad (1)$$

$$VDFB(r, i, aa, t) = PD.l(r, i, t) * XD.l(r, i, aa, t) \quad (2)$$

$$VMFB(r, i, aa, t) = PMT.l(r, i, t) * XM.l(r, i, aa, t) \quad (3)$$

In addition, the multi-production (“make”) matrix at supply and at basic prices are calculated. The multi-production matrix at supply prices in region r for activity a in year t (*MAKS*) is

determined by the supply prices $PP.l$ in region r for activity a in year t multiplied by the sum of production of activities $X.l$ in region r for activity a in year t over all the selected activities in r , see Eq. (4). The multi-production matrix at basic prices in region r for activity a in year t ($MAKS$) is determined by the sum product of basic prices $P.l$ and production of activities $X.l$ in region r for activity a in year t over all the selected activities in r , see Eq. (5).

$$MAKS(r, i, aa, t) = PP.l(r, a, t) * \sum_i X.l(r, a, i, t) \quad (4)$$

$$MAKB(r, i, aa, t) = \sum_i P.l(r, a, i, t) * X.l(r, a, i, t) \quad (5)$$

Calculating the demand for land for each activity, region and year

The demand for land and the production volume are expressed as the percentage change over time compared to the baseline year of 2014. Hence, for land demand we used $XF.l$ (i.e. GTAP = factor demand) and for the production volume $XD.l$ (i.e. GTAP = supply of domestic goods).

For all years after the year 2014 ($t = t00$), the demand for land area as an index $IndexLDEM$ is calculated based on Eq. (6):

$$IndexLDEM(r, a, t) = \left(\frac{XF.l(r, land, a, t)}{XF.l(r, land, a, t00)} - 1 \right) * 100 \quad (6)$$

With $XF.l(r, land, a, t00)$ being the demand for land area in the base year 2014. If the demand for land area for an activity is 0 in the base year, i.e. $XF.l(r, land, a, t00) = 0$, demand for land area will be set to 0 in all other years as well: $XF.l(r, land, a, t) = 0$.

In a similar way, the production volume $IndexProd$ is calculated. The demand for production volume $IndexProd$ is the sum of the factors for domestic purchase for an activity (XF) and domestic purchase of factors for production (XD) and imported volumes of factors (XM) for production of other activities:

$$IndexProd0(r, a, t) = \sum_f [XF.l(r, f, a, t) + \sum_i [XD.l(r, i, a, t) + XM.l(r, i, a, t)]] \quad (7)$$

For all years after the year 2014 ($t = t00$), the demand for land area as an index $IndexProd$ is calculated based on Eq. (8):

$$IndexProd(r, a, t) = \left(\frac{IndexProd0(r, a, t)}{IndexProd0(r, a, t00)} - 1 \right) * 100 \quad (8)$$

With $IndexProd(r, a, t00)$ being the demand for land area in the base year 2014. If the demand for production capacity for an activity is 0 in the base year, i.e. $IndexProd(r, a, t00) = 0$, demand for land area will be set to 0 in all other years as well: $IndexProd(r, a, t) = 0$.

The MagnetGrid exercise uses monetary indicators which are expressed in million USD. As the values in the database from G-RDEM (*GRDEMOutput.gdx*) are expressed in billions of USD, the indicators derived from G-RDEM were multiplied by a factor 1,000 to have the unit in *millions of USD* in line with the unit in GTAP database (Table 1).

Table 1. Conversions of G-RDEM indicators into MagnetGrid indicators

MagnetGrid indicators	Unit	Calculated as
Capital value(r, a, t)	<i>mn of USD</i>	$EVFB(r, f, a, t) * 1000$ for $f = Capital$
Land value(r, a, t)	<i>mn of USD</i>	$EVFB(r, f, a, t) * 1000$ for $f = Land$
Intermediate value(r, a, t)	<i>mn of USD</i>	$\sum_i [VDFB(r, i, a, t) + VMFB(r, i, a, t)] * 1000$
Value of labour(r, a, t)	<i>mn of USD</i>	$\sum_i [EVFB(r, i, a, t)] * 1000$ for $f = SKLab, UnskLab$
Production value(r, a, t)	<i>mn of USD</i>	$\sum_f [EVFB(r, f, a, t) + \sum_i [VDFB(r, i, a, t) + VMFB(r, i, a, t)]] * 1000$
Production tax	<i>mn of USD</i>	$(MAKB(r, a, t) * 1000) - Production\ value$
Land quantity	%	$IndexLand(r, a, t)$
Production quantity	%	$IndexProd(r, a, t)$

With indices for region r , activity a and year t . To calculate the capital value, the index of activities is equal to “Capital” in the G-RDEM results. To calculate the land value, the index of activities is equal to “Land” in the G-RDEM results. The value of labour is the sum of primary factor purchases by firms at basic prices for both skilled labour *SKLab* and unskilled labour *UnskLab*. Then we map the activities defined in G-RDEM model to activities in MagnetGrid, see Table 2.

Table 2. Conversion table of land-related activities in G-RDEM and sectors in MagnetGrid

GRDEM		MAGNET-GRID	
Name activity	Code activity i	Name Grid-sector	Code Grid-sector
Rice	pdr-a	Rice	Pdr
Wheat	wht-a	Wheat	Wht
Other grains	gro-a	Other grains	Gro
Oilseeds	osd	Oilseeds	Osd
Vegetables and fruit	v_f-a	Vegetables and fruit	v_f
Sugarcane and sugar beet	c_b-a	Sugarcane and sugar beet	c_b
Other crops	ocr-a	Other crops	Ocr
Plant based fibers	pfb	Plant based fibers	Pfb
Cattle	ctl-a	Cattle	Ctl
Raw milk	rmk-a	Cattle	Ctl
Wool	wol-a	Cattle	Ctl
Forestry	frs-a	Forestry	Frs

The outputs from G-RDEM are either available in percentage changes over time or in monetary units. For agricultural land demand and agricultural production, which are crucial variables for MagnetGrid downscaling, we make use of the temporal percentage change so that any physical value (i.e. unit of area for land demand, e.g. *hectare* and unit of mass for agricultural production e.g. *tonnes*) added in the baseline year of 2014 is multiplied by the percentage variation of the subsequent years (until 2050). The baseline values are sourced from FAO data [7] on crop specific annual agricultural production and harvested area per crop type at national level.

