

# Introduction to the EUCalc model

Cross-Sectoral Model description and documentation



<b>Project Acronym and Name</b>	EU Calculator: trade-offs and pathways towards sustainable and low-carbon European Societies - EUCalc
<b>Document Type</b>	Documentation
<b>Work Package</b>	all WPs
<b>Document Title</b>	Introduction to the EUCalc model Cross-Sectoral Model description and documentation
<b>Main authors</b>	Julien Pestiaux, Vincent Matton, Michel Cornet, Luis Costa, Bernd Hezel, Garret Kelly, Juergen Kropp, Ana Rankovic, Emily Taylor
<b>Partners in charge</b>	Climact
<b>Contributing partners</b>	PIK, SEECN, all
<b>Release date</b>	16 September 2019 (version 2.0)
<b>Distribution</b>	<i>All involved authors and co-authors agreed on the publication.</i>

### Short Description

This document aims to provide an overview of the EUCalc model. The first part should give the reader the main keys to understand what is the EUCalc model in terms of levers and scenarios creation. The modelling approach, presented in the second part, will focus the reader on a more detailed description of the modelling practices. It will give the reader a technical overview of the model. The document is continuously updated.

### Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

# TABLE OF CONTENT

<b><i>I. Overview of the European Calculator modeling approach</i></b> .....	<b>3</b>
<b>1. A brief summary</b> .....	<b>3</b>
<b>2. A lever-based bottom-up model</b> .....	<b>4</b>
2.1. Ambition levels .....	4
<b>3. Scenario creation</b> .....	<b>4</b>
<b>4. Output of the model</b> .....	<b>6</b>
<b><i>II. EUCalc modeling approach</i></b> .....	<b>7</b>
<b>1. A brief classification of energy models and where EUCalc fits</b> .....	<b>7</b>
<b>2. The EUCalc model</b> .....	<b>10</b>
2.1. Characteristics .....	10
2.2. Space resolution .....	10
2.3. Data-driven model structure .....	11
2.4. Exploring and reducing the solution space .....	12
2.5. Open to expert validation .....	13
<b>3. Feedback loops, country disaggregation and interaction</b> .....	<b>13</b>
3.1. Feedback loops .....	13
3.2. Disaggregation by Country .....	18
3.3. Historical data completion .....	20
3.4. Ambition curves shapes .....	22
<b>4. Interaction between modules</b> .....	<b>25</b>
4.1. Interaction with other models (GTAP-EUCalc) .....	26
<b><i>III. Quality management</i></b> .....	<b>27</b>
<b>1. Input data quality</b> .....	<b>27</b>
<b>2. Calibration on historical data</b> .....	<b>28</b>
<b>3. Modelization choice &amp; uncertainties</b> .....	<b>29</b>
<b><i>IV. Appendix A: KNIME modeling</i></b> .....	<b>31</b>
<b>1. Knime environment</b> .....	<b>31</b>
<b>2. Accessing the model</b> .....	<b>32</b>
<b>3. Definition of a KNIME workflow</b> .....	<b>32</b>
<b><i>V. Appendix B: List of levers</i></b> .....	<b>34</b>
<b><i>VI. References</i></b> .....	<b>36</b>

# I. OVERVIEW OF THE EUROPEAN CALCULATOR MODELING APPROACH

Part I of this document aims to provide an overview of the EUCalc model. It introduces the underlying key concepts and the modelling philosophy. For more technical descriptions, i.e. in terms of the implementation of the model we refer to Part II of this document.

## 1. A brief summary

The European Calculator (EUCalc model) is a model of energy, land, materials, product and food systems at European and member-states<sup>1</sup> levels for representing GHG emissions dynamics until 2050. The model can be applied for delineating emission and sustainable transformation pathways at a European scale but may also be used to study the impact of a specific member state on European-level policy.

While optimization models are often the norm in low carbon analysis (e.g., economic optimization), the massive uncertainties arising from taking a long-term horizon as 2050 or 2100 mean that optimizing on certain factors like costs is at the least extremely challenging, meaning these models should be complemented with other approaches to possible low carbon trajectories, particularly if one wants to include the potential of breakthroughs or non-linear changes. **Addressing these system dynamics with a bottom-up driver- and lever-based model provides a very powerful and complementary alternative.** The EU Calculator has these 2 concepts at its core; 1) it defines calculation sequences based on material, energy and emissions drivers, 2) and then it sets a range of ambition levels on the drivers that are most important and where the user can define projected levels. These drivers are called *levers* and they are at the center of the scenario creation logic.

Estimates of the end-use service demand (e.g. buildings heating, appliances usage, car road travel, freight demand, etc.), of the demographic evolution, and of the techno-economic trends are (exogenously) defined by the user for each region (or for Europe as a whole) and mapped by so-called levers representing various ambition levels for policy making. In addition, the user can choose for which climate scenario he/she wants to base his/her calculation (i.e., pursuing the ambition to stabilize the temperature rise to 1.5, 2, 3 degrees or above pre-industrial levels by 2100). Based on the *lever ambition* the user specifies, the EUCalc model supplies the energy to fulfill the resulting demand (demand-driven model). The calculation of GHG emissions is subsequently based on the amount and the type of energy used.

Rather than calculating optimal pathways the model allows the user to choose the *ambition level* of each individual lever (from a reference level up to maximum technical ambition) and thereby explores different scenarios of pathways to 2050. The investment costs of each pathway are estimated by adding the annual capital expenditures (e.g., for new infrastructures or assets), operational costs (e.g., maintenance) and fuel costs. Air pollution costs from PM2.5<sup>2</sup> emissions are also available. Other externalities (such as reduced noise, climate change damages, or biodiversity conservation) are not accounted for cost estimates, but included in the calculation workflow and available as output of the model.

---

<sup>1</sup> EU-28 + Switzerland

<sup>2</sup> Fine particles with a diameter of 2.5 µm or less

## 2. A lever-based bottom-up model

Before creating scenarios it is crucial to take a sectoral view to understand what types and levels of change are technically possible in each sector. For each defined lever a range of ambition levels was derived on the basis of expert/stakeholder elicitation (more than 1000 experts<sup>3</sup> from academics, businesses, and NGOs, etc.) and thorough literature studies describing a range of potential futures for the respective sectors. Our work has built on comments from these experts to better identify and understand the key implications for Europe of a move to a low carbon society. This concept is described below and each sector documentation goes at length on the reasoning behind the ambition levels defined. These *levers* and their possible *ambition levels* are the building blocks of the EUCalc model. They allow the user to construct possible pathways to 2050 and beyond. The approach looks not only at 2050 as an end point, but also at the sequence of changes that would need to occur over the next 40 years, i.e. for the implementation of a successful climate protection policy.

### 2.1. Ambition levels

*Levers* allow the user to interact with the model and to build scenarios by choosing *ambition levels* (cf. refer for a detailed list of levers to Appendix B). The section describes how *ambition levels* are defined at Country level and how trajectories are chosen between the base year (2015) and 2050.

The EUCalc model is controlled using a range of levers that represent changes one could make to mitigate climate change until 2050. Each lever has four different levels of effort. These four levels offer a broad variation of mitigation choices and sustainability impacts, e.g. including intermediate levels. Consequently, the model can provide a wide range of pathways arising from combinatorics of all levers and level settings (cf. Table 1).

<b>Level 1</b> This level contains projections that are aligned and coherent with the observed trends.	<b>Level 2</b> This level is an intermediate scenario, more ambitious than observed trends but not reaching the full potential of available solutions.
<b>Level 3</b> This level is considered very ambitious but realistic, given the current technology evolutions and the best practices observed in some geographical areas.	<b>Level 4</b> This level is considered as transformational and requires large additional efforts such as strong changes in the way society is organized, a very fast market uptake of deep measures, an extended deployment of infrastructures, major technological advances and breakthroughs (but without relying on new fundamental research), etc.

Table 1. Definition of ambition levels

## 3. Scenario creation

The EUCalc model is built to test a variety of low carbon trajectories or *scenarios* and to understand their key implications for policy planning. Those scenarios should support policy making by giving an indication of the required evolution of key indicators to reach the GHG reductions: scenarios explore the impact of switching certain group of parameters on/off so as to better understand the impact of certain choices (energy efficiency and lifestyle changes, technological options, etc.).

---

<sup>3</sup> The contributions of experts are gratefully acknowledged. However, as mentioned below, the responsibility of the analyses however lies with the partners of EUCalc. Therefore, the experts and stakeholders consulted do not necessarily endorse the analyses or the conclusions of our work.

Although the model assesses the cost implications of each scenario, i.e. based on the evolution of the investments and operational and fuel costs, it is an accounting-type model as opposed to optimisation or simulation models. This implies that EUCalc model does not adopt a cost optimisation approach and does not identify the least costly way of potential 2050 targets. The aim, instead, is to look at what might be practically and physically achievable in each sector over the next 30 years under different assumptions. The EUCalc model can be used via a so-called [Transition Pathways Explorer](#) (TPE) which allows users to explore their own choices in real time. This flexibility and transparency makes it particularly suitable for policy makers/decision makers to test a wide range of scenarios. The scientific underpinning and the online documentation makes it possible to be engaged in discussions with people interested in transformation challenges.

Across sectors, a large set of levers and trajectories are modelled (e.g. transport demand per person; insulation level for refurbished houses; lifetime of certain products like cars, efficiency and type of steel production; offshore wind capacity installations) driving energy demand and supply projections. Thus, they cover a broad range of possibilities, testing the boundaries of what happens in business-as-usual cases or what might be technically feasible. They are not based on specific assumptions about future policies and their impacts and should not be interpreted as such. The EUCalc approach aims to achieve as much consistency as possible across the different sectors in terms of 'level of ambition'. Therefore a 'level 2' effort in one sector is broadly comparable to a 'level 2' effort in a different sector.

Concretely, a *scenario* is created by choosing a combination of effort (or ambition) levels for the full set of drivers available to the user. These are grouped by category of issues (e.g., lifestyle and technology) and sectors (e.g., buildings and transport). Those drivers can be described as either *trajectories* on which authorities have little or no influence (e.g., demographic trends, evolution of energy prices) in contrast to *levers* which can be directly influenced. Both of these types of drivers can be defined by the user to project the evolution of all the outputs of the model, including energy consumption, production, and investment and operation costs. Higher ambition is always defined as having a stronger impact on reducing GHG emissions. Ambition levels for each *lever* range from [1] 'current trends' or minimum legal requirements to [4] maximum technical potential. The user of the model can therefore choose to limit the effort in one sector or category and to focus all efforts [e.g., maximum ambition of level 4] on another sector to reach the energy and GHG emissions target it has set for Europe or a particular member state.

This flexibility has its drawbacks: choices on the levels of levers and ambition levels must be made in a coherent manner, since the model itself does not reflect the full complexity of a real world system. and judgments are required to combine various ambition levels or sector trajectories. The users of the model must themselves make these judgements to avoid non-plausible combinations. Similarly, the model does not account for all possible feedbacks between different sectors. Changes in one sector might be expected to have a rebound effect in another sector, and not all of these are reflected in the model. The [Transition Pathways Explorer](#) (TPE) will include warnings on key possible issues and restrictions.

Various key dimensions are then visualized to understand the impact of the scenarios, and they can help the user to refine them: if the evolution of the energy demand is too strong, more focus can be set on energy efficiency or behavior changes, if the supply mix is not carbon free, the low-carbon levers will need to be pushed to a higher ambition level, if biomass imports are too high other alternatives need to be found, etc. Exogenous evolutions such as demography need to be defined consistently in light of the scenario being developed by the user. Global dynamics can also be reflected in the carbon budget.

