Ref. Ares(2024)8219449 - 19/11/2024

Nexus data vector of socio-economic data for each case study

Lead: Fondazione Università Ca Foscari (CAF)

Date: 8/10/2023

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

Project

Project Number	Project Acronym	Project Title
101003881	NEXOGENESIS	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:	Thematic Priority
H2020RIA	LC-CLA-14-2020

Title
Nexus data vector of Socioeconomic data for each case study

Contractual Delivery Date	Actual Delivery Date
31.10.2023	27.10.2023 8.10.2024 19.11.2024

Start Date of the project	Duration
01 September 2021	48 months

Organisation name of lead contractor for this deliverable	Document version
CAF	1

Dissemination level		Deliverable Type	
PUBLIC		OTHER	

Authors (organisations)

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

Roberto Roson (CAF), Fasika Mulu Nega (CAF), Walter Rossi Cervi (WR), Vincent Linderhof (WR),

Diti Oudendag (WR)

Reviewers (organisations)

Abstract

Deliverable 2.3 (D2.3) documents the release of a set of socioeconomic data vectors for each case study in NEXOGENESIS, making it accessible to the broader scientific community. This document describes the process of generation of the socio-economic variables for the NEXOGENESIS project. These variables primarily serve as inputs for System Dynamic Models (SDMs) developed for each CS, and subsequently for the machine learning engine. The G-RDEM model, a recursive dynamic computable general equilibrium model, is employed to estimate these variables, incorporating specific assumptions and long-term scenarios based on Shared Socioeconomic Pathways (SSP). Results are shared on the common data sharing drive (SURFdrive). This report also discusses the downscaling procedure for the European CS down to the regional NUTS2 disaggregation, as well as the coupling with the MAGNET model to get land-use results at the grid level. The methodology to assess uncertainty in the scenario data is described in detail, as well as a discussion is provided, about how the socio-economic data can be used to estimate variables in the SDM models of the various case studies.

Keywords

G-RDEM Model, CGE Modelling, Socioeconomic data, SSPs, GTAP database, CGEBox, NUTS

NEXOGENESIS STREAMLINING WATER RELATED POLICIES

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

Abbreviation/Acronyms

CGE	Computable General Equilibrium
CS	Case Study
GAMS	General Algebraic Modelling System
GDP	Gross Domestic Product
GGIG	GAMS Graphical User Interface Generator
GTAP	Global Trade Analysis Project
NXG	NEXOGENESIS
NUTS	Nomenclature of Territorial Units for Statistics
SDM	System Dynamic Model
SSP	Shared Socio-economic Pathway
WEFE	Water-Energy-Food-Ecosystem

NEXOGENESIS STREAMLINING WATER RELATED POLICIES

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

Contents

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

1 Introduction

1.1 Premise

The generation of socio-economic scenario data in the NEXOGENESIS project is based on the IPCC Shared Socioeconomic Pathways (SSP). These include *aggregate* macroeconomic and demographic indicators at the *national* levels, like GDP and population levels.

SSP scenarios *are not forecasts*, but rather reference datasets, that are employed to make different studies, realized by various researchers around the world, comparable and contrastable. As such, SSP scenarios should not be evaluated in terms of realism, or expectations.

A complex numerical macroeconomic model (G-RDEM) is used in NEXOGENESIS to greatly enlarge the spectrum of scenario variables available, beyond the few indicators provided by the SSPs. National GDP and population projections *are taken as given*, and a future, *hypotetical,* global general equilibrium is calculated, with equality of supply and demand in multiple, interdependent markets.

Computable General Equilibrium models, like G-RDEM have their parameters calibrated on the basis of *official economic national accounts*. Therefore, their output typically refers to *national* economic indicators, such as consumption, imports, investments, etc..

In NEXOGENESIS, we were able to get sub-national detail for *some* variables: production side variables at NUTS2 level for *European* case studies, and land-use patterns. As the spatial dimension of economic data does not (and cannot) match the river basin area considered in the various case studies, our strategy has been linking the socio-economic variables in SDM and NEPAT to one or more indicators, referring to one or two broader regions, were the basin falls. Consequently, scenario data have been provided as percentage variations, thereby assuming that these changes could (directly or indirectly) provide reasonable proxies for the variables of interest.

1.2 Description of socio-economic data

WP2 offers a set of socio-economic variables projections, related to the WEFE nexus sectors, for each CS. These variables include socio-economic indicators such as gross production, consumption, energy generation and use, and projected emissions. It's

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

important to note that these socio-economic projections do not function as forecasts but rather as quantifications of possible scenarios derived from qualitative descriptions, specifically SSPs, using a global model. These projections are aimed to serve as inputs for the development of SDMs that are specific to each case study.

The socio-economic variables are estimated using the G-RDEM model, a recursive dynamic CGE model that operates on specific assumptions including non-homothetic demand systems, endogenous saving rates, differentiated industrial productivity growth, interest payments on foreign debt, and variable input-output coefficients. This model is instrumental in creating long-term scenarios for socio-economic variables based on a qualitative scenario sourced from the Shared Socioeconomic Pathways (SSP), which categorize global social and economic scenarios into SSP1 (sustainability-focused growth and equality), SSP2 (middle of the road), SSP3 (resurgent nationalism), SSP4 (ever-increasing inequality), and SSP5 (rapid and unconstrained economic growth and energy usage). The model primarily employs population and GDP projections as its main drivers. In this context, socio-economic projections were formulated using population and GDP projections for two specific scenarios: SSP2 and SSP4. SSP4 is chosen, because it is proving to be the closest one to the recent historical data. SSP2 is chosen, because it is the most employed scenario in other research, thereby easing the comparability with findings in NEXOGENESIS.

Scenario variables, as well as most of the national accounts data used to calibrate parameters in the G-RDEM model, are only available at the national level. The geographical scale of the various case studies in NEXOGENESIS is much finer, though. To get more spatial detail, we undertook two steps. First, for the European CS, we employed a module in G-RDEM which allows getting some macroeconomic variables for the Eurostat NUTS2 regions. Second, we employed the MAGNET-Grid module, fed by variables obtained by G-RDEM, to estimate land-use patterns (at the grid level).