3.2. Downscaling G-RDEM agricultural land results

The G-RDEM downscaling based on MagnetGrid land-use model involves several steps to coherently integrate data inputs from diverse sources and with different formats, and generate simulation results. These steps are summarized (already considering the current integration with the G-RDEM) in sections 3.3 (model set up), 3.4 (spatial data module), 3.5 (spatial cost-benefit module) and 3.6 (land allocation module). Figure 1 shows a generic overview

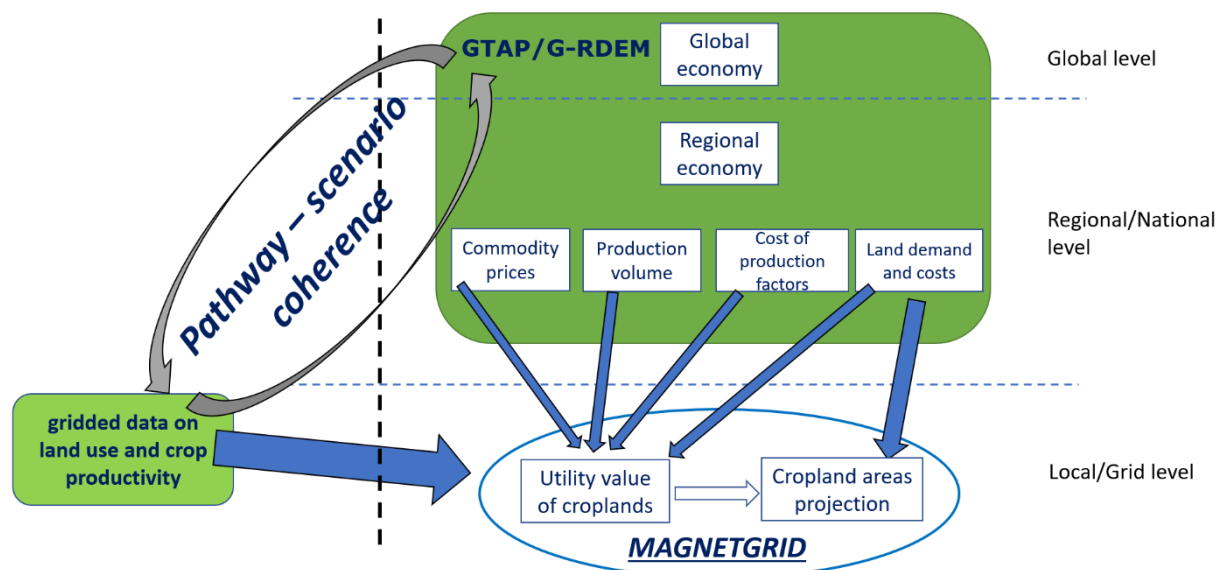


Figure 1. MagnetGrid's multimodel framework for the simulation of agricultural land-use patterns for regional models, derived from Diogo et al. [2].

3.3. MagnetGrid model set up

The MagnetGrid model set up (or scenario configuration) consists of organizing the initial state of the downscaling. In this stage, we build an interface with G-RDEM results and add basic input files.

- The **interface with G-RDEM** (or scenario building) is the preparation of table format archives (*INI_files*) containing key information on the scenario that is modelled. In this step, we inform MagnetGrid on the level of aggregation (regions and sectors) that G-RDEM results are produced to avoid an inconsistent representation of the macro-economic simulation. Moreover, *INI_files* also contain basic exogenous economic assumptions (e.g. discount rates) and scenario information (e.g. time steps).

- In the current model framework, the **basic input files** for setting up a scenario consist of global spatial datasets that allows for downscaling G-RDEM results to the grid level. These minimum required datasets are:

a) Initial land use maps containing both non-agricultural (NAg) land uses (exogenous in MagnetGrid, i.e. not dynamically modelled) and crop-specific (including pasturelands) distribution maps (endogenous in MagnetGrid, i.e. dynamically modelled), see Table 3. The latter represents the land baseline year and should be as close as possible to the G-RDEM baseline. Both maps are available in percentage of land area per grid cell, which allows for the quantification of the dynamic of land change.

b) Agro-ecological suitability maps are crop-specific biophysical information on maximum attainable yield considering both rainfed and irrigated conditions (i.e. with and without water restrictions). The productivity levels of crops have major influence on the economic returns of certain agricultural land due to economies of scale [8].

c) Irrigation maps are key in MagnetGrid as it informs the location of areas containing irrigated areas. Therefore, it will guide the use of either rainfed or irrigated agro-ecological suitability map for the given region.

d) World administrative division map contains the borders of all countries and territories in the world following the ISO 3166 code. This shapefile map will be used to extract the scenario geographical extent of the scenarios and also to align with the GTAP aggregated regions.

Table 3. Description of the basic input files

Spatial data input		Original data format	Original grid size	Reference year	Source
Land use maps	Crop-specific distribution maps	Raster	5 arcmin	2010	[9–12]
	NAg land uses	Raster	5 arcmin	2005 - 2014	[9–13]
Agro-ecological suitability maps		Raster	5 arcmin	Baseline condition = 2010	[14]
Irrigation maps		Raster	5 arcmin	2005	[15]
World administrative division map		Shapefile	-	-	

3.4. Spatial data module

The spatial data module consists of GIS operations to combine crop-specific distribution maps and agro-ecological suitability maps, so that they are representative to the agricultural sectors of G-RDEM (see Table 2). For crop-specific distribution maps, the combination is done by summing up the land area in each grid cell that is used by the different crops that are part of the same agricultural sector. Depending on the case study, the input crop-specific distribution maps also captures double cropping in their statical downscaling. This can create agricultural sectors map with areas larger than the actual grid cell area. In the current model set-up, we truncate all the sectors by the actual grid cell area. For the agro-ecological suitability maps, the maps of the crops belonging to the same sector are combined by taking the suitability index value from the crop with highest value in each grid cell. As a result the spatial data module create sectoral agro-ecological suitability maps and sectoral (crop-specific) distribution maps, which are the main inputs for the spatial cost-benefit analysis module.