Once several iterations of refinements have been made, the user can zoom in on sector specific alternatives as well: should transport be decarbonized based on electric cars or based on e-fuels?

The user can then explore scenario implications, getting a clear view on where his scenario will lead Europe. Typical questions are: can we reach net-zero emissions by 2050? What will this imply for each sector? Will this require importing clean energy? What is the role of the land-use sectors in my scenario? Will industry production decline/increase in certain sectors? Will we need to plant new forests? Extract more wood from them? Will all this lead to more or less jobs? What level of investments will it require? Will this be compensated by lower fuel costs?

Analyzing these outputs can help the user refine its scenario further, to reach what he considers the optimal transition for Europe.

If the user is not trying to define only one scenario then creating contrasting ones can help to better understand what these alternative paths can mean for Europe: a technology-based transition or one more focused on changes in societal patterns?

## 4. Output of the model

The EU-Calc model computes different types of *impacts*: the energy consumption and GHG emissions at a country level, resource depletion (water, fossil fuel, lands) and other environmental impacts such as biodiversity, and socio-economic impacts such as employment and air pollution impacts. For more details about the scope of those impacts' calculation and about the methodology, please refer to the related module documentation documents.

<b>Energy</b>	Total energy consumption in EU Energy consumption by energy vector and consumption sector
<b>GHG emissions</b>	Total GHG emissions in EU by country
	GHG emissions imported/exported outside of EU (linked to import/export of products/resources)
<b>Pollution</b>	Air pollution: Fine particulate matter
	Fertilizer input
<b>Resources use &amp; availability</b>	Water availability, consumption
	Minerals & rare earths availability
	Fossil fuels use
	Land and forest use
	Food waste
	Renewable resources: wood
<b>Socio-economic</b>	Competitiveness & employment
	Food import
	Material import
	Impact of air quality
<b>Environmental impacts</b>	Biodiversity : impact of land use allocation, species habitat
	World temperature change

Table 2. Modelized impacts

In terms of *scenarios*, the model includes the most important and known decarbonizing solutions grouped in the different levers. For more details on decarbonizing options included in each module, please refer to the specific sector module documentation (lifestyle, building, transport, land-use, supply, manufacturing, etc.). Table 3 summarizes the macro-solutions that are included across those modules.

<b>Avoid</b>	Demand reduction (e.g. lower demand of goods and manufactured products, lower transport demand, lower temperature in buildings, etc.)
	Sharing economy & Circular economy
<b>Shift</b>	Fuel switch to biofuels, e-fuels, electricity
	Shift to existing cleaner technologies (e.g. buildings renovation, transport modal shift, etc.)
	Shift to new breakthrough technology
	Automation (of buildings, vehicles, etc.)
<b>Improve</b>	Energy efficiency improvements
	CCU/CCS

Table 3. Ambition of the model in terms of decarbonization options & scenarios

## II. EUCALC MODELING APPROACH

### 1. A brief classification of energy models and where EUCalc fits

The EUCalc model is a simulation model, driven by people *activities* in a given *context* and reflects the impact of using technologies to perform the activities on *energy, emissions, socioeconomic impacts* and *environment & resources*. It also assesses links to the economy, to policies and to transboundary flows.

To start a simulation users define the inputs to the calculator by making choices using a number of *levers*. These levers typically make a change in either the supply or demand of energy in a particular sector, for example building nuclear power stations, or reducing the distance people travel by car.

The model resolution for real time results is the following:

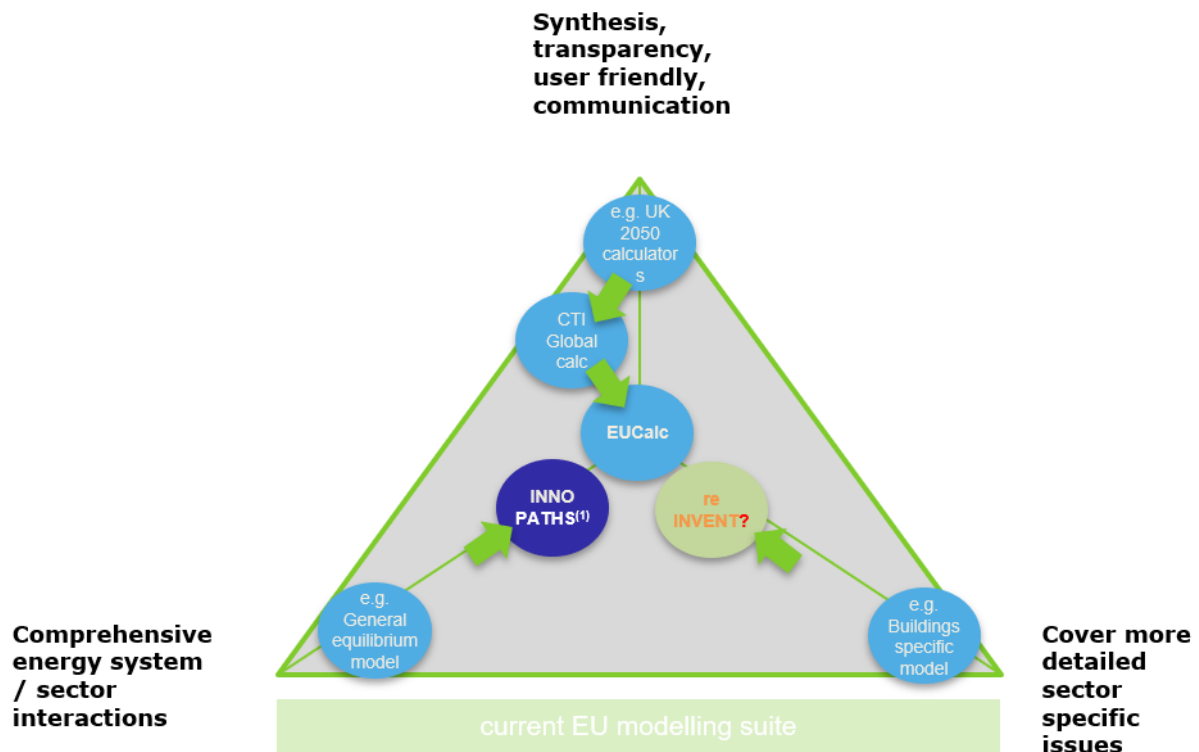
- The spatial resolution of the calculator is country level for all EU-28 + Switzerland;
- In terms of time resolution, historical data is collected until 2015 (which is the reference year), then projections and ambition levels are assessed yearly between 2020 and 2050. In addition, energy consumption and production profiles detail the evolutions along different days and within the days.

The model is described as directed graph in the sense it resembles a calculation tree where each node is calculated once. This is required to perform the calculations with the required resolution in real time. This has two major implications.

- It does not perform feedback loops. Because many feedback loops have been discussed during the model design, the strategy to address them is detailed in section II.3.1.
- It does not perform optimisations. However, it is deeply linked to several models performing optimisations (e.g. GTAP provides several inputs or suggested lever positions). Likewise, it is expected to provide inputs for optimisation models such as GTAP and TIMES (see section II.4.1).

Models are generally classified along three desired objectives. These goals are unfortunately not simple to combine in one model, which means that very often approaching one of them makes it harder to reach the other two fully.





NOTE: (1) Energy system decarbonisation simulator online tool

Figure 1. Model classification.

The EUCalc model's origin are DECC 2050 calculators. These excel and web-based simulation models provided great value by being synthetic, transparent, and user friendly.

These models are typically used as eye-openers, especially in the first phase of the analysis to get a grasp of the impact of the various levers. It is typically complemented by:

- optimization models such as TIMES/PRIMES/GTAP to answer questions such as: “what is the cheapest way of”, “in which order should I perform this”;
- sector specific models on each of the issues addressed to better operationalize the pathway recommendations in the sector; for example, in “buildings”, or “air quality”.

Through EUCalc, we added more sectors, more interlinkages, and a more in-depth modelling of each specific issue.

A combination of all the lever choices creates a scenario. The model outputs for a given input scenario are named *pathways*, because the focus is on the final impact and overall evolution trend. For each pathway, the calculator displays the implications over time (for example in terms of energy, emissions, resource use, job creation and land-use).

For further details on each specific module, please refer to the module specific documentation available [here](#).

Figure 2 gives a global overview of the EUCalc model architecture.

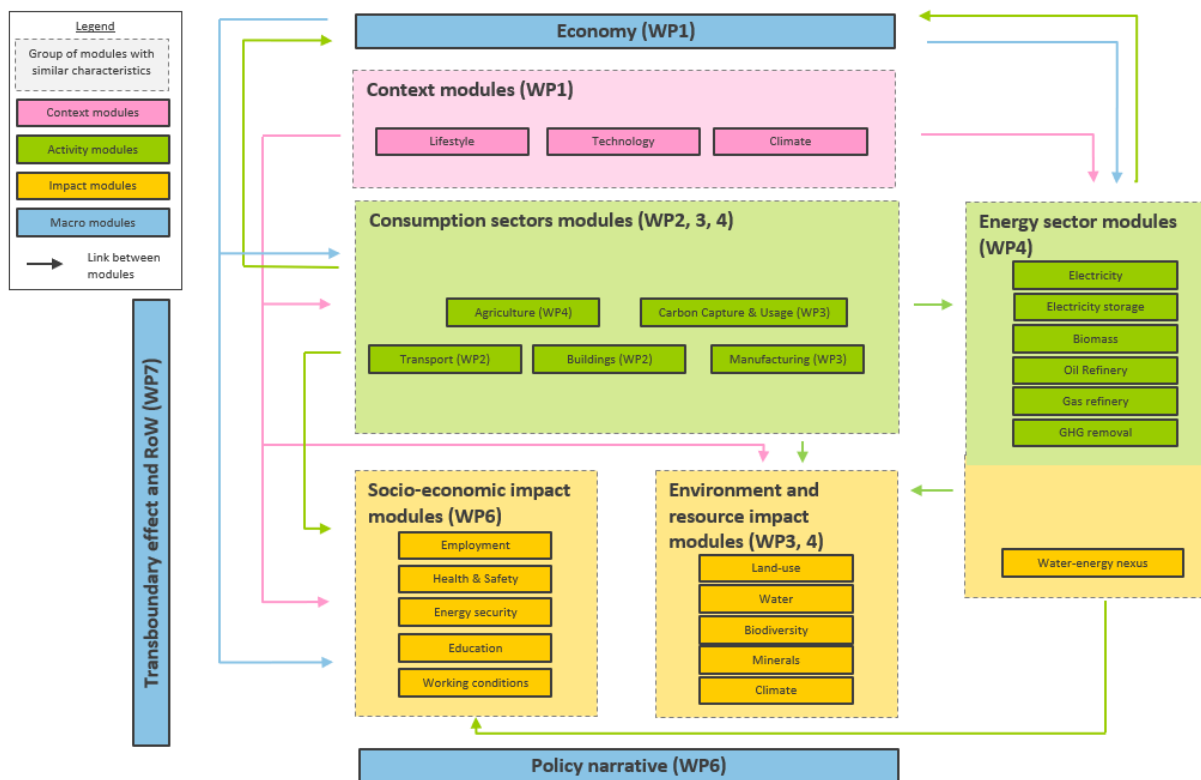


Figure 2. Overview of model architecture.