1.3 Modelling and data generation process

1.3.1 Description of the database

The data generation process, based on the G-RDEM model, relies on data sourced from the GTAP 11 database. The GTAP Data Base represents a comprehensive Social Accounting Matrix of the global economy for a specific year, constructed from a multitude of sources, including national input-output tables, trade data, macroeconomics, and energy statistics. The database mainly records market transaction values. It encompasses 64 production sectors, five production sectors, 121 regions, and 20 aggregated regions. However, for

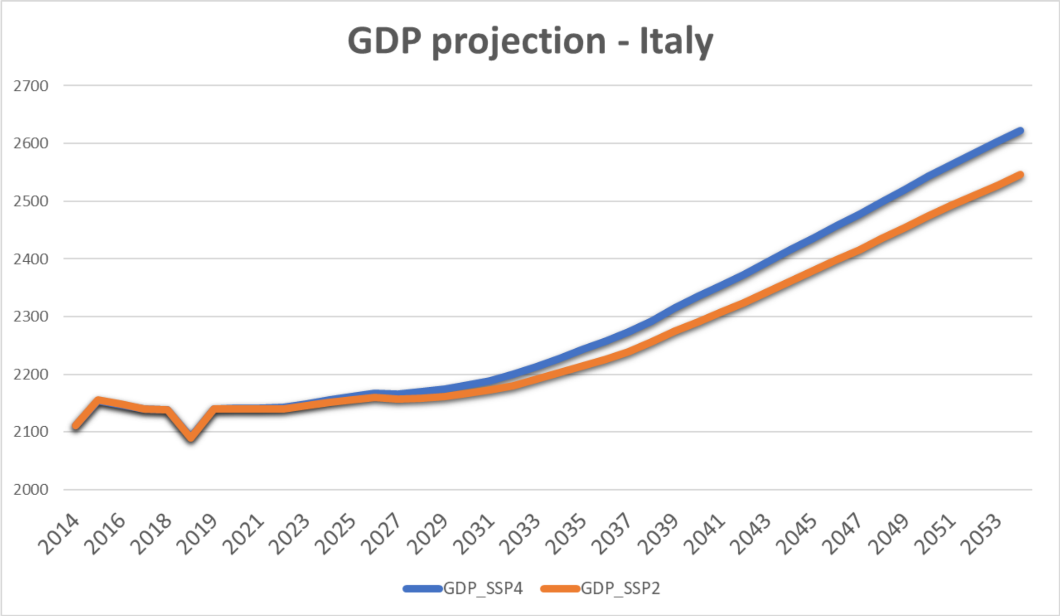
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

NEXOGENESIS, we have tailored an aggregation comprising 10 regions, seven of which are countries associated with case studies:

- Adige– Italy
- Nestos– Bulgaria and Greece [Transboundary]
- Jiu– Romania
- Lielupe– Latvia and Lithuania [Transboundary]
- Inkomati-Usuthu– South Africa

Additional countries, not associated to case studies, have been consolidated in the global G-RDEM model into broader regions, such as North America, the European Union, and the rest of the world. Consequently, our database is structured to encompass 10 regions, 64 production sectors, and five factors of production.

External drivers for projections in the model are the national GDPs and population projections, provided by the Shared Socioeconomic Pathways (SSPs). Specifically, we focus on SSP2 and SSP4 scenarios. The GDP and population projections are shaped by different interpretations of the future, resulting in distinct trends for these two key variables. By way of illustration, the figures below show the evolution of the Italian GDP and population in the two different cases.



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

Figure 1. GDP projection based on two socio-economic pathway scenarios: SSP2 and SSP4

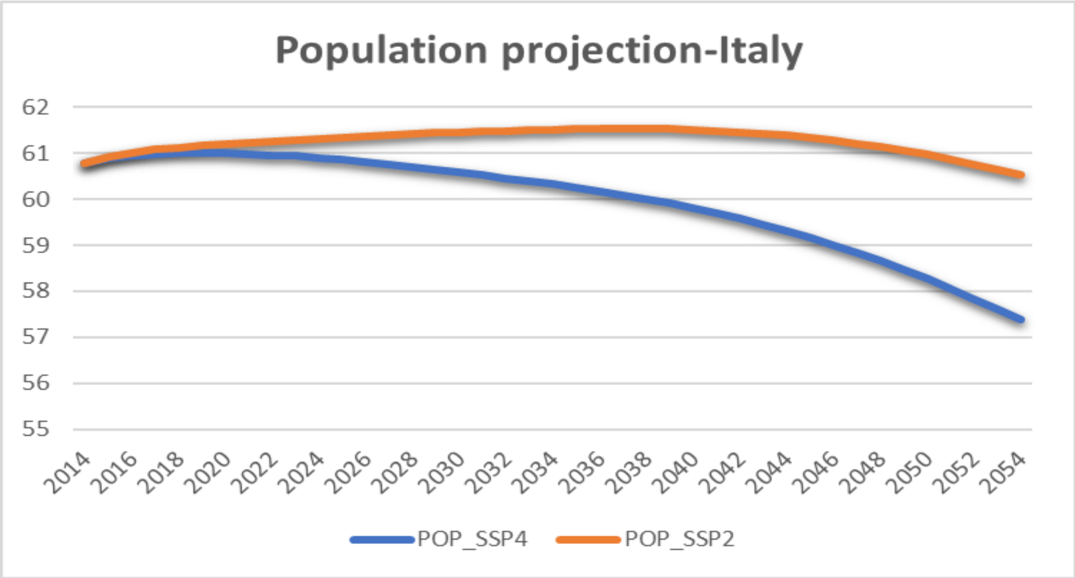


Figure 2. Population projection based on two socio-economic pathway scenarios: SSP2 and SSP4

We execute the G-RDEM model on the CGEBox modeling platform. CGEBox is an open-source code base designed for CGE modeling in GAMS, using GTAP data and offering a user-friendly interface through GGIG. This modelling platform also allows the integration of additional modules like GTAP-AEZ, GTAP-E, GTAP-Melitz, etc., alongside G-RDEM. Thus, we utilized this feature to incorporate additional modules to introduce supplementary assumptions pertaining to emissions, energy, and other relevant factors. For instance, we incorporated CO^2 and Non-CO^2 modules to project total emissions by households and production sectors. We have estimated socioeconomic variables for the ten regions linked to our case studies, spanning from the baseline year of 2014 (2015) up to 2054 (2055).

We summarize in Table 1 the variables, produced by our estimates, and relevant for the NEXOGENESIS case studies. These are the variables that can be found in the common repository, as explained in more detail later in this document. To abstract from differences in units of measures, as well as from heterogeneity in the composition of sectors, our findings are expressed as percentage changes (unit-free) from the variables baseline levels.

Variables	**Dimension**	**Description**
Physical gross output	64 production sectors, sub-regions, quantified as a percentage change relative to the baseline year.	The total quantity of goods and services produced annually by production sectors at regional and subregional level.
Value added	64 production sectors, sub-regions, quantified in value but	The share of production sector to the total GDP of regions, and sub

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

	represented as a percentage change relative to the baseline year.	regions for each case studies.
Total demand	64 production sectors, sub-regions, quantified in value but represented as a percentage change relative to the baseline year.	Total commodity demanded for private, investment, and government consumption.
Energy demand	64 production sectors, sub-regions, quantified in value but represented as a percentage change relative to the baseline year.	The private, investment and government consumption demand of energy goods and services.
Labour income	64 production sectors, sub-regions, quantified in value but represented as a percentage change relative to the baseline year.	Total labour return/salaries based on skill levels: skilled and unskilled labour
Total Co2 and non-Co2 emission	64 production sectors, sub-regions, quantified in value but represented as a percentage change relative to the baseline year.	The total Co2 and non-Co2 emission from household consumption and production activity.