3.5. Spatial cost-benefit module

This module creates the two main spatial data required to carry out the land use change simulation. At first, the spatial cost-benefit analysis module carries out the valuation of the local economic returns, i.e. Net Present Value (NPV) of each agricultural sector in every grid cell. This is done by combining sectoral agro-ecological suitability maps and sectoral crop-specific distribution maps with G-RDEM national agro-economic projections. Hence, in this module the main agro-economic indicators from G-RDEM (e.g. land prices, agricultural production costs, revenues) are downscaled from the national to grid level. As a result, NPV sectoral maps are produced. The main downscaling operations are described by the following equation 1 [2]:

$$NPV_{c,j,t} = -Inv_{c,j,t} + \sum_{y=1}^n \frac{R_{c,j,y} - C_{j,y} + NST_{j,t}}{(1+r)^y} = -Inv_{c,j,t} + \frac{R_{c,j,t} - C_{j,t} + NST_{c,j,t}}{\frac{r \cdot (1+r)^n}{(1+r)^{n-1}}} \quad (9)$$

where:

$Inv_{c,j,t}$ is the average investment costs per unit of land area (in USD\$/ha) to convert land in gridcell c into agricultural land use j , in time-step t ;

$R_{c,j,t}$ is the expected annual gross revenues per unit of land area (in USD\$/ha) of agricultural land use j in gridcell c , in time-step t ;

$C_{j,t}$ is the expected annual production costs (in USD\$/ha) per unit of land area of agricultural land use j , in time-step t ;

$NST_{j,t}$ is the average net subsidies and taxes per unit of land area (in USD\$/ha) related to agricultural land use j in gridcell c , in time step t ;

$\frac{r \cdot (1+r)^n}{(1+r)^{n-1}}$ is the capital recovery factor, i.e. the ratio of a constant annuity to the present value of receiving that annuity for a given length of time;

r is the discount rate;

n is the lifetime of the project (in years).

In a second moment, the exogenous NAg land uses and the sectoral (crop-specific) distribution maps are updated to avoid grid cells summed up to larger extent than the grid cell area. Hence, if the summation of both maps exceeds the grid size, we assume that the sectoral (crop-specific) distribution maps remain stable, whereas the amount of exogenous NAg land uses is reduced/adjusted. Figure 2 shows the resulting process of adjusting the grid cells for the input land use maps in MagnetGrid.

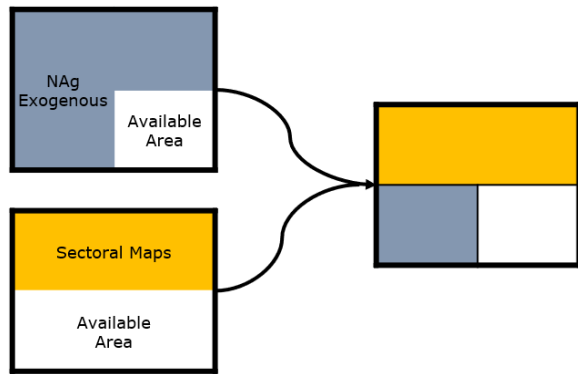


Figure 2. Schematic representation of pre-processing and adjusting land use grid cell for the land change simulation in MagnetGrid.

3.6. Land Allocation

The land allocation method, is a discrete choice model that explains the probability for a certain agricultural land use being chosen in a particular location, according to the utility of that specific agricultural land use in relation to the total utility of all possible alternative agricultural land-use types in that location [2]. Therefore, the discrete choice model can thus be formulated in a spatially-explicit way.

In summary, the land allocation module is mainly divided in three steps:

- 1) valuation of the local utility (U) by summing up sectoral NPV maps derived from the spatial cost-benefit module with land opportunity costs (i.e. net economic benefits from the previous land use that would be foregone) and sunk costs (i.e. investments for land and capital assets that have already been made in the previous land use). These costs can only be calculated in the temporal dynamic land allocation module because they are related to the previous ($t - 1$) land use of the simulated year (t).
- 2) assessment of land that is still available for future land expansion (i.e. land that is currently not occupied by the endogenous land classes nor by the exogenous land classes).
- 3) allocation algorithm that carries out the spatio-temporal land dynamic simulation

In the land allocation module, balancing demand and supply is the core. In this module, Parameter a_j can be interpreted as the demand balancing factor that ensures that the total

amount of allocated land for the endogenous land-use type (or sector) j with $j \in J$ equals the sector-specific land demand. Parameter b_c is the supply balancing factor that ensures that the total amount of allocated land in cell c does not exceed the amount of land that is available in that particular cell. The goal is to allocate the demands of all sectors without exceeding the totals of the supply and exogenous functions (i.e. occurrence of NAg land uses) for each grid cell. The size of a grid cell is defined as A_c and the exogenous land use types e in a grid cell for $e \in E$ is $M_{c,e}$. The available land for allocating endogenous land-use types is $L_c = A_c - \sum_e M_{c,e}$. Appropriate values a_j and b_c are found through an iterative approach simulating a bidding process between competing land uses of the sectors accounting. Exogenous NAg land uses are not competing, these are used to limit the total available land use by the sectors.

In each iteration, b_c is changed in such a way that deviations between total demand and available land are gradually reduced, comparable to slow cooling as implemented in the simulated annealing algorithm.

In the first iteration starting values are set: $a_j = 1$ for all land use sectors j and $b_c = (1/\exp(200))$ for all grid cells c .