As described in the following paragraphs, the general model logic is as follows:

The model starts by assessing the energy demand based on three **context** modules (*lifestyle, technology and climate*). They each provide contextual data to the activity and impact modules.

Then twelve **activity** modules cover both consumption sectors (*agriculture, transport, building, manufacturing and CCU*) and energy sectors (*electricity production, electricity storage, biomass production, refinery of oil and gas, and GHG removal technologies*). To perform the activities, these modules use technologies which compute emissions, based on an assessment of the production and consumption of energy, products, materials & resources.

Then, eleven **impact** modules which include social impacts (such as employment, health and safety, energy security, education and working conditions), resource impacts (such as land-use, water, biodiversity, minerals and climate) and water-energy nexus module, then compute various types of impacts based on data from input modules and on results from activity modules.

Finally, three **macro** modules (economy, transboundary effect and policy narrative), use results from activity modules and impact modules to provide results such as fuel prices, GDP, import and export balances, and policy narratives. The way the economic module is linked is not finalised. As context, it can specify the link between GDP and the activity demand drivers (the demand elasticity to price) and computes fuel prices. As activity, it can define the energy costs as function of the supply technology uses. As economic output, it can specify the value added.

Along the activities and impact modules, a transboundary module assesses the imports and exports.

A policy narrative module assesses the links between the policies and scenarios (ideally with each lever position).

## 2. The EUCalc model

### 2.1. Characteristics

The model relates emission reduction with human lifestyles, the exploitation and/or conservation of natural resources, job creation, energy production, agriculture, costs, etc. in one highly integrative approach and tool which enables decision makers to get real-time guidance on possible policy choices underpinned by comprehensive trade-off analyses.

The model strives to achieve the following 6 characteristics:

<b>Answers questions in near-real time</b>	The 'time' is the delay experienced by the user between user input and the moment the results from the model calculations are visible. The Pathways Explorer has to present as a real-time experience to the user, while the model may run in hours/days to pre-compile data.
<b>Covers a wide array of topic</b>	See Figure 2 with modules architecture.
<b>Is granular enough to provide specific answers</b>	The model enables the user to set the levers for Europe plus Switzerland as a block and for each of the 28 countries separately. The model also assesses interactions between Member States and Switzerland and with the rest of the world.
<b>Goes deep enough to provide added value</b>	The model enables to answer key questions regarding the energy transition. The model provides more than sufficient analytical depth to be considered credible by stakeholders. In addition, the model aims to provide as much analysis depth as possible within the technical constraints.
<b>Is collaboratively built</b>	The model is built in parallel by the 10 organisations of the consortium with the inputs of significant numbers of expert stakeholders. The 10 partner organisations are not all required to have high level programming skills.
<b>Is transparent enough to enhance stakeholder buy in</b>	The model is developed through a participative process involving expert stakeholders in each of the sectors addressed. Stakeholders consulted have a good technical understanding of the main assumptions underpinning modelling choices performed in their area of expertise. The model is as transparent as possible to enable stakeholder and end-user buy-in. Transparency can be observed in: <ul style="list-style-type: none"> <li>· The input data used in the model;</li> <li>· The rationale of the calculation applied on the input data;</li> <li>· The ease of use by the end user of the tools used;</li> <li>· The open-source access of the tools</li> </ul>

*Table 4. Characteristics of the model*

### 2.2. Space resolution

The EUCalc model allows the user to select and simulate decarbonization pathways for EU28 + Switzerland. As the envisioned decarbonization pathways impose changes in both demand and supply, levels and structures of production and consumption at sectoral and country levels would also be altered. This, in turn, would change the economic dependencies concerning the EU28 MS + Switzerland at sectoral levels and lead to altered trade patterns. Furthermore, as transboundary flows of goods and services also embody energy consumption and GHG emissions, projecting international trade

impacts is also an important consideration in evaluating the options and tradeoffs of EU decarbonization pathways and their “emission effectiveness” in a global context.

Within the broader scope of the calculator, the transboundary effects are quantified, including intra- and extra-EU trade flows. To do so, a modified version of the GTAP-E model (nicknamed “GTAP-EUCalc”) is developed, simulating perturbations to a projected baseline of the world economy in 2050. Results are shown specifying imports and exports along the commodity dimensions, expressed in both monetary and CO<sub>2</sub>e terms (see section II.4.1). Additionally, the EUCalc model aims at addressing trade linkages with the rest of the world at a greater granularity than one “average” region. In fact, transboundary results for major trade partners (e.g. USA and China) are computed.

### 2.3. Data-driven model structure

To understand the structure of the EUCalc model, one should understand the distinction between the *historical data*, based on various historical databases and the *levers trajectories* based on projection from 2020 onwards until the target year (2050). All runs of the EUCalc model exploit the same *historical data* while each run will have different projections based on the *levels of the levers* the user chose.

Each instance of the model will have exactly the same structure (variables and equations) but different output based on a variation of the projections data.

#### Historical data (OTS – Original Time Series)

Based on various historical databases (see specific sectors for more details), the historical data are values from 1990 to the base year (2015). The historical data are collected for each MS and each year. Those data are present in the calculation process for multiple reasons:

1. The historical data are following the same workflows as the projection. Using official numbers (for GHG emission for example) it is then possible to calibrate the EUCalc results based on the historical figures. High calibration rate meaning high potential of errors in the calculation, this parameter is used to spot issues and to adapt the calculations in order to have correct number for the projection results.
2. In some sectors, the historical data are used to calculate the future time series. The historical data are then used to derive a trend for the position of the levers (see Section 3.5)

One of the challenges of the historical data is to fill the missing values. To fulfill this task, each module is responsible for choosing the most appropriate filling methodology. Standard filling methods are presented in Section 3.4.

#### Projections (FTS – Future Time Series)

Based on stakeholder consultations and scientific literature, the level of the levers are values from 2020 to the target year (2050). The projections are collected for each MS, each year and each position of the lever (1 to 4). There are multiple ways to validate the projections:

1. Stakeholder consultations: asking the expert of the domain and validating numbers during workshops;
2. Using historical data to calculate trends for the future;
3. Scientific literature to validate the future trends and the evolution of the sectors.

You’ll find in each sector documentation how they are building their projections. The main methods to build the projections are presented in Section 3.5.

Additionally, an economic baseline for 2050 is built. Against this base, counterfactual scenarios derived from user-defined EUCalc pathways in the GTAP-EUCalc model are simulated to compute

transboundary effects. The economic baseline, as proposed in D7.1, validated in the expert consultation and detailed in WP7 content document, represents a BAU scenario computed by using macroeconomic drivers, consistent with the EU Reference Scenario for the EU and with SSP2 for ROW.

### Categories of input

The model requires a large amount of input data from a various number of sources to be able to compute its outputs. Those inputs are either historical data, based on various historical databases, or projections and trajectories to 2050.

The model inputs can be classified in different categories:

- Activity: number of households and the insulation rate of buildings, freight transport demand, etc.
- Technology share: share of different heating technologies, share of solar PV, wind or hydro in electricity production capacity, etc.
- Energy consumption/energy efficiency: energy consumption to produce 1kg of meat, energy consumption for the production of a car, etc.
- Fuel shares: share of biofuels and e-fuels in gasoline, gasoil, etc.
- Emission factors: GHG, NOx, particulate matter emissions per unit of fuel used.
- Unit costs: CAPEX and OPEX per unit of activity
- Other socio-economic factors: average number of jobs per unit of activity in a sector, or per euro spend in that sector.
- Other environmental factors: average land-use per unit of activity in a sector.

All those input data allow to compute various outputs which can be observed through the Transition Pathways Explorer:

Dimension	Production	Consumption	Imports	Exports
Energy (per vector)	✓Energy supply	✓Consumption sectors	✓	✓
Emissions (GHG, particular matter)	Local	/	Indirect emissions	/
Products, Materials & Resources	✓	✓	✓	✓
Costs	✓	/	/	/
Employment	✓	/	/	/
Value added	✓	/	Trade Balance	

Table 5. Categories of input

For the detailed list of input and output of each module, please refer to the specific module documentation.

### 2.4. Exploring and reducing the solution space

Lever incompatibilities happen when the choice of levers is not completely MECE. This means that different levers quantities are linked in the real life, but not in the model. This can lead to some incompatibilities between two levers positions.

We identify different types of lever incompatibilities:

- If [Lever 1] is moved to a higher ambition, then [Lever 2] should be moved to a higher ambition;

- If [Lever 1] moved to a higher ambition, then [Lever 2] should be moved to a lower ambition;
- Some positions of [Lever 1] and [Lever 2] are completely incompatible.

Those lever incompatibilities are generating warnings in the Transition Pathways Explorer to let the user know that the levers position is out of the solution space.

## 2.5. Open to expert validation

One of the unique benefits of the European Calculator co-creation process has been how it has helped, in previous iterations, to engage more effectively with stakeholders; not just so that they understand the project but also so they provide expert, real world perspectives on the data/assumptions and scenario setting. Managed sensitively the Calculator can be used as a useful consensus building tool and getting stakeholders involved at an early stage is vital to this.

It is for this reason that the EUCalc embeds a co-creation process with stakeholders who are leading experts in their field, organized through a series of workshops, one for each main module. Early engagement with expert stakeholders also adds value by building a network of supporters for actions within the project related to communication, dissemination, the future exploitation of the EUCalc tool.

The key numbers from the co-creation participation process are:

- Number of Expert Stakeholder Workshops: 10
- Number of experts informed & mapped: > 1000
- Number of expert workshop participants: 171
- Breakdown by target group of workshop participants: 29% academia, 22% private sector, 27% science CSOs and professional associations, 22% European Institutions and others
- Gender balance: 2/3 Male and 1/3 Female

The co-creation process was underpinned by three deliverables:

Del 9.2 defined a set of [standards for stakeholder mapping](#) in order to maximize the value of the co-creation process, in terms of a triple benefit effect;

- i) supporting the EUCalc's scientific soundness (credibility),
- ii) increasing its relevance (related to saliency) and
- iii) ensuring unbiased transparent conduct that considers, among other factors, different perspectives and positions (related to legitimacy).

Del 9.3 provides a summary of discussions and key takeaways from the internal training [workshop on stakeholder engagement](#) as action engagement to compliment D9.4.

D9.4 outlines a consistent [methodology for implementation of EUCalc co-design process](#). Its purpose is to provide information/guidelines to EUCalc partners regarding the process of preparation and implementation of each of the ten expert co-creation workshops and a Public Call for Evidence, organized during the development phase of the European Calculator - to elicit expert feedback in the different sectors and modules included in the Calculator.

All expert stakeholder interactions taking place within the EUCalc project have been documented published in the format of sector related [expert workshop deliverables](#).

## 3. Feedback loops, country disaggregation and interaction

### 3.1. Feedback loops

In real life, the interactions between the sectors are tightly intertwined and lead to feedback loops. These interactions between sectors are required but generate complexity. This complexity brings an

increase of the calculation time and thus must be reduced to allow real time computation. There is 2 ways to remove a feedback loop:

1. the first option is to create a lever dedicated to it, allowing the user to decide the ambition for a specific branch of the loop;
2. the second is to reduce the number of interactions between sectors, cutting the 'weak' links between the sectors.

We illustrate in the first section several examples enabling to reduce complexity by reducing the number of interactions between sectors or by adding new levers. The second section is dedicated to the list of feedback loops that were removed from the model and the scientific justification for it.