Table 1. Description of socioeconomic variables

1.3.1.2 Illustration of results

The projections are formulated as percentages from baseline values of the baseline year, which is 2014. To illustrate, consider the graph below (Figure 3), showing the trend of total domestic demand in Italy from 2014 to 2054, applicable to the Adige case study. One can see a rising domestic demand for commodities such as petroleum, wheat, cereal grain, and other goods. It also shows a declining trend in domestic demand for products like oil, energy-intensive items, and various manufactured goods.

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

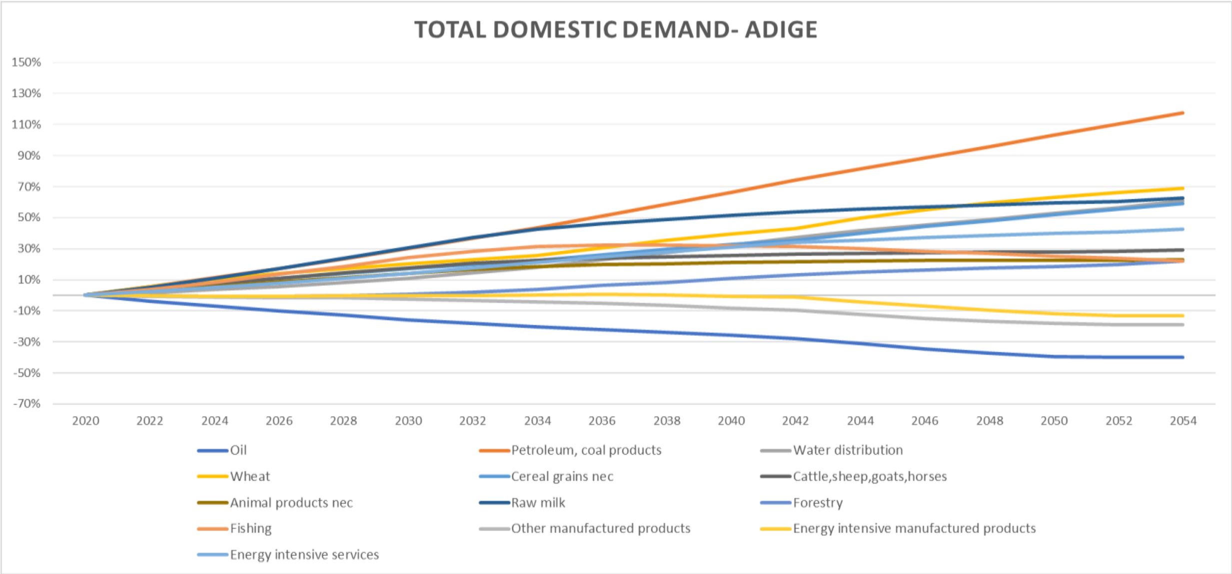
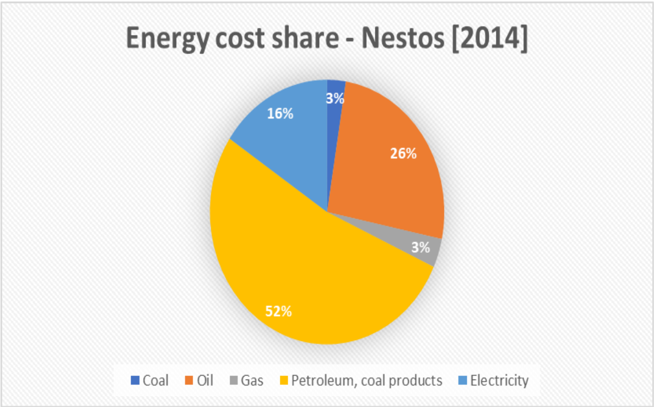


Figure 3. Total domestic demand for Adige case study [Italy] based on the SSP4 scenario.

For the two transboundary case studies, we have got a single estimate, computed as the weighted average of the results obtained for the two countries involved. The weighted average is calculated, based on the share of each production sector or commodity.

For example, Figure 4 presents the energy cost share for the Nestos case study, derived as the weighted average of the energy cost share in Bulgaria and Greece.



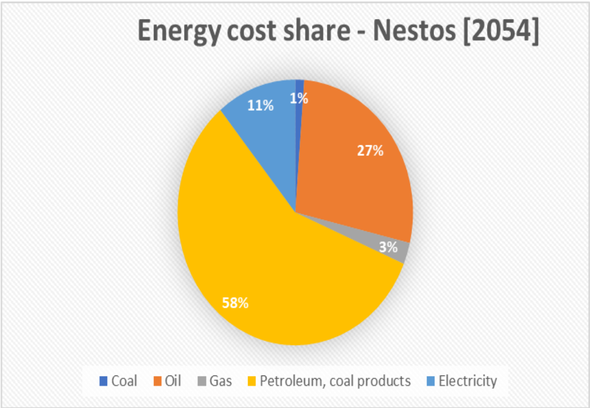


Figure 4. Energy cost share for Nestos case study [Bulgaria and Greece] based on the SSP4 scenario.

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

The GTAP database primarily provides economic data at the national level, whereas the case studies in this project consider river basins at a smaller geographical scale.

1.3.2 Downscaling the database at subregional level

The generated database is at the national level, consistent with the scope of our macroeconomic model. However, the case studies in this project focus on river basins at the subregional level, which requires further downscaling of the database. Therefore, to get more significant results, we applied to the G-RDEM model an additional module for regional downscaling. Presently, the module supports regional disaggregation at the NUTS2 (Nomenclature of Territorial Units for Statistics) level. This classification only incorporates second-level regional disaggregation, i.e., the basic regions in each country associated with the case studies (CSs), with population not exceeding 3 million. However, the CSs in this project are based on more refined regions that are closer to the river basins. Consequently, the database generated at the NUTS 2 regional level is mapped to the corresponding subregions. Since the results are expressed as percentage changes (unit-free), the data is downscaled by applying these percentage changes to the corresponding shares of each variable at the river basin level, using local data. A detailed explanation of how to apply the socioeconomic variables to the SDM variables is provided in section 5.

The European case studies are therefore associated to one or more sub-national regions, where applicable:

- Adige : Alto Adige, Trentino and Veneto
- Nestos: Yugozapaden and Anatoliki Makedonia, Thraki
- Lielupe: Latvia and Lithuania
- Jiu : Sud-Vest Oltenia
- Inkomati-Usuthu : South Africa

Notice that NUTS regional disaggregation is applicable exclusively to European regions, leaving South Africa as a single, aggregated region. For South Africa, we explore the possibility of using a different model, called DEMETRA, which would have allowed us to introduce more detailed regional disaggregation and some features relevant to South Africa, such as internal migration. However, after conducting a thorough feasibility analysis, we realized that this model present challenges in terms of data requirements, the ability to capture long-term structural changes, and modelling consistency compared to other case studies. By way of illustration, we show in the figures below some of our findings.