$M_{c,j}$ is the area used by endogenous land-use sectors j in grid cell c , which is calculated each iteration step after updating a_j and b_c . According to the formula in Eq. (10):

$$M_{c,j} = a_j * b_c * \exp(\beta * U_{c,j}) \text{ for all combinations of } j \text{ and } c \quad (10)$$

The β factor is used to spread the claims of land use types. It can be set by the user and a value below 1 requires more iterations. However, when the procedure of downscaling is executed with smaller adjustments per iteration, there is a better chance that the simulation will lead to a successful allocation. For this version of MagnetGrid, $\beta = 0.25$, although other values can be considered depending on the quality of the allocation results. When $M_{c,j}(s)$ exceeds the area of c , the values is set to that area.

$$D_j = \sum_c M_{c,j} \quad (11)$$

In each iteration, a_j is updated:

$$a_j(s) = \frac{D_j}{\sum_c b_c(s-1) \exp(\beta * U_{c,j}(s))} \quad (12)$$

For updating b_c the variable M_{all} is calculated and applied:

$$M_{all,c}(s) = \sum_j a_j(s) \exp(\beta * U_{c,j}(s)) \text{ for } s \quad (13)$$

Now b_c can be updated:

$$b_c(s) = (1 - \text{SlowFactor}(s)) * b_c(s-1) + \frac{\text{SlowFactor}(s) * (1 - \text{Exo}_c)}{M_{all,c}(s)} \quad (14)$$

The *Slowfactor* is used to come more gradually to a solution, which is a pragmatic approach to avoid a situation that a solution cannot be reached. For each iteration step s (total N_s steps):

$$SlowFactor(s + 1) = 1 - \min(0.99999, \frac{0.99999 * s}{round(N_s/2)}) \quad (15)$$

The iteration procedure stops for three reasons:

- The maximum number of iterations N_s are achieved, $s = N_s$;
- A solution is found, where all land demands are allocated, $Demand_j - \sum_c M_{c,j} = 0$;
- There is no solution. This means that the absolute value of the discrepancy between demand and supply is positive, and the value of the discrepancy did not change for 3 subsequent iterations: $Demand_j - \sum_c M_{c,j} > 0$.

S	pdr	gro	osd	c_b	v_f	ocr	pbf	ctl	AbsoluteSumDeviation
48	-2	-15	-8	0	-29	-18	0	-80	152
49	-2	-10	-5	0	-19	-12	0	-53	101
50	-1	-5	-3	0	-10	-6	0	-27	52
51	0	0	0	0	0	0	0	0	0 [solution!]

Figure 3. An example of a successful stop of the iteration procedure:

At the beginning of the iteration procedure, the absolute sum of the deviations drops very rapidly from step to step due to the graduality mechanisms of the algorithm. As more iterations are carried out, the improvements at each iteration are becoming minor when the results are reaching the solution. In the example in Figure 3, all land demand is allocated in 51 steps without violating land supply constraints.

4. Results of downscaled socioeconomic indicators

4.1 Information on the results

Meta-information on the downscaling of socio-economic indicators with MagnetGrid

With the downscaling of socioeconomic indicators, 8 files with indicators were produced, see Table 4. The data files provide results for different combination of SSP-scenarios, RCP-scenarios and indicators. For the SSP scenarios, two scenarios [16] were considered:

- SSP2 Middle of the road, The world follows a path in which social, economic, and technological trends do not shift markedly from historical patterns
- SSP4: Inequality—A road divided. Highly unequal investments in human capital, combined with increasing disparities in economic opportunity and political power, lead to increasing inequalities and stratification both across and within countries.

The results of the socioeconomic indicators from the SSP/RCP scenarios were derived from the results of the G-RDEM models provided at the end of September 2023. New and updated SSP scenarios from the G-RDEM models might have slightly deviating results.

Table 4. File names for the different set of results per SSP scenario, RCP scenario and per indicator

File name	SSP scenario	RCP scenario	Indicator
SSP2RCP26_AREA.xlsx	2	2.6	Land use
SSP2RCP26_prod.xlsx	2	2.6	Production
SSP2RCP85_AREA.xlsx	2	8.5	Land use
SSP2RCP85_prod.xlsx	2	8.5	Production
SSP4RCP26_AREA.xlsx	4	2.6	Land use
SSP4RCP26_prod.xlsx	4	2.6	Production
SSP4RCP85_AREA.xlsx	4	8.5	Land use
SSP4RCP85_prod.xlsx	4	8.5	Production

In this section, we present the results of the downscaling for the five case studies. For all case studies, a similar set of results are produced and stored in the data repository.

The indicator variable names are listed in Table 4. Projections for 2020, 2030, 2040 and 2050 were prepared and the results between these points in time were interpolated. The base year is 2015.

Table 5. The list of variables included in the database for each case study

Variables	Unit	Description
YEAR		year of analysis
AGRI_SECTOR		crop type
REGION		country
BASIN - LAND AREA	ha	crop area (see AGRI_SECTOR) within the river basin of a given country (see tab name)
BASIN - CHANGING RATE		rate of change (compared to 2015) within the river basin
BASIN - PERCENTAGE	%	percentage of change (compared to 2015) within the river basin
COUNTRY - LAND AREA	ha	crop area (see AGRI_SECTOR) within the country (see tab name)
COUNTRY - CHANGING RATE		rate of change (compared to 2015) in the country
COUNTRY - PERCENTAGE	%	percentage of change (compared to 2015) in the country

There are three single country river basins and two transboundary river basin case studies. As the MagnetGrid data is organised by country, Table 5 indicates the relevant countries per river basin. The downscaled socioeconomic results in the data repository are organised by COUNTRY.

The list of agricultural sectors included in the downscaling are:

- sugar crops (c_b)
- paddy rice (pdr)
- vegetables and fruits (v_f)
- cereals (gro)
- other crops (ocr)
- fiber crops (pbf)
- oil seed crops (osd)
- wheat (wht)
- cattle (ctl)

Table A.1 in the Annex provides more detailed information on the crop- and animal-based products that are included in the agricultural sectors.