### 3.1.1. How to spot and remove feedback loops?

Let us look at some effects of a population increase in Box 1:

1. More people eating leads to more chicken consumption;
2. More chicken livestock breeding leads to more cereals consumption; (the cap on cereals availability is influenced by a warming climate)
3. More cereals production leads to more fertilizers consumption; (the yield of cereals production is influenced by a warming climate)
4. More fertilizers production leads to more ammonia consumption;
5. More ammonia production leads to more electricity demand;
6. More electricity demand leads to more wind turbines demand;
7. More wind turbines demand leads to more steel demand;
8. More steel demand leads to more electricity demand;

In the example above, we arrive at a situation with a feedback loop from 8 to 6. These feedback loops are typically rerun many times and significantly contribute to the calculation time. For this reason, we will avoid them whenever possible.

Box 1. Example of population increase loop

The example above can be illustrated as in Figure 3.

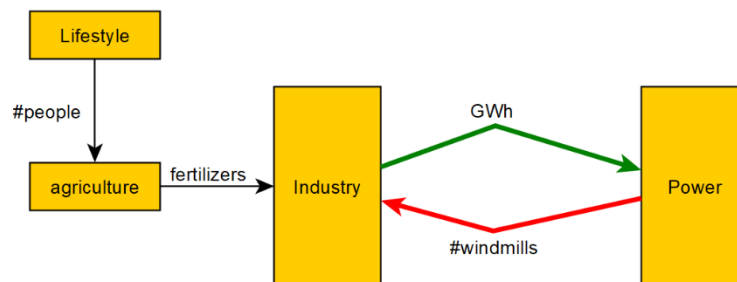


Figure 3. Identification of a loop between Industry and Power.

After identifying the loop, we need to consider multiple options as illustrated in Figure 4, Figure 5 and Figure 6 below:

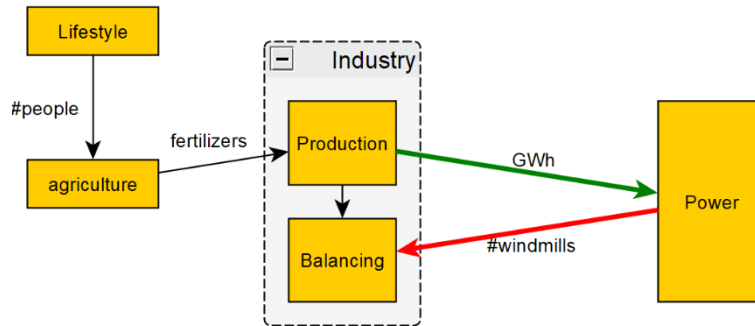


Figure 4. The power to build the wind turbines is negligible, the weak link is the red link.

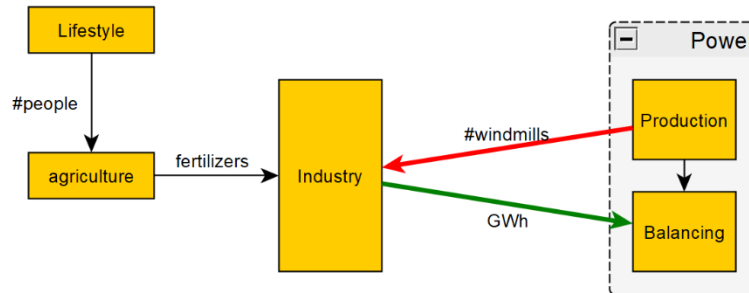


Figure 5. The GWh from industry is negligible and could be handled by a balancing strategy. The weak link is the green link.

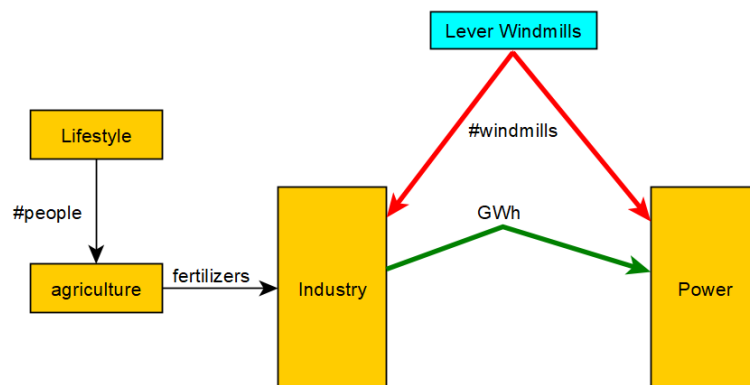


Figure 6. The lever 'Windmills' is providing the number of windmills needed to Power and Industry.

The sector leaders of the sectors are then responsible for choosing the better solution between the multiple options. Let us look at another example (and loop) with the effects of an increase steel demand in Box 2.



1. More steel demand leads to more freight demand;
2. More freight demand leads to more demand for ships (and ships building materials);
3. Building & using more ships requires more electricity;
4. Supplying more electricity requires the deployment of new power plants;
5. Building power plants requires steel;

In the example above, several links do not make a material difference because they do not explain a significant portion of the influenced variable. At a global level, ships represent only 2% of steel and power plants represent only a minor contribution to steel demand, which means the necessity of keeping links 3 & 5 should be challenged.

Furthermore, like in the previous example, we can also assess modelling freight demand through a lever, which breaks the link 1.

*Box 2. Example of increasing steel demand*

The examples here above are feedback loops between 2 sectors, but we must also consider the feedback loops that may occur between a larger number of sectors. The easiest example here is the climate. The temperature influences multiple sectors that are emitting GHG emissions which influence the climate and the temperatures. Regarding those 'main' feedback loops, it is the responsibility of the consortium to decide on how to deal with this kind of loops. See the next section for the decisions that were taken regarding the loops.

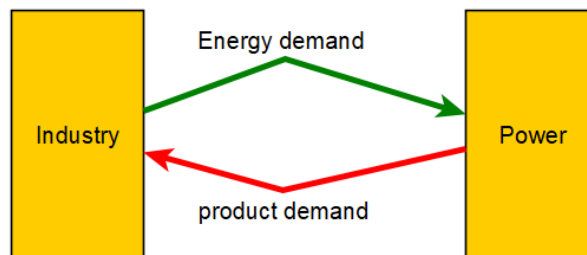
### *3.1.2. Feedback loops list*

#### Feedback loops between sectors

In this section, we present multiple loops that were broken in the model. All the red arrows are the links identified as the weaker links by the sectoral modelers.

#### *Industry - Power*

This feedback loop is represented in the Figure below: Industry is providing his energy demand to Power while Power is telling Industry the quantity of products (windmills, pv, thermal, marine...) needed to fulfill the demand.



*Figure 7. Feedback loop Industry - Power.*

To avoid this conflict, a lever was created. This lever will determine the quantity of plants available to the power module and to the industrial sector.

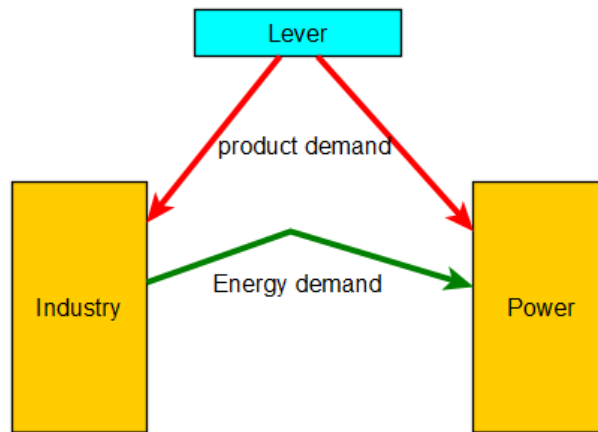


Figure 8. Lever determining quantity of plants available to the power module and to the industrial sector.

### Agriculture - Industry

In this feedback loop, the issue is coming from the fertilizers demand (see Figure below).

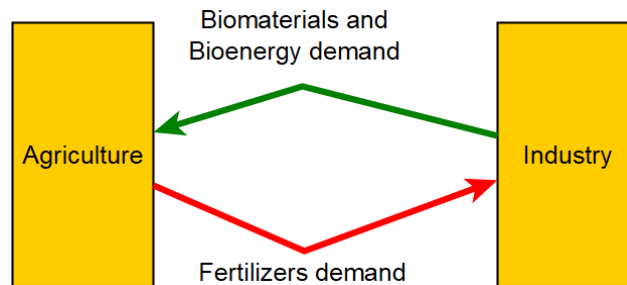


Figure 9. Feedback loop Agriculture - Industry.

To avoid this loop, the weak link was found to be the biomaterials and bioenergy demand coming from the ammonia sector. The solution proposed by the consortium is the one presented in Figure below.

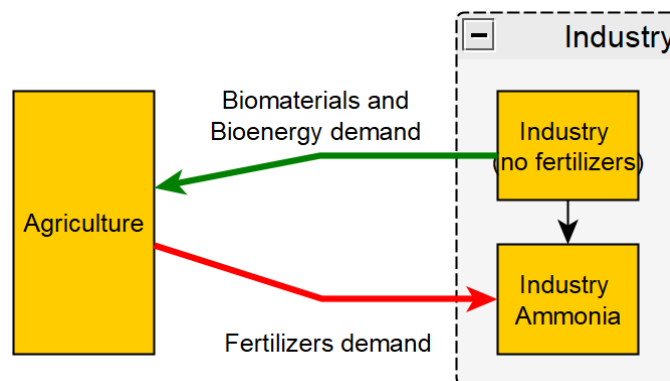


Figure 10. Lever determining biomaterials and bioenergy demand coming from the ammonia sector.

### Agriculture - Power

To avoid a possible loop between agriculture and power regarding the demand of biomass, the solution selected was to keep the biomass production/demand inside the agriculture module. The result is that the energy demand to the power sector include the electricity from biomass technologies.

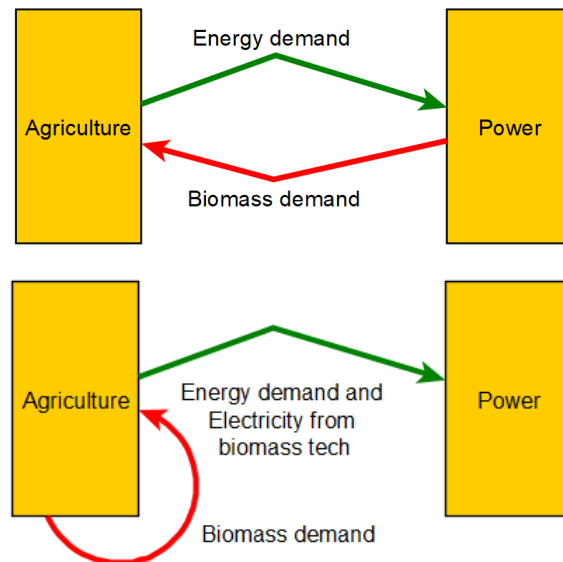


Figure 11. Biomass production/demand inside the agriculture module.

### District Heating - Agriculture- Industry

In the following case, the output of district heating to agriculture and Industry have been ignored to avoid a loop between multiple sectors.