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

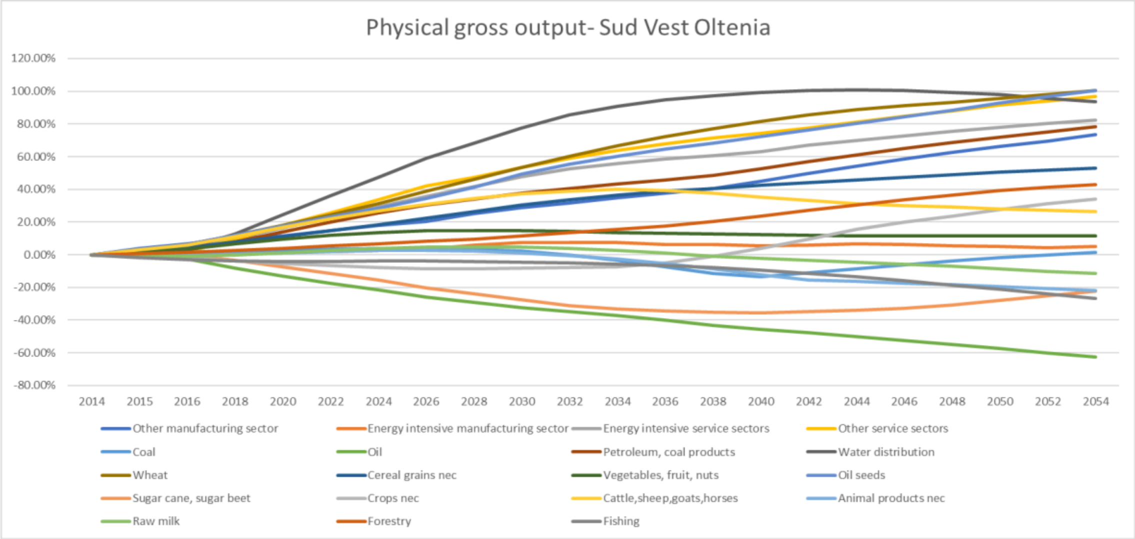


Figure 5. Physical gross output by production sectors for Jiu case study [Sud Vest Oltenia region in Romania] based on the SSP4 scenario.

Figure 5 shows the projected changes in gross output for various production sectors in the Sud Vest Oltenia region of Romania. The data spans from 2014, our baseline year, up to 2054. Several sectors exhibit substantial growth in total output (relative to their base level), including fishing, wheat cultivation, crop production, and livestock farming (cattle, sheep, goats, and horses), as well as oilseeds and petroleum production (whose initial level, however, is very low). In contrast, sectors like oil production, energy-intensive manufacturing, energy-intensive services, and animal product sectors are characterized by a declining trend in total output.

Figure 6 shows the shifting demand for energy products in the Yugozapaden region of Bulgaria. The results depict an overall increase in the demand for all energy products, with electricity registering the most significant surge, followed by petroleum, coal, and coal products, all projected to increase by 2054.

NEXOGENESIS STREAMLINING WATER RELATED POLICIES

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

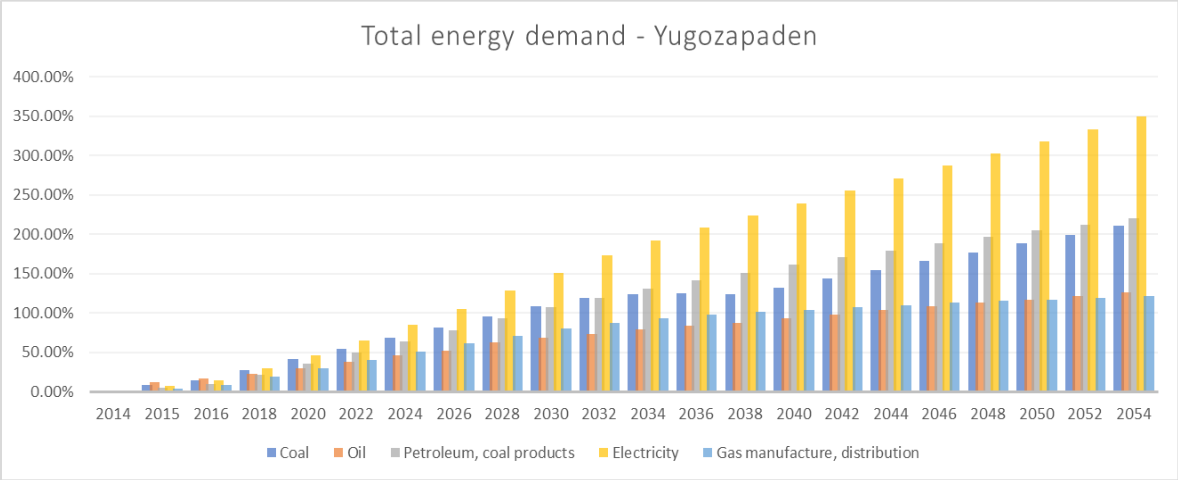


Figure 6. Total energy demand by the five energy carriers for Nestos case study [Yugozapaden region in Bulgaria]- based on the SSP4 scenario.

We analyzed both CO^2 and Non-CO^2 emissions for the case studies, considering two sources of emissions: emissions stemming from household activities and consumption, and emissions originating from production activities. The emissions trend associated with production activities is displayed in Figure 7 below for all regions linked with our case studies.

Notice a substantial, relative increase in CO^2 and Non-CO^2 emissions in the Yugozapaden region in Bulgaria, followed by South Africa and the Trentino and Alto Adige in Italy. Conversely, there is a relatively smaller, relative increase in total emissions in the Sud Vest Oltenia region in Romania, as well as in Latvia and Lithuania.