Table 6: Relevant countries in the downscaled socioeconomic data bases

#	Case study	COUNTRY
1	Nestos river basin	Bulgaria
		Greece
2	Lielupe river basin	Latvia
		Lithuania
3	Jiu river basin	Romania
4	Aldige river basin	Italy
5	Incomati-Usuthu river basin	South Africa

Land area trends

Figure 4 displays four panels of land demand trends over time for the different river basins in different countries and climate/socio-economic pathways. These graphs aggregated all the land demand (i.e. they are not crop-specific) within the river basin boundaries in order to identify major differences across the scenario. As a common result across all scenarios, the Inkomati-Usuthu river basin presents the largest room for land expansion driven by a large demand for the cattle sector, whereas the European river basins presented a much more constrained area for expansion. Moreover, there are major differences between SSP 2 and SSP 4, whereas the differences driven by climate scenarios (RCPs) are not largely perceived in the land area within the river basin. even though it will have a major impact on the production quantities. A more in-depth analysis will be carried out in a more scientific publication highlighting the added value of downscaling socio-economic indicators in Nexogenesis.

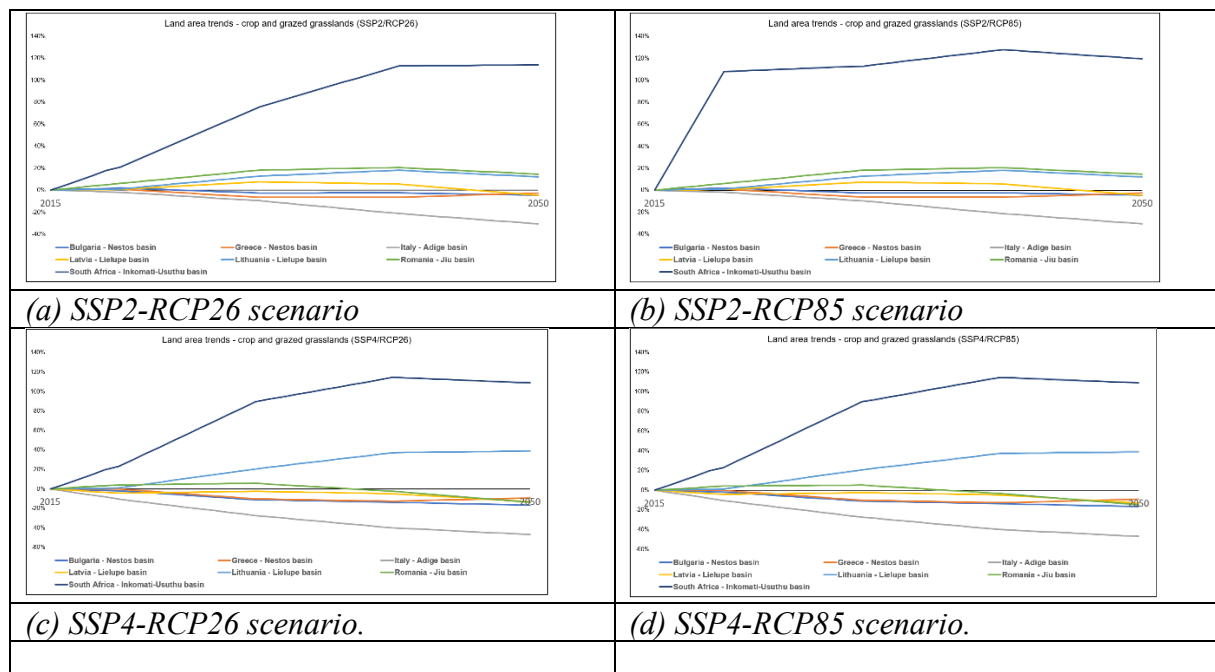


Figure 4. land area trends over time per country/river basin for four combinations of SSP and - RCP scenario: (a) SSP2-RCP2.6, (b) SSP2-RCP8.5, (c) SSP4-RCP2.6 and (d) SSP4-RCP8.5.

Example of land use map

In principle, the results from the downscaling of economic indicators can be plotted on maps, both at the national and the river basin level. Figure 4 shows an example of the spatio-temporal variation of cropland and grazed grasslands for the Adige river basin. The downscaling of socioeconomic indicators for Italy was derived from G-RDEM results for Italy under the SSP4 scenario. Figure 4 shows the spatial distribution of the demand driven land use types (cropland and grazed grasslands, which are endogenously modelled). According to the G-RDEM projection, the combination of arable agriculture (excluding rice) and the grazing livestock sector will slightly increase the land demand in Italy by approx. 10%, especially triggered by the oil seed (*osd*) and vegetables and fruits (*v_f*) sectors. Moreover, Figure 4 shows the land use types that are exogenous in MagnetGrid (i.e. water bodies, urban areas and existing forests), and assumed to be fixed over time. The Adige river basin is marked with a black border in the northern part of Italy. Information on river basin level and national level is derived from the aggregation of the relevant grid cells.

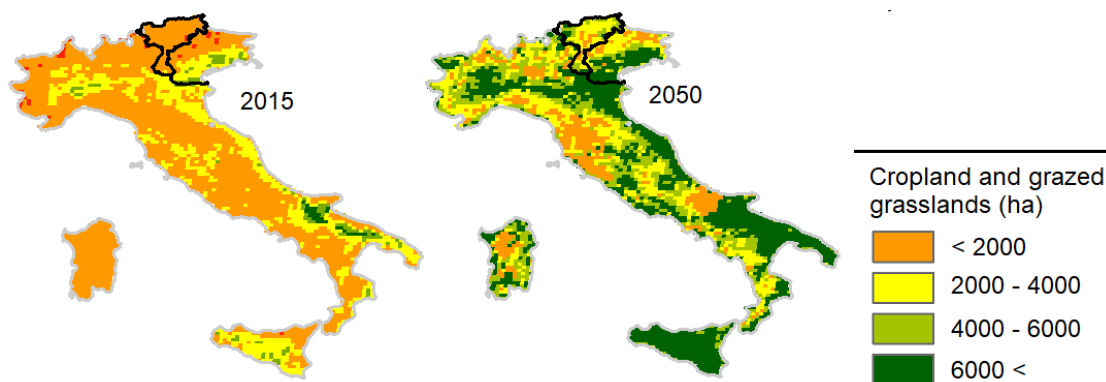


Figure 5. Spatio-temporal distribution of cropland and grazed grasslands in 2015 (baseyear) and in 2050 (SSP4). Grid cell size = 10,000 ha.