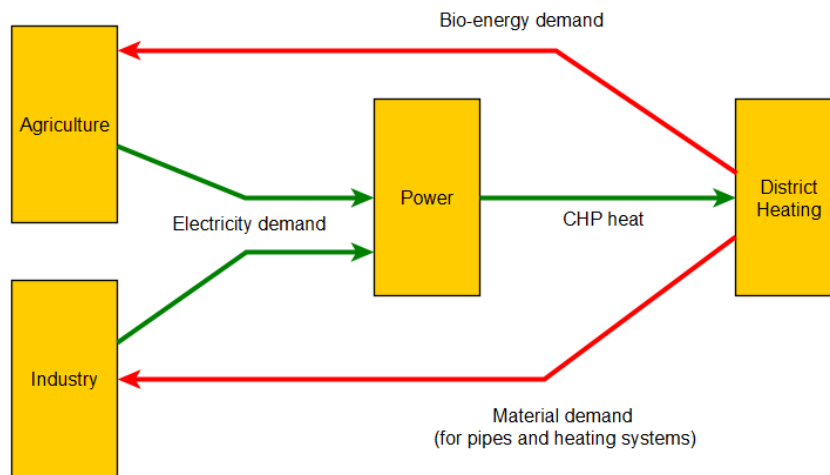


Figure 12. Output of district heating to agriculture and Industry.

### Feedback loops with the overall model

#### Climate Loop

The Climate loop is one of the major feedback loops running all over the module. To avoid this loop, the climate was converted into a lever: the user chooses the target temperature he wants to achieve with his simulation and the web application show him the carbon budget allowed to stay under this target.

### 3.2. Disaggregation by Country

Based on the EU-wide levels of ambition, we use the Science-based target concepts [Science Based Target, 2015] to disaggregate ambition levels at the country level.

The Science-based target uses two concepts to describe the targets evolution: the convergence and the compression concepts (see Figure 13 and Table 6).

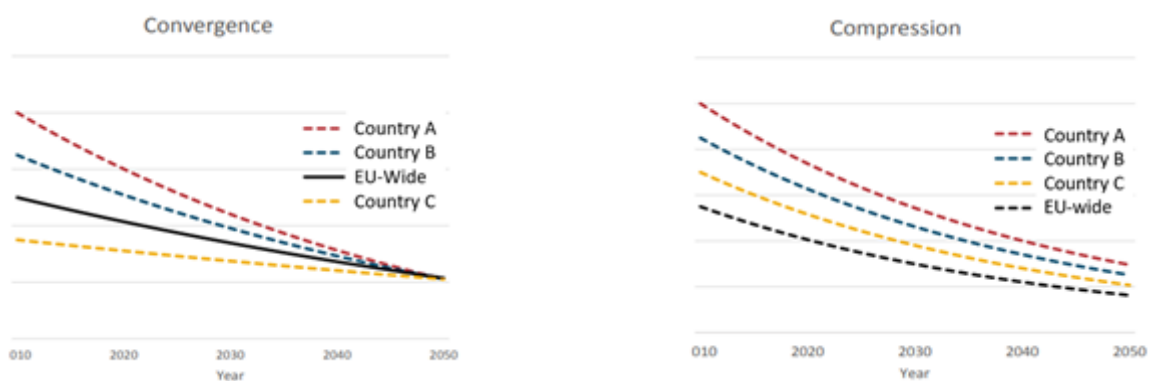


Figure 13. Concepts to describe the targets evolution: convergence and compression.

Convergence	Compression
The absolute 2050 ambition is the same for all countries (e.g. x kwh electricity/km for small electric vehicles) in 2050.	The relative 2050 ambition is the same for all countries (e.g. -30% passenger.km/year by 2050 vs 2015 in each country)
This results in some countries having to do greater efforts than others, depending on their 2015 situation.	This results in all countries having to do the same relative efforts based on their 2015 situation

Table 6. Convergence vs Compression

As described in Table 6, the convergence is better suited when country-specific parameters have little to no influence on the long-term evolution of the lever value. This is usually accepted for technological levers such as energy efficiency of a given technology for example. The compression is, on his side, better suited when local or country-specific parameters have an important influence on the long-term evolution of the lever value. This could be the case for transport demand, for example, for which urbanization rate, population density or local topography have an influence

We mostly use a hybrid calculation based on a weighted average of convergence and compression results.

Generally, and if no literature source is found to justify another logic, we will use the following logic:

- High convergence for parameters that are highly influenced by economic factors (e.g. car ownership and car modal share): indeed, economic convergence between countries is not only an ethical objective of EU [European Parliament, 2018], but is also practically observed [Butkus et al., 2018].

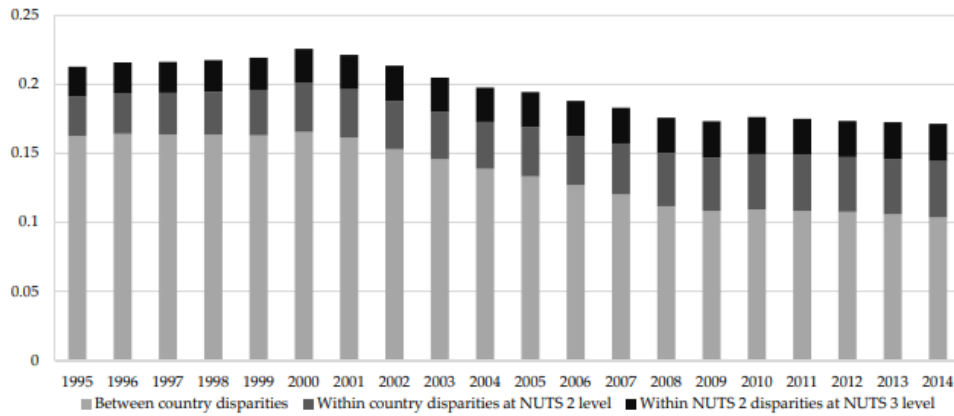


Figure 14. Evolution of EU disparities between 1995 and 2014 [Butkus et al., 2018].

- High compression for parameters that are mostly influenced by local parameters such as geographical parameters, regional territory planning, etc. (e.g. public transport networks).

The weights of the hybrid calculation that are used are specified for each lever in the following sections.

### 3.3. Historical data completion

One of the challenges of the historical data is to fill the missing values. When collecting the data from official data sources, the results are rarely coming without missing values. The Figure 15 is an example of what is common to collect and how to fill it to avoid gaps while Figure 16 is a graphical representation of one filling method.

Metric ID	Value	1990	2005	...	2014	2015	2016	...	2050
tra_passenger_distance			Y	Y	Y				
tra_freight_demand		X	X	X	X	X			
Other Metric		Z		Z	Z	Z			

↓ Base year

Metric ID	Value	1990	2005	...	2014	2015	2016	...	2050
tra_passenger_distance		Ye	Y	Y	Y	Ye			
tra_freight_demand		X	X	X	X	X			
Other Metric		Z	Zi	Z	Z	Z			

Figure 15. Historical data with missing values filled to ensure data is present for all years up to base year - 'e' meaning extrapolation - 'i' meaning interpolation.

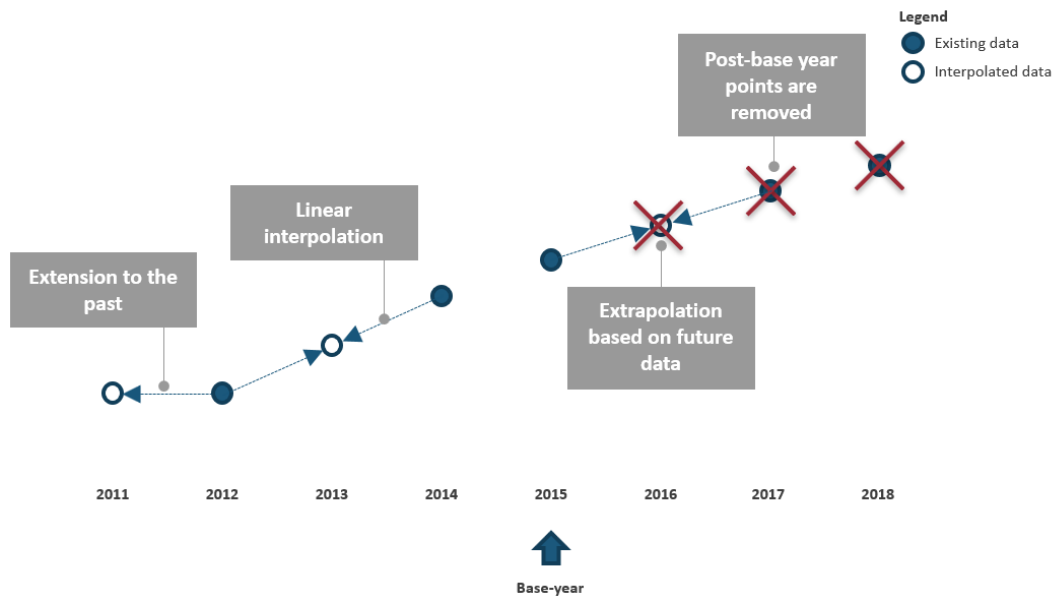


Figure 16. Linear interpolation for completion of missing values in historical dataset.

To fulfill this task, each module is responsible of the filling methodology used his model. They are 4 standard filling methods implemented in the model:

1. Linear interpolation: this method is the basic method filling the missing 'inter-year' value by linear interpolation and the past/future missing values by extension;
2. Linear extrapolation: this method is the basic linear extrapolation of all the missing values on a single straight line;
3. Exponential weighted moving average<sup>4</sup>: In this function missing values get replaced by moving average values;
4. Similarities with other countries: this method use similarities with other MS to fill the missing values.

<sup>4</sup> see the R function `na.ma` using 'exponential' weighting for further details here: <https://www.rdocumentation.org/packages/imputeTS/versions/2.7/topics/na.ma>

## Missing Value Completion

Linear extrapolation vs. Linear interpolation vs. Weighted Moving Average

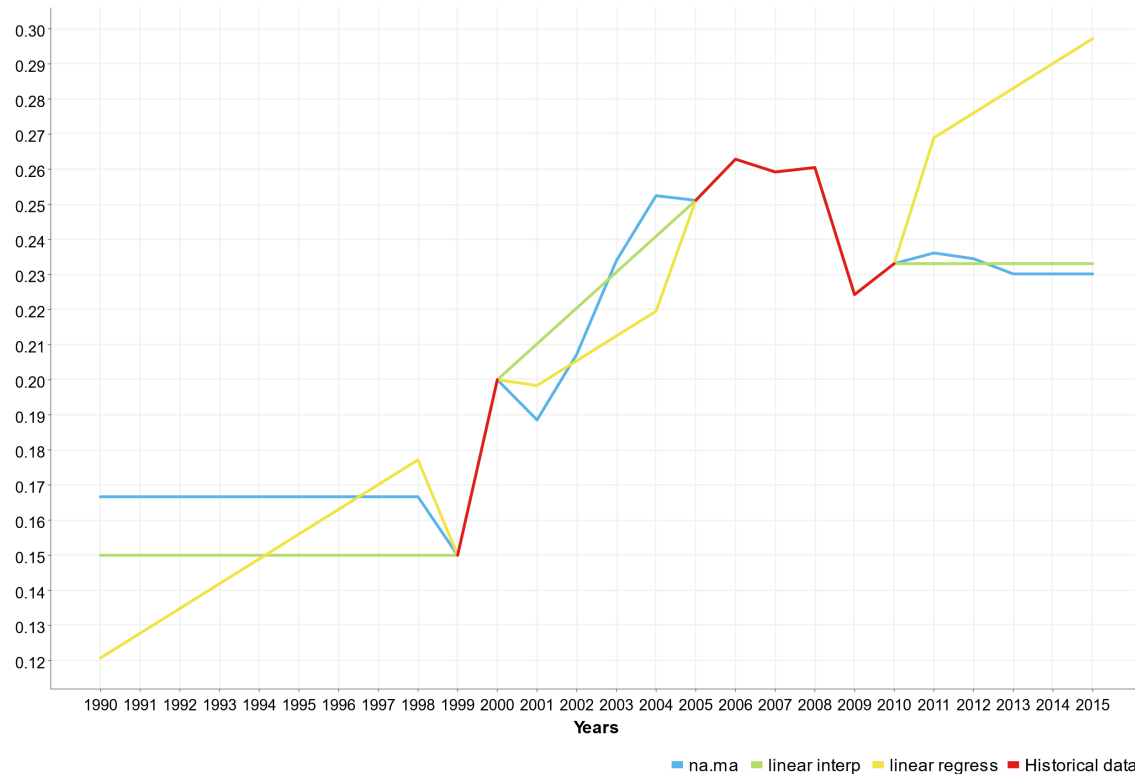


Figure 17. Completion of missing values (red) using linear interpolation (green), linear regression (yellow) and Weighted moving average (blue).