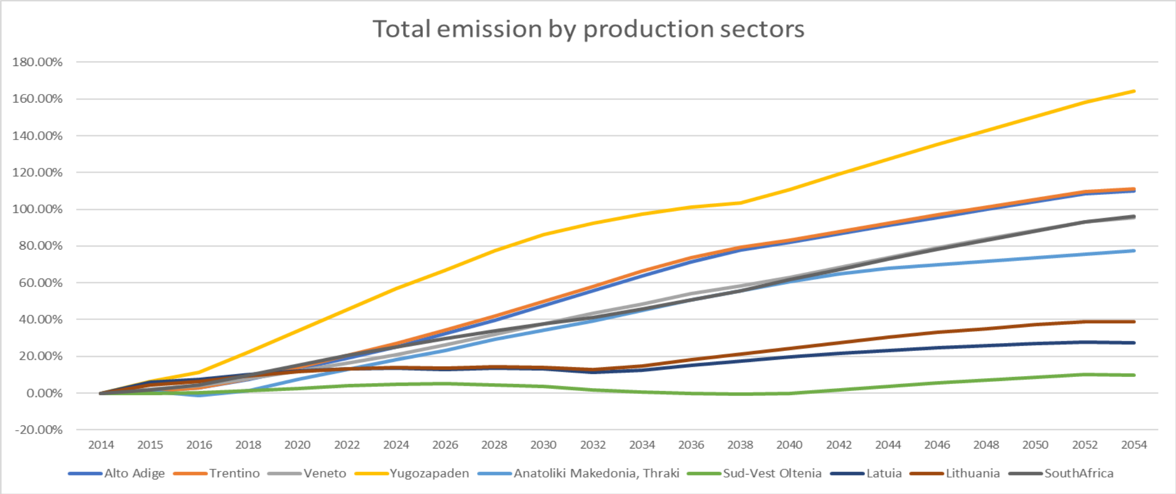
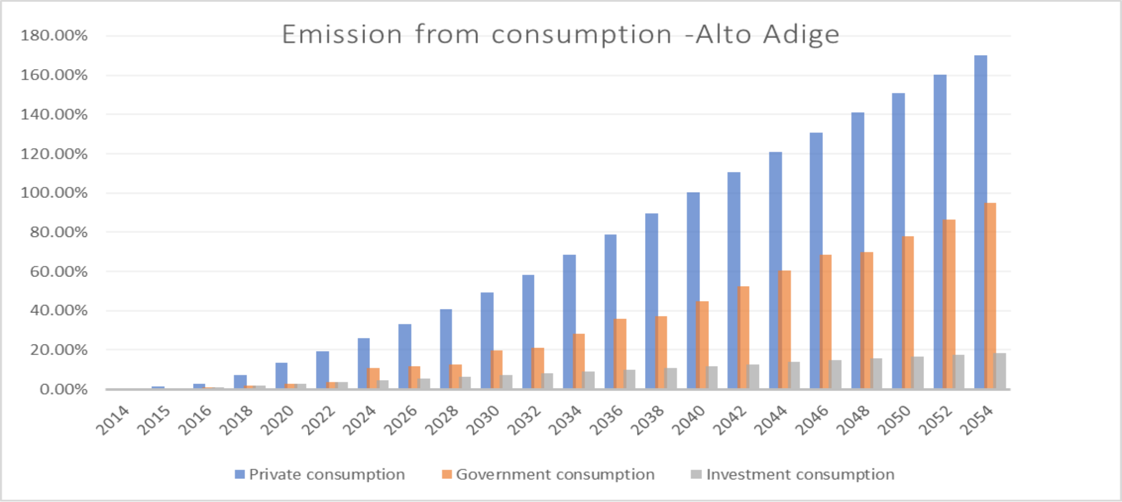


Figure 7. Total emission from production by sub-regions associated with case studies - based on the SSP4 scenario.

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

Figure 8 shows the (percentage changes of) emissions resulting from consumption in the Alto Adige region of Italy. These emissions are categorized into three distinct types: emissions from private consumption, government consumption, and investment



consumption. There is a notable (relative) increase in emissions originating from private consumption, with government consumption following as the second-highest relative increment to emissions.

Figure 8. Total emission from consumption for Adige case study - based on the SSP4 scenario.

2. Land use downscaling using the MAGNET model.

We adapted the MAGNETGrid (GTAP based) downscaling method to G-RDEM, in order to provide spatially explicit results on land related indicators for the river basin case studies.

The entire description of the downscaling method and some preliminary results (see two examples for the Adige case study in Figure 9 and 10) are available in the document of the milestone MS16. At the moment (M26 – October 2023), we are currently applying the downscaling method in all case studies and all four scenarios (SSP2 – RCP2.6, SSP2 – RCP8.5, SSP4 – RCP2.6, SSP4 – RCP8.5) and to share the results with the SDMs and the case studies in M28 (December 2023).

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

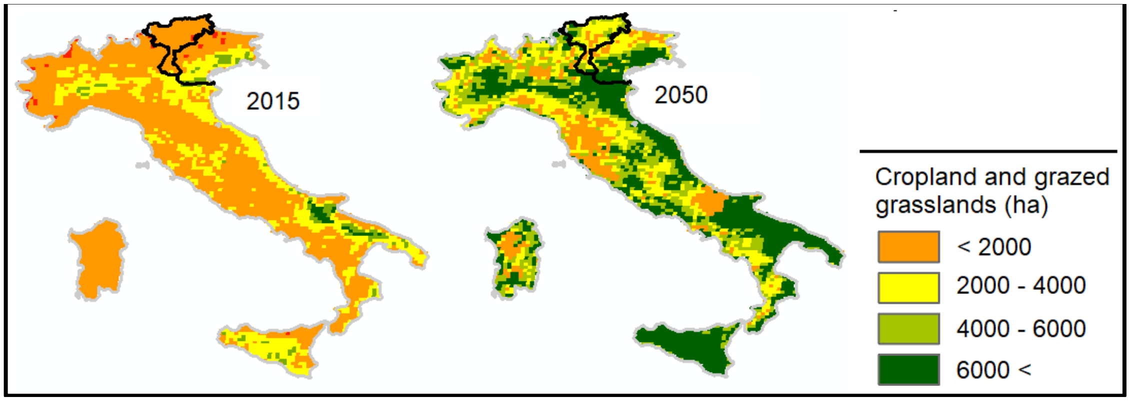


Figure 9. Spatio-temporal distribution of cropland and grazed grasslands over time. Grid cell size = 10,000 ha. Based on the G-RDEM projections, the combination of arable agriculture (excluding rice) and the grazing livestock sector will slightly increase the land demand in Italy by approximately 10%, especially triggered by the *osd* (oil seed) and *v_f* (vegetables and fruits) sectors.

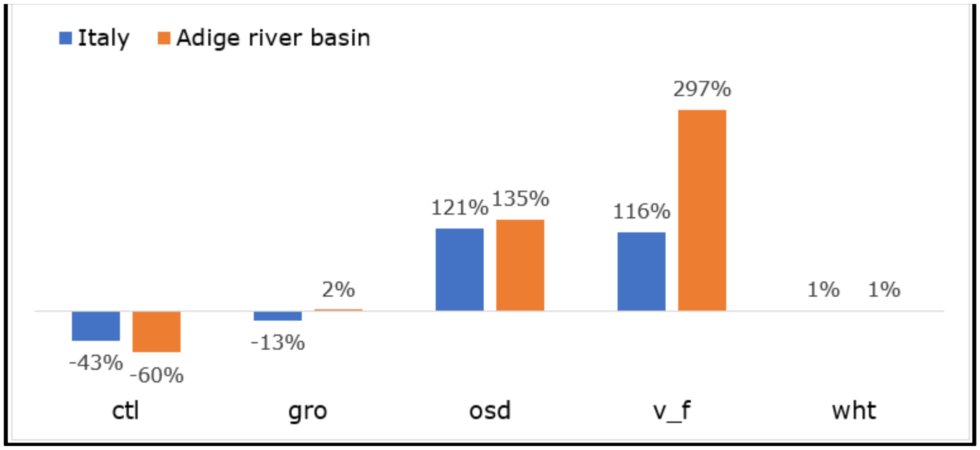


Figure 10. Variation in land demand over time per agro sector in Italy and in the Adige River basin. ctl: grazed grasslands; gro: cereals; osd: oil-seed crops; v_f: vegetables and fruits; wht: wheat. This graph shows that the largest demand driven sector will proportionally demand more land within the Adige River basin in comparison to the rest of Italy. Most likely these new areas for arable crops will expand at expense of grazed grasslands, which may have its area reduced by 60% in the Adige River basin area.