Maps as displayed in Figure 4 are not directly part of the results, as the SDM models do not use spatial (or grid cell) level data.

4.2 Examples of results for case studies

This section describes a random number of examples for the case studies derived from the downscaled socioeconomic indicators provided in the data repository.

- Development of land use for wheat (*wht*) under the SSP2 scenario and the RCP2.6 scenario for the Nestos case study
- Development of production of vegetables and fruit (*v_f*) under the SSP2-scenario and the RCP8.5-scenario for the Lielupe river basin
- Development of land area for grazing land for cattle (*clt*) over time under the SSP4-scenario and the RCP8.5-scenario for the Jiu river basin
- Development of cereals production (*gro*) over time under the SSP4-scenario and the RCP8.5-scenario for the Adige river basin
- Development of land area for cereals (*gro*) over time under the SSP2-scenario and the RCP8.5-scenario for the Inkomati-Isuthu river basin

Nestos river basin

Table 7 shows a sub-sample (i.e. baseline and simulated years) of the results for the wheat sector in the SSP2/RCP2.6 scenario for the Nestos river basin. Downscaling national results on supply and demand in transboundary river basins provides more detailed information on upstream dynamics affecting downstream environmental flows, especially when several activities are competing for water resources.

Table 7. Development of land area for wheat (wht) over time under the SSP2-scenario and the RCP2.6-scenario for the Nestos river basin

			Basin			Country		
			Land area (ha)	Changing rate (base 2015)	Share of change (base 2015)	Land area (ha)	Changing rate (base 2015)	Share of change (base 2015)
YEAR	AGRI SECTOR	REGION	BASIN - LAND AREA (ha)	BASIN - CHANGING RATE	BASIN - PERCENTAGE	COUNTRY - LAND AREA (ha)	COUNTRY - CHANGING RATE	COUNTRY - PERCENTAGE
Bulgaria								
2015	wht	BGR	4,613.4	1	0%	878,089	1	0%
2020	wht	BGR	23,186.8	5.03	403%	1,163,272	1.32	32%
2030	wht	BGR	16,780.7	3.64	264%	1,112,618	1.27	27%
2040	wht	BGR	19,022.7	4.12	312%	1,099,886	1.25	25%
2050	wht	BGR	17,690.0	3.83	283%	1,075,188	1.22	22%
Greece								
2015	wht	GRC	7,945.2	1	0	489,068	1	0%
2020	wht	GRC	10,476.9	1.32	32%	562,187	1.15	15%
2030	wht	GRC	10,662.6	1.34	34%	528,395	1.08	8%
2040	wht	GRC	9,836.4	1.24	24%	523,391	1.07	7%
2050	wht	GRC	9,410.4	1.18	18%	502,166	1.03	3%
Nestos river basin calculated								
2015	wht		12,558.6	1	0%			
2020	wht		33,663.6	2.68	168%			
2030	wht		27,443.3	2.19	119%			
2040	wht		28,859.1	2.30	130%			
2050	wht		27,100.4	2.16	116%			

Remark: Selection of data from the worksheets "Bulgaria" and "Greece" in the file "SSP2RCP26_AREA.xlsx".

Lielupe river basin

Table 8 shows a sub-sample (i.e. baseline and simulated years) of the results for vegetables and fruits in the SSP2/RCP8.5 scenario for the Lielupe river basin. This example shows that for Lielupe river there is more relevance of food crop production and associated impact in Latvia than Lithuania. Understanding these regional differences are key for promoting adequate water governance within the river basin area.

Table 8. Development of production of vegetables and fruit (v_f) over time under the SSP2-scenario and the RCP8.5-scenario for the Lielupe river basin

			Basin			Country		
			Production (tonnes)	Changing rate (base 2015)	Share of change (base 2015)	Production (tonnes)	Changing rate (base 2015)	Share of change (base 2015)
YEAR	AGRI_SEC-TOR	REGION	BASIN - PROD (t)	BASIN - CHAN- GING RATE	BASIN - PER- CENTAGE	COUNTRY – PROD (t)	COUNTRY - CHAN- GING RATE	COUNTRY - PERCEN- TAGE
Latvia								
2015	v_f	LVA	187,829	1	0%	690,936	1	0%
2020	v_f	LVA	477,249	2.54	154%	681,635	0.99	-1%
2030	v_f	LVA	663,112	3.53	253%	663,114	0.96	-4%
2040	v_f	LVA	647,633	3.45	245%	647,633	0.94	-6%
2050	v_f	LVA	664,081	3.54	254%	664,081	0.96	-4%
Lithuania								
2015	v_f	LTU	133,628	1	0%	869,359	1	0%
2020	v_f	LTU	162,007	1.21	21%	887,311	1.02	2%
2030	v_f	LTU	101,278	0.76	-24%	934,770	1.08	8%
2040	v_f	LTU	102,813	0.77	-23%	978,788	1.13	13%
2050	v_f	LTU	254,957	1.91	91%	1,066,187	1.23	23%
Lielupe river basin calculated								
2015	v_f		321,456	1.00	217%			
2020	v_f		639,257	1.99	531%			
2030	v_f		764,390	2.38	655%			
2040	v_f		750,446	2.33	641%			
2050	v_f		919,038	2.86	807%			

Remark: Selection of data from the worksheets "Latvia" and "Lithuania" in the file "SSP2RCP85_prod.xlsx".

Jiu river basin

Table 9 shows a sub-sample (i.e. baseline and simulated years) of the results for the area of grazing land (*clt*) under the SSP2/RCP8.6 scenario for the Jiu river basin. This example shows that the grazing area in the Jiu river basin is growing less in the first decades compared to the average in Romania, and decline more strongly than the national average in the last three decades (2030-2050) in this scenario. In 2050, the grazing areas decline to 35% of the surface for the base year 2015 in the river basin. At the national level, the total grazing area in 2050 under this scenario is 57% of the total surface of grazing land in the base year 2015.