### 3.4. Ambition curves shapes

As specified in section II.2.4, the model is driven by historical data and projections. The value of the levers for each year between the base-year end the target year, must be computed and validated. The validation process has already been mentioned before. This section is dedicated to the way the sectors are building their projections using ambition curves shapes. Here are the two main methods used by the different sectors:

- Build of projection using historical data: the historical data are used to calculate trends to the future. This is specific to each sector. It won't be developed in this section.
- Configuration file with multiple parameters: multiple parameters are used to build the projections of a lever. Those parameters (starting time, duration, final ambition and shape of uptake) are presented in the section hereunder.

#### 3.4.1. Curve shapes

It has been observed that uptake of new technologies is usually not linear but has a "s-shaped" trajectory (see Figure 18): the new technology starts slowly by reaching the innovators and early adopters who are a minority, then it accelerates and reaches the majority and finally it re-decelerate and reaches the laggards [Roger, 1995], [Felton, 2008].

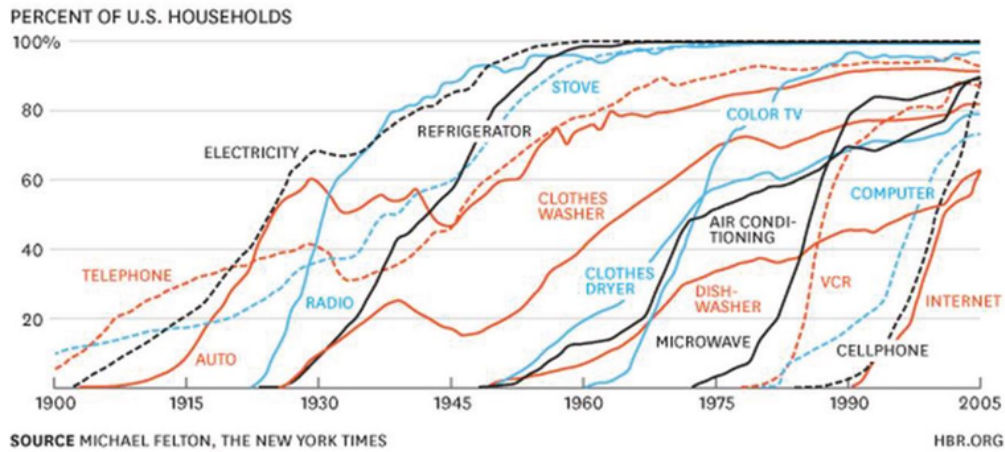


Figure 18. Technology adoption rate in U.S. [Source]].

For this reason, we have decided to implement different types of ambition levels curve shapes, in order to be as realistic as possible and not limit trajectories to simple linear curves.

The ambition levels trajectories between the base year (2015) and 2050 will depend on different parameters:

- Starting time: when will the new trend or new technology start to spread?
- Duration: how long will it take to reach its maximum potential?
- Final ambition: what is the maximum potential we expect?
- Shape of uptake: will it evolve smoothly, or is it most likely to start slowly and accelerate after this starting phase?

In order to reflect the variety of situations, the different levers can take different types of shapes (see Figure 19): linear evolution (L-curve), S-shaped evolution (s-curve), or half-S-shaped curve (HS-curve) in case the trend is already considered to be in its acceleration phase. If none of those curves are adapted, each lever can also be implemented with a custom curve.

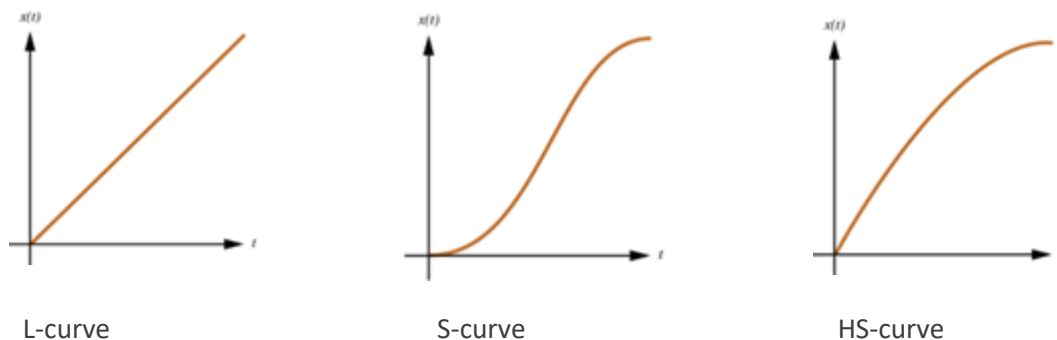


Figure 19. Ambition levers curve shapes.

S-curve will usually be used for the diffusion of new technologies, and other types of curves will be used when necessary, based on expert judgment.

The Figures 20 and 21 are two examples of generated projection based on parameters.



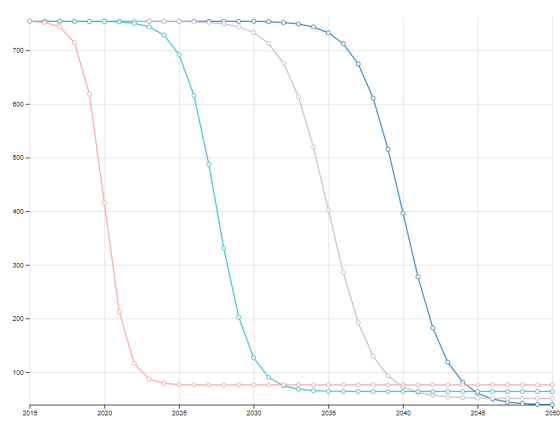


Figure 20. s-curve projection between 2015 and 2050 for 4 levels scenarios.

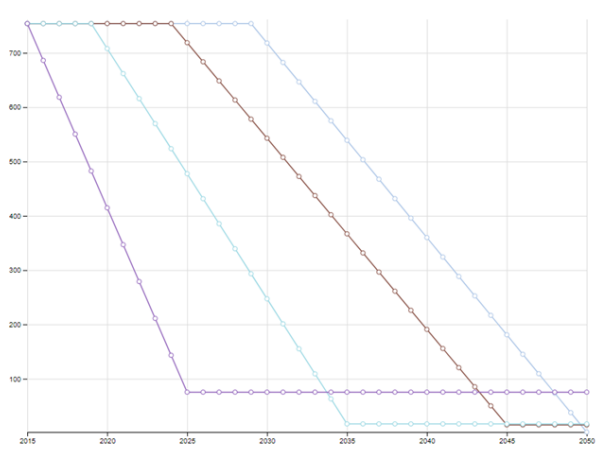


Figure 21. Linear projection between 2015 and 2050 for 4 levels scenarios.

## 4. Interaction between modules

Figure 22 illustrates the main interactions between modules.

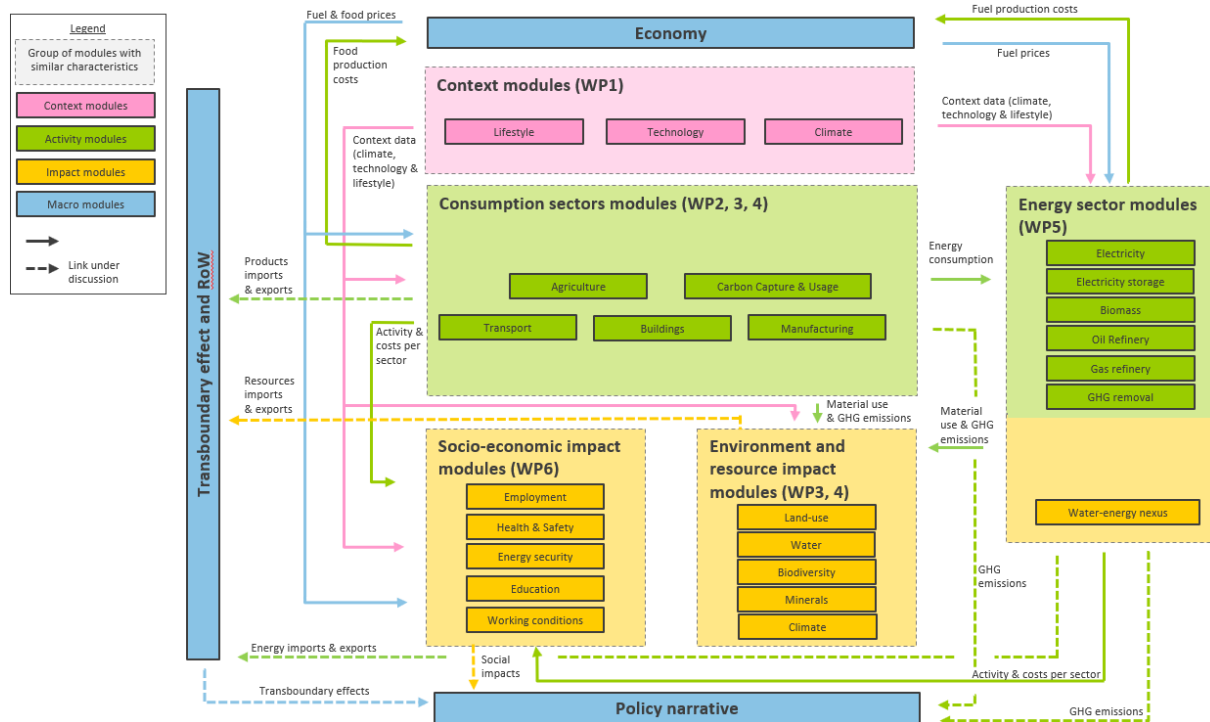


Figure 22. Interactions between EU-Calc modules

The **context modules** give contextual data about climate, lifestyle and technology to the consumption sectors modules, the energy sector modules and the social and resource impact modules.

The **consumption sector modules** receive data from the context modules and from the economy module. Then it provides energy consumption to the energy sector module, material use and GHG emissions to the resource impact modules, activity and costs to the socio-economic modules, product quantities and technological/behavioral changes to the transboundary module, and GHG emissions to the policy narrative module.