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

3. Uncertainty and sensitivity analysis

As part of the analysis conducted on NEXOGENESIS on socio-economic scenarios, an assessment of uncertainty on scenario variables has been conducted. To appropriately understand the meaning and objectives of this exercise, however, it should be kept in mind that socio-economic projections are not forecasts. Rather, they are quantifications of "possible futures", which were first described in qualitative terms, and from where more macroeconomic variables have been obtained, within an internally consistent framework, provided by a global, general equilibrium model.

The typical methodology in uncertainty and sensitivity analyses is conducting Monte-Carlo simulations, by randomly shocking a number of key parameters and exogenous variables, which then allows getting endogenous variables in terms of statistical distributions. However, in NEXOGENESIS, and specifically in the G-RDEM model employed to generate the scenario variables, a full-fledged Monte-Carlo analysis would be unfeasible and useless, for two main reasons. First, the model is complex and heavy, so that running multiple instances would require substantial computing time and resources. Second, there are so many given parameters in the model system, that the number of runs could easily grow exponentially.

To avoid these issues, we devised an alternative, we believe more sensible, modelling strategy. This is because results in the CGE model are mostly steered by one main exogenous driver, defining the SSPs: income per capita, at the country level. Therefore, we implemented a Monte-Carlo analysis, but limited to this key variable.

Starting from a given time series of national GDP at 5-years steps (2015-2055), we perturbed the growth rates by multiplying them by random variables, normally distributed, with mean 1 and standard error rising over time. This process was repeated 10000 times, thereby providing 10000 observations of GDP in each (case study relevant) country, for all time steps. Then, mean, and standard error were computed for each step. The mean is, of course, approximately the same as the deterministic GDP figure. The standard error, however, gets larger over time and provides a sort of "likelihood interval" around the mean.

By way of illustration, Figure 10 below shows this "cone of uncertainty" for the SSP4 projections of the Italian GDP (2015-2055).

NEXOGENESIS STREAMLINING WATER RELATED POLICIES

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

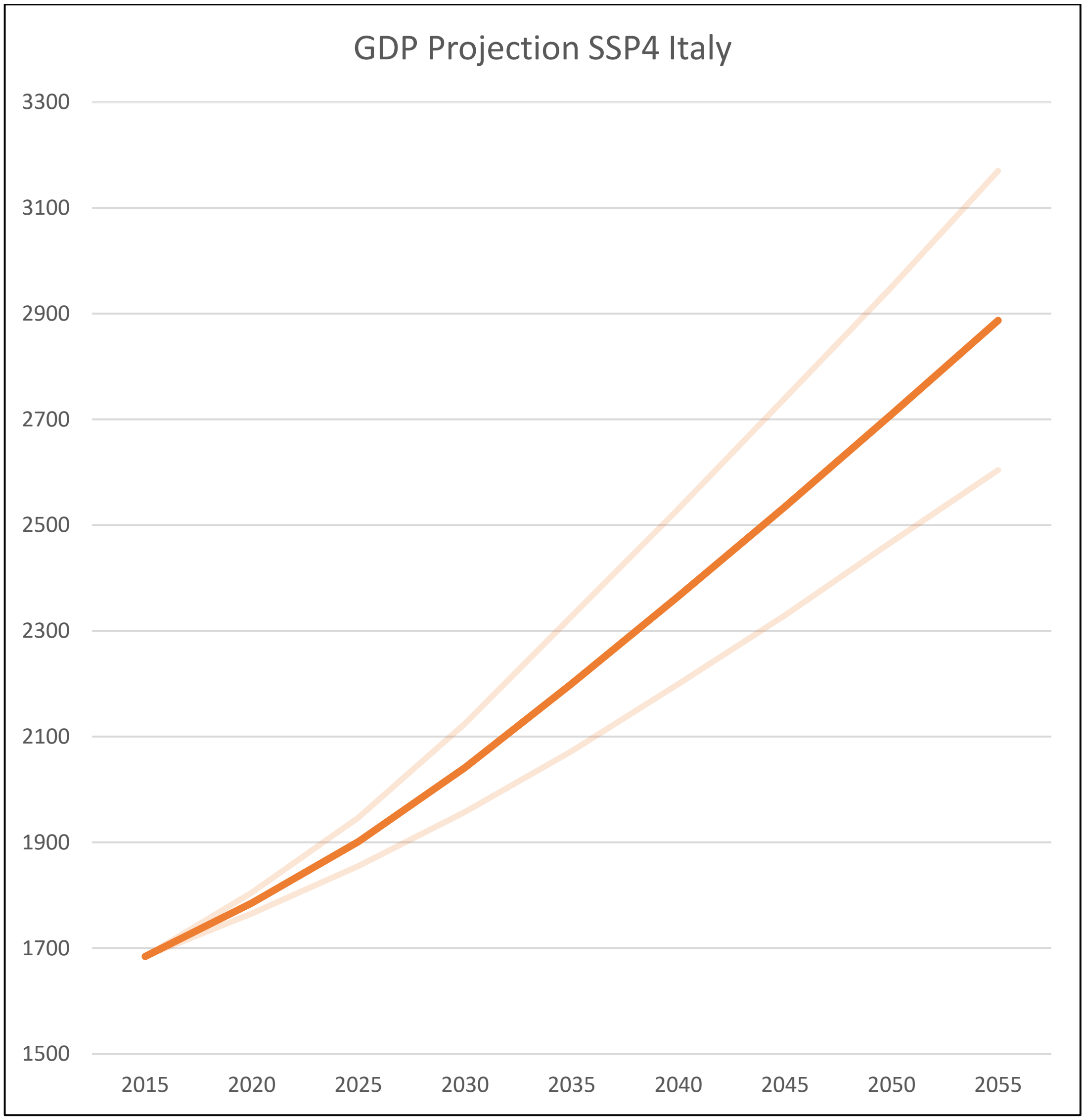


Figure 11. “Cone of uncertainty” (+/- s.e) for Italian GDP, based on SSP4 scenario.

For most of the endogenous variables in the CGE model, different but dependent on the GDP, we assumed that the causal link could be locally approximated by a proportional relationship. This implies that variable-specific standard error would be the same as that of the GDP *in relative terms*. In other words, if the GDP growth rate deviates from its central value by 5% in a certain year, so would do the dependent variable.

We deemed it necessary to introduce one exception to the method above, when referring to the households’ consumption. Indeed, it is well known that consumption of various items is not proportional to income. There are “necessities” (e.g., food) whose consumption grows less than proportionately, there are “luxuries” (e.g., tourism services) that grow more than proportionately. Using the GTAP data base as a panel data set, we econometrically

NEXOGENESIS STREAMLINING WATER RELATED POLICIES
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003881

estimated the item-specific "income elasticity", which is the percentage variation in consumption for a 1% rise in income or total expenditure. Results are shown below and applied to get the relevant standard error for these variables.