Table 9. Development of land area for grazing land for cattle (clt) over time under the SSP4-scenario and the RCP8.5-scenario for the Jui river basin

			Basin			Country		
			Land area (ha)	Changing rate (base 2015)	Share of change (base 2015)	Land area (ha)	Changing rate (base 2015)	Share of change (base 2015)
YEAR	AGRI_SEC-TOR	REGION	BASIN - LAND AREA (ha)	BASIN - CHAN- GING RATE	BASIN - PERCEN- TAGE	COUNTRY - LAND AREA (ha)	COUNTRY - CHAN- GING RATE	COUNTRY - PERCEN- TAGE
2015	ctl	ROU	176,577	1	0%	4,463,471	1	0%
2020	ctl	ROU	202,837	1.15	15%	6,546,467	1.47	47%
2030	ctl	ROU	169,821	0.96	-4%	5,940,207	1.33	33%
2040	ctl	ROU	112,358	0.64	-36%	4,267,581	0.96	-4%
2050	ctl	ROU	62,110	0.35	-65%	2,566,333	0.57	-43%

Remark: Selection of data from the worksheet "Rumania" in the file "SSP4RCP85_AREA.xlsx".

Adige river basin

Table 10 shows a sub-sample (i.e. baseline and simulated years) of the results for the production of cereals (*gro*), including maize, pearl millet, small millet, sorghum and other cereals, under the SSP4/RCP8.5 scenario for the Adige river basin. Although the share for the Adige river basin is slightly more than 1% of the total Italian cereals production, this example shows that the production of cereals in the Adige river basin grows gradually in the period 2020-2050 with a 42% increase compared to 2015. However, cereals production in Italy in the same period will grow faster under this scenario. For 2050, for instance the cereals production in Italy is expected to have grown by 63% compared to the base year.

Table 10. Development of cereals production (gro) over time under the SSP4-scenario and the RCP8.5-scenario for the Adige river basin

			Basin			Country		
			Production (tonnes)	Changing rate (base 2015)	Share of change (base 2015)	Production (tonnes)	Changing rate (base 2015)	Share of change (base 2015)
YEAR	CROP	REG	BASIN - PROD (t)	BASIN - CHANGING RATE	BASIN - PER- CENTAGE	COUNTRY - PROD (t)	COUNTRY - CHANGING RATE	COUNTRY - PERCENTAGE
2015	gro	ITA	128,582	1	0%	10,005,466	1	0%
2020	gro	ITA	117,759	0.92	-8%	10,821,089	1.08	8%
2030	gro	ITA	144,577	1.12	12%	12,773,039	1.28	28%
2040	gro	ITA	168,872	1.31	31%	14,797,403	1.48	48%
2050	gro	ITA	182,420	1.42	42%	16,270,057	1.63	63%

Remark: Selection of data from the worksheet "Italy" in the file "SSP4RCP26_prod.xlsx".



Inkomati-Usuthu river basin

Table 11 shows a sub-sample (i.e. baseline and simulated years) of the results for land used for cereals production (*gro*), including maize, pearl millet, small millet, sorghum and other cereals, under the SSP2/RCP8.5 scenario for the Inkomati-Usuthu river basin. Although the share for this river basin is slightly more than 1% of the total South Africa area for cereals production. This example shows that the area for cereals production in the Inkomati-Usuthu river basin is expected to fluctuate heavily during 2020-2050. However, the area of cereals production in South Africa is expected to grow gradually to twice the surface observed in the base year 2015.

Table 11. Development of land area for cereals (gro) over time under the SSP2-scenario and the RCP8.5-scenario for the Incomati-Isuthu river basin

			Basin			Country		
			Land area (ha)	Changing rate (base 2015)	Share of change (base 2015)	Land area (ha)	Changing rate (base 2015)	Share of change (base 2015)
YEAR	AGRI SECTOR	REGION	BASIN - LAND AREA (ha)	BASIN - CHANGING RATE	BASIN - PERCENTAGE	COUNTRY - LAND AREA (ha)	COUNTRY - CHANGING RATE	COUNTRY - PERCENTAGE
2015	gro	ZAF	98,419	1	0%	2,037,761.6	1	0%
2020	gro	ZAF	71,194	0.72	-28%	3,374,830.8	1.66	66%
2030	gro	ZAF	151,640	1.54	54%	3,701,926.8	1.82	82%
2040	gro	ZAF	90,178	0.92	-8%	3,929,306.1	1.93	93%
2050	gro	ZAF	135,721	1.38	38%	4,108,443.7	2.02	102%

Remark: Selection of data from the worksheet "South Africa" in the file "SSP2RCP85_AREA.xlsx".

5. Application with case studies SDMs

The utilization of the results from MagnetGrid in the SDMs was developed in cooperation with the case studies. The downscaling framework (see green blocks in Figure 6) does not include local data on land use, which is important to give a more realistic assessment of the case study conditions. This would also require a data harmonization (see grey blocks in Figure 6) process in order to reduce mismatches across the different spatial datasets, which can be time-intensive rather data-intensive and time consuming. Hence, each case study will be analyzed individually to make sure we accommodate as much local stakeholder input as possible without compromising the project timeline, such as land uses of locally grown crops, for instance. In

addition, G-RDEM has the potential to provide sub-national macro-economic projections at NUTS2 level, which would not change the current downscaling framework, but could demand more processing time.

Recurrent meetings with case studies have been ongoing twice a month from the beginning of the project through WP3 activities to define case study and SDM development needs based on stakeholder elicitation. From the *Nexogenesis* project meeting in September 2023 in Tours/France, we have presented preliminary results and verified final feedback from the case studies to consolidate the specific indicators, and associated procedure. The feedback from the case studies and common agreements were incorporated in the modelling framework.