The **energy sector modules** receive context data from the context modules and the energy consumption from the consumption sectors modules. It provides the fuel production costs to the economy module, which in turn provides fuel prices. Just as the consumption sectors, it provides material use and GHG emissions to the environment and resource impact modules, activity and costs to the socio-economic impact modules, energy quantities and technological/behavioral changes to the transboundary module and GHG emissions to the policy narrative module.

The **socio-economic impact modules** receive contextual data from context modules, fuel and food prices from the economic module and activity and costs from the activity modules. It produces various socio-economic impact indicators as output which are used by the policy narrative module. The **environment and resource impact modules** receive contextual data from context modules and material use and GHG emissions from the activity sectors. It produces various environmental impact indicators and gives resource production data to the transboundary module.

The **economy module** computes fuel and food prices based on fuel production costs and food production costs provided by the related activity module. It then gives the prices to the consumption and energy sectors and to the social impact module.

The **policy narrative module** receives social and environmental impact data to produce its outputs. Finally, the **transboundary effect module** receives the consumption/production data, together with technological/behavioral changes, from activity sectors and from resource sectors.

For further details on each sector interaction, we reference the reader to the sector specific documentation.

#### 4.1. Interaction with other models (GTAP-EUCalc)

The GTAP-EUCalc general equilibrium approach differs from the EUCalc modularized approach, in which the lever setting reflects a range of ambition levels expressed by the end-user. The combination of the two adds value to the EUCalc with respect to the current existing calculators.

Pathways defined in the EUCalc model are essentially driven by a set of assumptions (levers) on lifestyle choices on food consumption, transportation, buildings, materials and manufacturing, and direct and indirect energy demand. These lifestyle choices drive the supply side in the individual modules to generate a particular set of emission outcomes. The baseline projections from WP7 provide a set of trade results to be included in the base case calibrations of traded products in the individual modules. The EUCalc model then provides users with a set of levers on both the demand and supply side to create their own decarbonization pathways. The lever settings on the demand side (i.e. lifestyle choices) are essentially considered changes to the demand curves relative to the baseline in the GTAP-EUCalc model, whereas changes on the supply side imposed in the individual EUCalc model are modeled in the GTAP-EUCalc as changes in the supply function either through changes in the cost structure of producers and/or through total or biased productivity progresses. Consistent with the overall set up of WP7 and the generally accepted economic theory, the simultaneous shifts of the demand and supply functions arising from implementing the EUCalc pathways then endogenously determine the excess demand (i.e. imports) or excess supply (exports) of each and every product for each and every country/region included in the model. In addition, the specific modeling structure featuring the well-known Armington specification allows for generating bilateral trade flows between any pair of trading countries.

To be able to implement the interface described above, the GTAP-E model has been modified to be able to accommodate the sectoral coverage of the various modules. The new so-called GTAP-EUCalc model differs from its predecessor by incorporating an isoelastic aggregate land supply function, a new Cobb-Douglas private demand system (as it is not easy to re-parameterize the constant difference of elasticities demand system used for projecting the 2050 economic baseline for purposes of simulating the very large demand structural changes implied by many of assumed lifestyle levers in WP1) with an embedded twist to target variations in consumption shares, a twist parameter in each nest of the constant elasticity of substitution firms' structure, and an additional set of equations to include non-CO2 emissions in the model and to measure changes in overall GHG emissions.

# III. QUALITY MANAGEMENT

Several processes are set up to ensure the quality of the model and of project deliverables and to facilitate stakeholder buy-in.

Each deliverable is reviewed internally by at least one institution other than the authors' prior external publication. In addition, two calls for evidence are performed to increase the scientific robustness of the results.

A pre-Call for Evidence was launched in January 2019. This is a review of most of the outputs which will be published by the project (i.e. the model, the content documents specifying the assumptions). This review is mostly internal to the consortium.

A public Call for Evidence is now ongoing (September 2019). This is a review of the whole model and of the presentation of its outputs. This review is internal and external (public) and targets all the stakeholders which have already been identified during the co-creation process.

The model quality management consist in various quality checks:

1. Input data quality;
2. Results robustness and calibration on historical data;
3. Modelization choices and uncertainties.

## 1. Input data quality

The European calculator project makes use of several heterogeneous data sources as basis for its calculations. It is therefore of high importance to ensure the quality of input data, which are the foundation of the model.

To ensure good quality of input data, we have set up a quality management plan which consists of a series of quality control checks:

- **Reliability control:** for reliability control, we evaluate three things:
  - Data source reliability: this is an informed expert judgment based on experience. Usually we favor official and open-datasets for transparency.
  - Data rating: we use a data rating scale from A to E based on and IPCC methodology [IPCC, 2001] (see Figure 23)

Rating	Definition	typical error range
A	an estimate based on a large number of measurements made at a large number of facilities that fully represent the sector	10 to 30 %
B	an estimate based on a large number of measurements made at a large number of facilities that represent a large part of the sector	20 to 60 %
C	an estimate based on a number of measurements made at a small number of representative facilities, or an engineering judgement based on a number of relevant facts	50 to 150 %
D	an estimate based on single measurements, or an engineering calculation derived from a number of relevant	100 to 300 %
E	an estimate based on an engineering calculation derived from assumptions only	order of magnitude

Figure 23. Data rating proposition.

- Clarity of the data scope: this consists in verifying that the scope of the data is clearly defined. It is a very important step as we have to make sure that, when using datasets from different sources, all input data are coherent.
- **Completeness control**: for this test, we identify the gaps in the dataset in terms of countries, period, and technology it covers.
- **Consistency control**: for this test, we identify outliers in the dataset, using the interquartile method (for example, see [Rousseeuw, 1993]).
- **Accuracy control**: for the accuracy control, we use two methods:
  - Cross-checking with other sources.
  - Review by experts.
- **Timeliness control**: this check consists in verifying the time when the dataset was created and last updated.

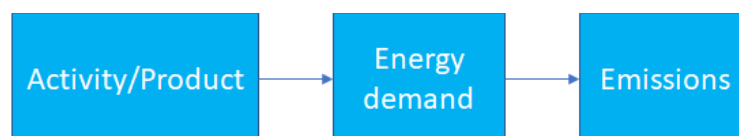
The results of those data quality controls are detailed in the specific modules documentation.

In addition to those quality controls, the database is made available during the calls for evidence and is presented during the sectoral workshops.

## 2. Calibration on historical data

Calibration is the process by which we check that the model is able to reproduce historical activity/product demand, energy consumption and GHG emissions. The goal of calibration is to highlight model weaknesses or/and data errors. The model/data are then adapted until a satisfactory calibration factor is obtained throughout the different model dimensions (activity/product, energy and GHG emissions). A 100% calibration factor means that the model results are fully in line with the available reference data.

The general model logic is the following (Figure 24). We first consider the demand for products or activity. Then, depending on the energy intensity of this activity or product, we compute the total energy demand it implies. Finally, the GHG emissions are computed from the total energy demand considering the various energy vectors and their respective emission factors.



*Figure 24. Activity/product consumption determines energy demand which in turn determines GHG emissions.*

Calibration thus follows the same order: activity/product variables are calibrated first, then energy demand and finally emissions.

The first step is to compare the value of uncalibrated variables computed for years prior to the base year (i.e. 2015) to reference values for these years. This allows to determine the calibration rate, i.e. the corrective factor by which we should multiply the results of the model to stick to historical reference data (Figure 25).

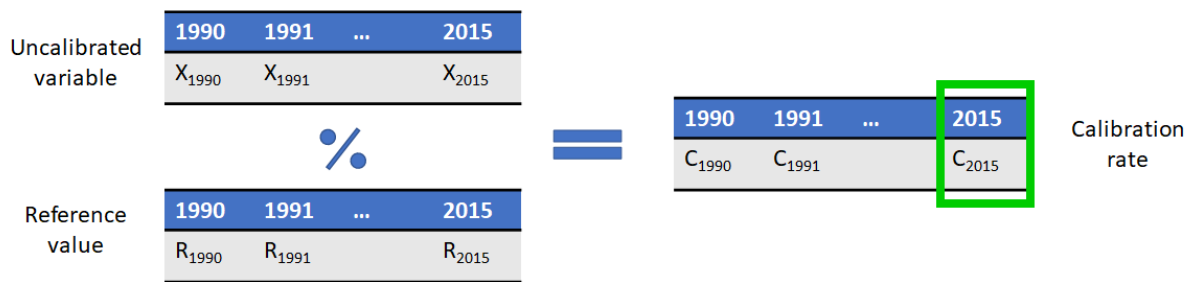


Figure 25. Computation of calibration rate.

The second step is to apply the calibration rate to the model projections, i.e. the results for subsequent years, 2015 to 2050. The rationale is to apply the calibration rate determined for the base-year to all subsequent years (Figure 26).

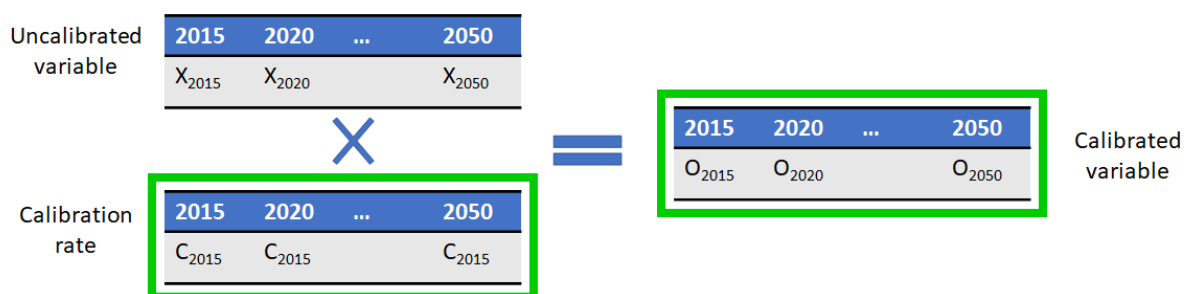


Figure 26. Calibration of model projections.

However, calibration does not ensure that results for the projected years until 2050, are correct. Indeed, the scenarios that are explored by the model include many new social patterns and technology developments that are not captured in the historical data used for the calibration. The calibration process does therefore not guarantee that such new trends are correctly modelled.

On the model, several dashboards highlight the consistency of key model results amongst themselves and their alignment with official external sources such as the energy balances and emissions inventory based on the calibration results.

Other robustness processes, such as quality checks between modules, are also incorporated to ensure mutual understanding of data flows between modules and towards the Pathways Explorer (both in terms of which variable types are transferred, and their expected value range).

For further module-specific information and computation details, the reader should refer to module documentation.

### 3. Modelization choice & uncertainties

In addition to input data quality and results calibration, the modelization choices also have to be subjected to quality management. The goal of this quality management is to make sure:

- we integrate the most important trends and future decarbonization solutions to the model and don't miss major evolutions;
- the new trends are integrated and modelled correctly in the model.