Estimated elasticity parameters for agricultural commodities.			
Number	**Code**	**Estimated elasticity par**	**Description**
1	pdr	1.182613	Rice: seed, paddy (not husked)
2	wht	0.9526314	Wheat: seed, other
3	gro	1.03387	Other Grains: maize (corn), sorghum, barley, rye, oats, millets, other cereals
4	v_f	1.531067	Veg & Fruit: vegetables, fruit and nuts, edible roots and tubers, pulses
5	osd	0.7483922	Oil Seeds: oil seeds and oleaginous fruit
6	c_b	1.225351	Cane & Beet: sugar crops
7	pfb	1.757455	Fibres crops
8	ocr	0.6810341	Other Crops
9	ctl	1.113775	Cattle and bovine animals
10	oap	0.901719	Other Animal Products
11	rmk	0.5631789	Raw milk
12	wol	0.6081909	Wool: wool, silk, and other raw animal materials used in textile
13	frs	0.2681766	Forestry
14	fsh	1.108727	Fishing

Table 2. Estimated elasticity parameters of private consumption for agriculture commodities.

In practice, the uncertainty assessment we conducted is reflected in the series of tables delivered to the various case study units where, in addition to the central value of the (percentage growth) for each scenario socio-economic variable, we associate the corresponding standard error. Here below we show some examples of the findings one could find in the data repository.

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

Physical gross output- Latvia										
Production Sectors	**2015**	**2015_**	**2020**	**2020_**	**2030**	**2030_**	**2040**	**2040_**	**2050**	**2050_**
	Level	s.e	Level	s.e	Level	s.e	Level	s.e	Level	s.e
ALL sectors	5.06%	0%	20.04%	2.21%	44.83%	6.10%	67.42%	7.56%	84.47%	6.65%
Other Extraction (former omn)	-1.56%	0%	3.32%	0.01%	14.35%	0.02%	30.97%	0.02%	45.20%	0.02%
Metal products	12.49%	0%	28.88%	0.04%	55.56%	0.10%	95.06%	0.13%	128.56%	0.12%
Computer,electronic,optical prod	27.70%	0%	43.24%	0.06%	59.32%	0.16%	94.88%	0.21%	122.97%	0.19%
Electrical equipment	28.32%	0%	44.24%	0.02%	64.03%	0.06%	108.46%	0.08%	144.60%	0.08%
Machinery and equipment nec	18.79%	0%	35.11%	0.03%	60.80%	0.08%	108.80%	0.10%	142.92%	0.09%
Motor vehicles and parts	17.83%	0%	40.82%	0.02%	80.53%	0.07%	125.55%	0.09%	156.49%	0.08%
Transport equipment nec	2.80%	0%	13.37%	0.00%	40.71%	0.01%	77.94%	0.02%	90.92%	0.01%
Manufactures nec	6.90%	0%	18.61%	0.03%	38.02%	0.07%	68.48%	0.09%	93.17%	0.09%
Bovine meat products	17.38%	0%	25.59%	0.00%	27.33%	0.01%	24.51%	0.01%	16.49%	0.00%
Meat products nec	14.21%	0%	21.87%	0.01%	31.57%	0.03%	38.41%	0.04%	32.53%	0.03%
Vegetable oils and fats	26.95%	0%	24.23%	0.01%	25.82%	0.02%	37.37%	0.02%	55.04%	0.02%
Dairy products	19.39%	0%	40.86%	0.03%	72.63%	0.08%	70.58%	0.09%	47.92%	0.07%
Sugar	16.15%	0%	17.00%	0.00%	24.72%	0.00%	34.72%	0.00%	40.81%	0.00%
Food products nec	8.87%	0%	15.28%	0.05%	24.29%	0.12%	27.12%	0.13%	26.03%	0.10%
Beverages and tobacco products	5.78%	0%	40.25%	0.02%	104.62%	0.07%	136.08%	0.09%	145.83%	0.07%
Textiles	32.23%	0%	66.37%	0.01%	127.01%	0.03%	180.71%	0.04%	201.28%	0.04%
Wearing apparel	31.57%	0%	60.93%	0.01%	111.22%	0.04%	159.72%	0.05%	184.30%	0.05%
Leather products	46.50%	0%	107.49%	0.01%	222.70%	0.02%	326.86%	0.03%	368.85%	0.03%
Wood products	31.03%	0%	63.53%	0.13%	98.90%	0.35%	147.48%	0.46%	190.39%	0.43%
Paper products, publishing	17.34%	0%	40.05%	0.03%	77.03%	0.09%	113.37%	0.11%	138.48%	0.10%
Chemical products	29.93%	0%	51.20%	0.05%	70.67%	0.12%	89.97%	0.14%	98.66%	0.11%
Basic pharmaceutical products	40.30%	0%	65.03%	0.02%	94.76%	0.06%	141.57%	0.08%	183.51%	0.07%
Rubber and plastic products	23.28%	0%	46.19%	0.02%	75.84%	0.06%	106.35%	0.07%	125.16%	0.06%
Mineral products nec	8.09%	0%	15.99%	0.02%	27.36%	0.06%	47.08%	0.07%	62.62%	0.06%
Ferrous metals	27.46%	0%	38.44%	0.02%	48.78%	0.06%	74.38%	0.07%	95.28%	0.06%
Metals nec	39.70%	0%	43.17%	0.02%	38.31%	0.04%	53.93%	0.04%	64.11%	0.04%
Petroleum, coal products	23.81%	0%	39.12%	0.00%	59.64%	0.00%	85.91%	0.00%	122.41%	0.00%
Electricity	9.76%	0%	25.83%	0.05%	53.17%	0.14%	80.44%	0.18%	108.69%	0.17%
Transport nec	4.31%	0%	18.46%	0.14%	41.55%	0.37%	58.18%	0.45%	69.74%	0.38%
Water transport	15.36%	0%	30.58%	0.03%	59.23%	0.10%	94.65%	0.13%	129.75%	0.12%
Air transport	7.76%	0%	31.30%	0.02%	76.69%	0.07%	115.22%	0.09%	151.36%	0.08%
Water distribution services	9.94%	0%	27.47%	0.01%	60.99%	0.04%	88.24%	0.05%	114.17%	0.05%
Construction	-28.92%	0%	-32.90%	0.12%	-30.76%	0.28%	-15.94%	0.36%	-8.64%	0.31%
Trade	-0.26%	0%	28.27%	0.20%	80.54%	0.66%	110.55%	0.83%	130.76%	0.72%

Table 3. Gross output by production sectors along with its corresponding standard errors for Lielupe case study [Latvia region].

4. A guide to data in the common project repository

So far, we have introduced four result sets available in Excel file format, accessible in the common drive [SURFdrive]. Each of these result sets offers both graphical and tabular representations of the estimated socio-economic scenarios. Furthermore, they include a separate sheet encompassing model descriptions, regional and sectoral aggregations, along with a summary of the results. The files can be found in the subdirectory T2.1, inside the directory WP2.