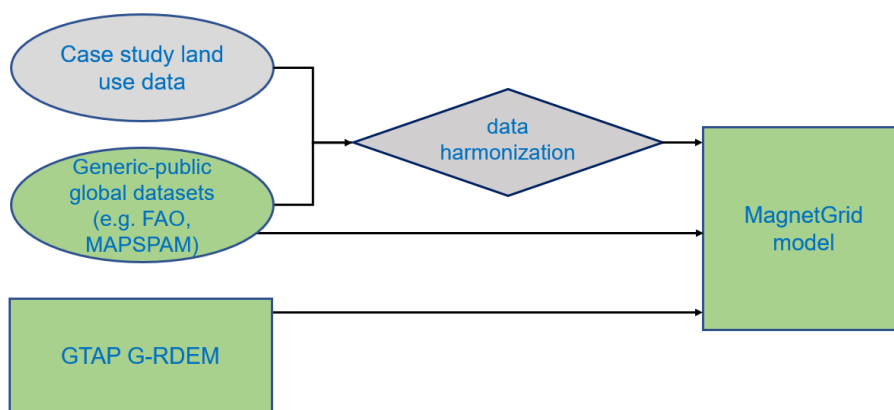


Figure 6. Workflow of model integrations for generating land use simulations.

The green blocks in Figure 6 show the established process of integration of MagnetGrid and G-RDEM. The gray blocks in Figure 6 show the processes that are yet to be enabled jointly with the case studies.

6. References

- [1] Britz W, Roson R. G-RDEM: A GTAP-Based Recursive Dynamic CGE Model for Long-Term Baseline Generation and Analysis. SSRN Electron J 2018. <https://doi.org/10.2139/ssrn.3167781>.
- [2] Diogo V, Hennen W, Verma M, Oudendag D, Kuiper M. MagnetGrid: Model description and user guide. The Hague: 2019.
- [3] Woltjer GB, Kuiper M. The MAGNET model: Module description. Wageningen LEI 877 Wageningen UR 2014:144.
- [4] Schaphoff S, Forkel M, Müller C, Knauer J, Von Bloh W, Gerten D, et al. LPJmL4 - A dynamic global vegetation model with managed land - Part 2: Model evaluation. Geosci Model Dev 2018;11:1377–403. <https://doi.org/10.5194/gmd-11-1377-2018>.
- [5] Walmsley TL, Aguiar AH. Introduction to the Global Trade Analysis Project and the GTAP Data Base. 2012.
- [6] Diogo V, Koomen E, Kuhlman T. An economic theory-based explanatory model of agricultural land-use patterns: The Netherlands as a case study. Agric Syst 2015;139:1–16. <https://doi.org/10.1016/j.agsy.2015.06.002>.
- [7] FAO. FAOSTat 2023. <https://www.fao.org/faostat/en/#data> (accessed March 1, 2023).
- [8] van der Hilst F, Faaij APC. Spatiotemporal cost-supply curves for bioenergy production in

- Mozambique. Biofuels, Bioprod Biorefining 2012;6:405–30. <https://doi.org/10.1002/BBB>.
- [9] Pesaresi M, Syrris V, Julea A. A New Method for Earth Observation Data Analytics Based on Symbolic Machine Learning. Remote Sens 2016;8:399. <https://doi.org/10.3390/rs8050399>.
- [10] Lamarche C, Santoro M, Bontemps S, d’Andrimont R, Radoux J, Giustarini L, et al. Compilation and Validation of SAR and Optical Data Products for a Complete and Global Map of Inland/Ocean Water Tailored to the Climate Modeling Community. Remote Sens 2017;9:36. <https://doi.org/10.3390/rs9010036>.
- [11] You L, Wood S, Wood-Sichra U, Wu W. Generating global crop distribution maps: From census to grid. Agric Syst 2014;127:53–60. <https://doi.org/10.1016/j.agry.2014.01.002>.
- [12] Ramankutty N, Evan AT, Monfreda C, Foley JA. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochem Cycles 2008;22:1–19. <https://doi.org/10.1029/2007GB002952>.
- [13] United Nations Environment Programme. World Database on Protected Areas 2022.
- [14] Fischer G, Nachtergaele; F, Velthuisen; H van, Chiozza; F, Franceschini; G, Henry; M, et al. Global Agro Ecological Zones (GAEZ v4): Model documentation. 2021.
- [15] Siebert S, Henrich V, Frenken K, Burke J. Update of the digital global map of irrigation areas to version 5. Rheinische Friedrich-Wilhelms-Universität, Bonn, Ger Food Agric Organ United Nations, Rome, Italy 2013:171.
- [16] O’Neill BC, Kriegler E, Ebi KL, Kemp-Benedict E, Riahi K, Rothman DS, et al. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. Glob Environ Chang 2017;42:169–80. <https://doi.org/10.1016/j.gloenvcha.2015.01.004>.

Annex : Meta-information of the socioeconomic data

Table A.1. Agricultural food groups in the economic models and agricultural products per group

Agricultural sectors	Variable name	Number of products	List of agricultural products included in the agricultural sector			
sugar crops	c_b	2	1	sugarbeet	2	sugarcane
paddy rice	pdr	1	1	rice		
vegetables and fruits	v_f	16	1	banana	9	pigeon pea
			2	bean	10	plantain
			3	cassava	11	potato
			4	chickpea	12	sweet potato
			5	cowpea	13	temperate fruits
			6	lentils	14	tropical fruits
			7	other pulses	15	vegetables
			8	other roots	16	yams
cereals	gro	5	1	maize	4	small millet
			2	other cereals	5	sorghum
			3	pearl millet		
other crops	ocr	5	1	arabica coffee	4	tea
			2	cocoa	5	tobacco
			3	robusta coffee		
fiber crops	pbf	3	1	cotton	3	other fibers
			2	flax		
oil seed crops	osd	8	1	coconut	5	rapeseed
			2	groundnut	6	sesame seed
			3	oil palm	7	soybean
			4	olive oil	8	sunflower
wheat	wht	2	1	barley	2	wheat
cattle	ctl	1	1	pastureland		