Socio-economic variables are presented in the datasets as percentage change variations from the base year. This is due to a combination of factors. First, data employed in the G-RDEM macroeconomic model are expressed as values, and prices are initialized at unity. Therefore, monetary units of measure are used for all quantity variables, whereas by expressing them as percentage variations, we are making our estimates independent from the units of measure. Second, as variables in the SDM/NEPAT models do not generally coincide with the variables generated by G-RDEM, a percentage variation can be easily applied to local variables, to the extent that they are correlated with some macro variables. For example, if local demand for renewable energy is considered in some case study, one

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

may want to estimate future levels of demand, by applying to the initial level the calculated variation for the aggregate regional demand for energy. Of course, this process is not a mechanical one, and here is where stakeholders' assessment/validation is necessary, so that – for example – the change in local demand may be estimated to be more or less proportional than the aggregate one, or causally related to other variables.

1. NXG_Data_WP2_T2.y_first GRDEM results_V1.xlsx - SURFdrive

- **Regional Aggregation:** Results are aggregated into 10 regions at the national level.
- **Factor and Sectoral Aggregation:** Data is categorized into 64 production sectors and 5 factors of production.
- **SSPs Scenario:** The scenario used is SSP4.
- **Unit of Measurement**: All values are represented as percentages (%) relative to the baseline year.
- **Time Period:** The data covers the period from 2014 (baseline year) to 2054.
- **Results:** This file contains both graphical and tabular presentations of selected variables. [see Table 1]

2. NXG_Data_WP2_T2.y_Second_GRDEM_NUTS2_results_V2.xlsx - SURFdrive

- **Regional Aggregation:** Results are further disaggregated into 9 sub-regions based on the NUTS2 regional classification.
- **Factor and Sectoral Aggregation:** Data is categorized into 64 production sectors and 5 factors of production.
- **SSPs Scenario:** The scenario used is SSP4.
- **Unit of Measurement**: All values are represented as percentages (%) relative to the baseline year.
- **Time Period:** The data covers the period from 2014 (baseline year) to 2054.
- **Results:** This file contains both graphical and tabular presentations of selected variables. [see Table 1]

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

3. NXG_Data_WP2_T2.y_Third_GRDEM_SSP2_NUTS2_results_V3.xlsx - SURFdrive

- **Regional Aggregation:** Results are further disaggregated into 9 sub-regions based on the NUTS2 regional classification.
- **Factor and Sectoral Aggregation:** Data is categorized into 64 production sectors and 5 factors of production.
- **SSPs Scenario:** The scenario used is SSP2.
- **Unit of Measurement**: All values are represented as percentages (%) relative to the baseline year.
- **Time Period:** The data covers the period from 2014 (baseline year) to 2054.
- **Results:** This file contains both graphical and tabular presentations of selected variables. [see Table 1]

4. NXG_Data_WP2_T2.y_Fourth_GRDEM_SSP2_NUTS2_results_V4.xlsx - SURFdrive

- **Regional Aggregation:** Results are further disaggregated into 9 sub-regions based on the NUTS2 regional classification.
- **Factor and Sectoral Aggregation:** Data is categorized into 64 production sectors and 5 factors of production.
- **SSPs Scenario:** The scenario used is SSP2.
- **Unit of Measurement**: All values are represented as percentages (%) relative to the baseline year.
- **Time Period:** The data covers the period from 2014 (baseline year) to 2054.
- **Results:** This file contains graphical and tabular presentations of selected variables. Additionally, it includes results of standard errors and estimated elasticity parameters of private consumption from the Uncertainty and Sensitivity Analysis.

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

Data sharing: Data will be available through the Zenodo Repository, in compliance with H2020 regulations. Therefore, these databases will be made accessible once the corresponding deliverables are available via the Open Access (OA) repository through the OpenAIRE service.

5. How to apply the Socio-economic Projections to variables in the case study SDMs.

The set of variables considered in the projections does not generally match the set of variables in the system dynamic models (SDMs) of the various case studies. First, the geographical scale is different, which for economic data is related to administrative regional boundaries, not to physical entities, like river basins. Second, sectors are aggregated so that, for instance, we may get information about trends in regional energy production (or consumption), but not about hydroelectric energy. Furthermore, scenarios are expressed as unit-free, percentage changes, requiring the knowledge of baseline levels for all variables in the SDM.

Once the baseline levels for variables in the SDM have been determined, the scenario projections can inform about the future evolution of them, once a correlation between one variable in the SDM and one (or more) variables in the scenario has been identified.

For instance, the simplest case is when a sectoral growth rate could be taken as a reasonable proxy for a SDM variable, e.g., when the growth rate of regional production of fruits can be assumed as the growth rate of apples production in the river basin.

In other cases, a more elaborate relationship should be defined. If we know that, for example, physical constraints do not allow the expansion of hydroelectric power generation locally, then the SDM growth rate could be a fraction (possibly zero) of the sectoral growth in the region.

Summing up, the application of scenarios to SDM variables necessitates: (1) base levels of variables, in some unit of measure; (2) linking local to regional variables; (3) establishing a causal relationship, over which the proxies are based.

6. Contribution to Objectives 2 and 3 of the project

Objective 2 of the NEXOGENESIS project states:

"*Reduce uncertainties of how new policies and stakeholder behavior affect the nexus through the integration Self-Learning Nexus Assessment Engine (SLNAE) output and policy*

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

feasibility assessments, along with validation of findings by stakeholders. This will lead to better cross-sectoral policy-making and governance, leveraging synergies and avoiding trade-offs to support trade developments. Support the goals of the EU 2030 vision for smart, sustainable and inclusive growth. Strong stakeholder engagement aids towards SDG17."

Socio-economic reference scenarios, as provided by WP2, are an essential element in the construction of the NEPAT tool (formerly SLNAE). They represent a possible future state, against which policy impacts can be assessed, as well as possible different policy priorities that may emerge, in the various cases, at a longer horizon.

Objective 3 of the NEXOGENESIS project states:

"*Develop and apply a new WEFE Nexus Footprint in a similar vein to the urban water blueprint to track progress of policy objectives, e.g. greening of the CAP, SDG goals. Detailed analysis of constituent indicators, provided by the SLNAE, provides the user with insights regarding where interventions and investments are necessary*."

Nexus footprints may refer to present time, business-as-usual future, or policy-determined future. Socio-economic reference scenarios, as provided by WP2, are of course fundamentals in the construction of future indicators.

7. Conclusion

The generated socio-economic database provides key variables for the development of SDMs tailored to each case study, including detailed projections of production, demand for consumption goods, energy, and emissions associated with the nexus of water, energy, and land use. These projections, while not forecasts, offer insight into potential future scenarios. We accounted for uncertainty through integrated uncertainty and sensitivity analysis. This ensures robust, well-rounded results that support policy development in key areas and contribute to the assessment and mitigation of climate change impacts.

Looking ahead, the database will be further refined through collaboration with other work packages (WPs) and case study (CS) team members to enhance data mapping and ensure the results remain valid and consistent. This collaborative effort will ultimately aid in the development of NEPAT, a critical tool that will provide deeper insights for policy analysis and development, helping shape effective strategies for addressing nexus challenges.

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