In order to reach those objectives, we proceed in different steps:

- **Analysis of the current situation:** the first step is to analyse the current situation in each sector. This analysis allows to identify the sub-sectors that are the most important in terms of energy consumption and GHG emissions, and the global trends of those subsectors (are their activity increasing or decreasing? Is there a subsector which is growing faster than the others?). This information is very useful to decide which trends should be modeled in priority (we will first concentrate on the trends concerning major sub-sectors).
- **Identification of major trends for the future:** to identify the major trends and expected evolutions in a sector, we base our analysis on an extensive literature review, on other sectoral models' comparison, and on exchanges with experts (through the workshops and through bilateral discussions).
- **Prioritization of trends to be included in the model:** to decide which trends should include in the model in priority, we ask ourselves 3 questions:
  - Which sub-sectors will this trend impact? If the impacted sub-sectors are important in terms of energy and emissions, it will be of higher priority than if the trend only impacts small sub-sectors.
  - What is the development potential of this trend? If the trend is expected to stay very marginal and does not have a large development potential, its will be considered as a lower priority.
  - Does this trend have a clear impact on activity, energy consumption and GHG emissions of the sector? If the identified trend does not have a clear impact on the activity level, energy consumption or GHG emissions, it is not included in the model. A good example of such a case is the car automation case: it can either drastically reduce passenger activity or largely increase it, depending on the context in which it will develop.
- **Modelization of those trends:** once we have identified the trends that will be included in the model, we still need to ensure that they are modelled correctly. Just as for the identification of major trends, we base our analysis on an extensive literature review, on other sectoral models' comparison, and on exchanges with experts. A good way to assess if our modelization is coherent, for example, is to reproduce a scenario from another study in our model, and to compare the results and analyse the differences.

# IV. APPENDIX A: KNIME MODELING

To access the KNIME model, please refer to the user guide for the knime model document which is available [here](#).

## 1. Knime environment

[KNIME Analytics Platform](#) is an open source software for creating data science applications and services. Intuitive, open, and continuously integrating new developments, KNIME makes understanding data and designing data science workflows and reusable components accessible to everyone.

In [KNIME Analytics Platform](#), individual tasks are represented by nodes. Each node is displayed as a colored box with input and output ports, as well as a status, as shown in Figure 27. The input(s) are the data that the node processes, and the output(s) are the resulting datasets. Each node has specific settings, which we can adjust in a configuration dialog. When we do, the node status changes, shown by a traffic light below each node. Nodes can perform all sorts of tasks, including reading/writing files, transforming data, training models, creating visualizations.

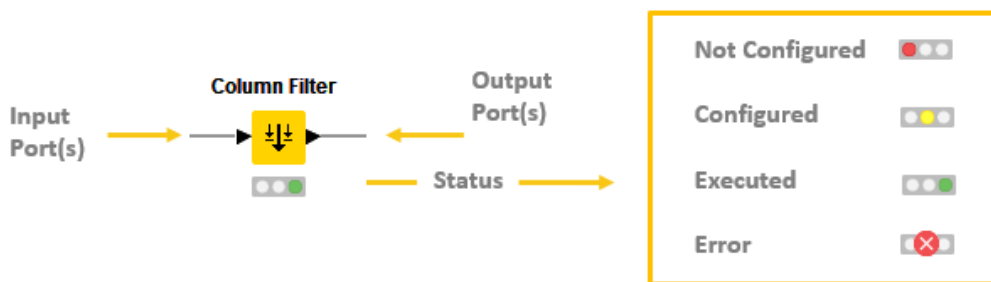


Figure 27. Node ports and model status.

A collection of interconnected nodes constitutes a workflow, and usually represents some part - or perhaps all - of a particular data analysis project.

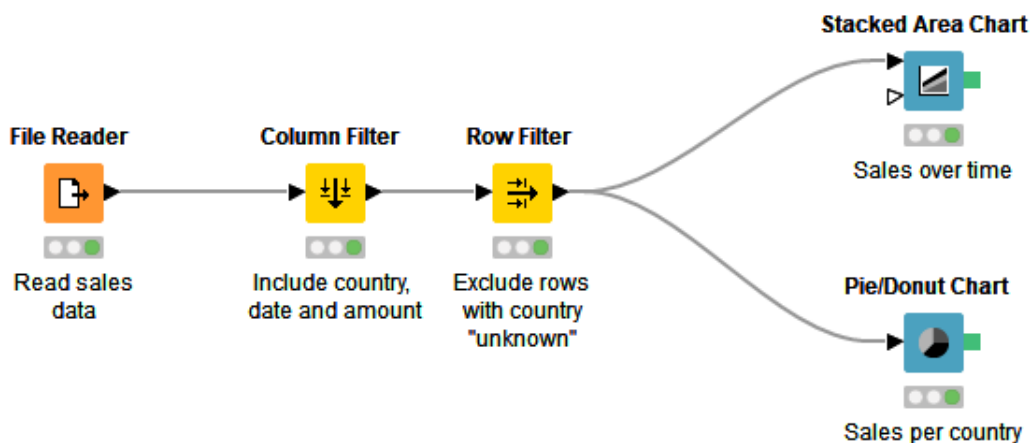


Figure 28. Example workflow.

The example workflow in Figure 28 reads data from a CSV file, filters a subset of the columns, filters out some rows, and visualizes the data in two graphs: a stacked area chart and a pie chart, which you can see in Figure 29.



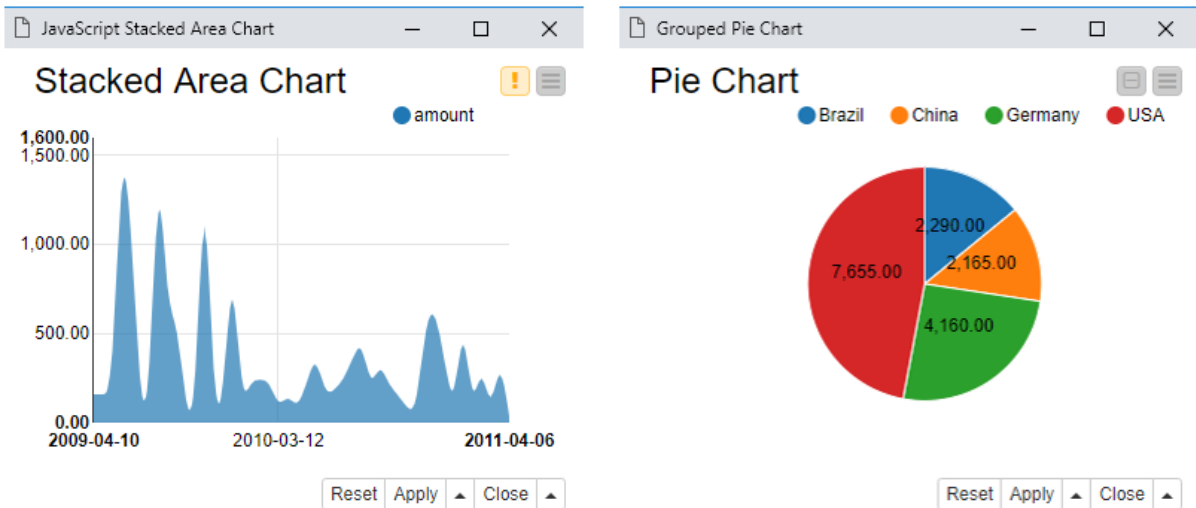


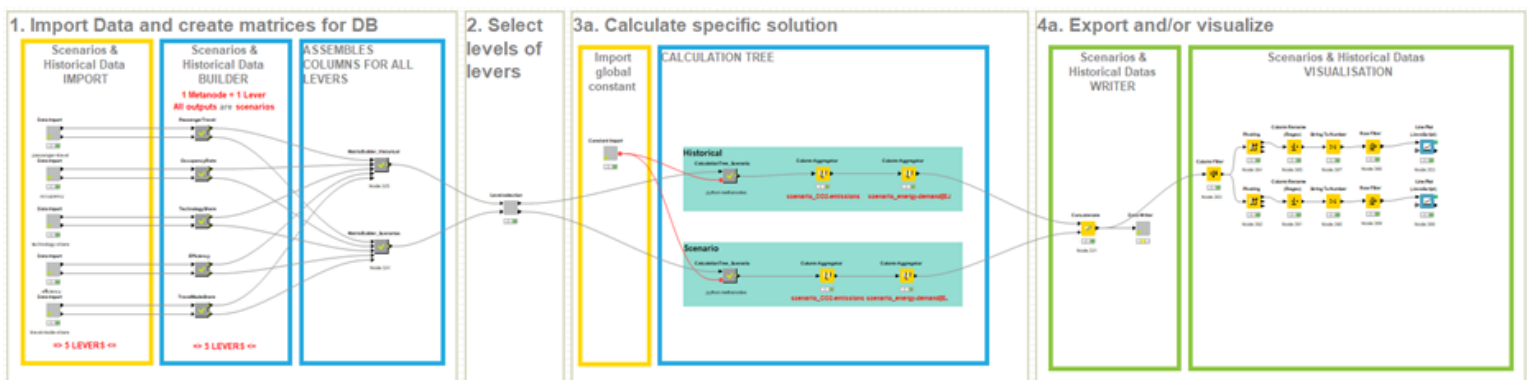
Figure 29. Output views of the example workflow.

## 2. Accessing the model

To get you a copy of the project up and running on your local machine for development and testing purposes, please go to the [Git repository](#) of the EUCalc project and follow the instruction of the [README](#).

## 3. Definition of a KNIME workflow

The workflows in KNIME are built to be readable from left to right with a color coding for functional meaning:



- **Yellow:** Input and cleaning
- **Green:** output/reporting/formatting
- **Red:** error/warning processing
- **Blue:** methodological processing steps that are not input, output or errors.

Figure 30. KNIME workflows.

Moreover, those colored boxed are described using clear logical steps. Note that this 'in-model' documentation only answers the question of 'what the workflow does?'. For further details of the calculation logic, please refer to the WP detail documentation.

You will also find calculation logic inside the modules metanodes in the following form:

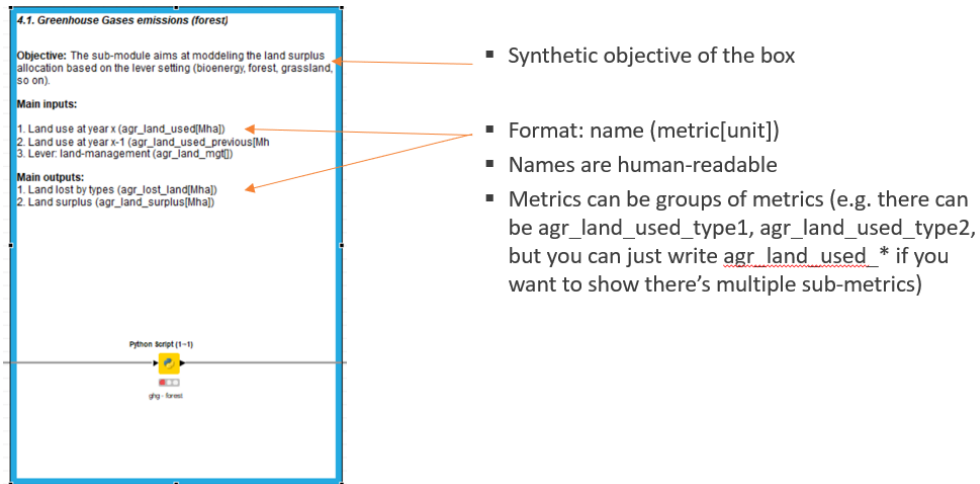


Figure 31. Calculation logic inside modules metanodes.

## V. APPENDIX B: LIST OF LEVERS

Topic	Lever group	Individual lever
Key behaviours	Travel	Passenger distance
		Mode of transport
		Occupancy
		Car own or hire
	Homes	Living space per person
		Percentage of cooled living space
		Space cooling & heating
		Appliances own
	Diet	Appliance use
		Calories consumed
	Consumption	Type of diet
		Use of paper and packaging
		Product substitution rate
Food waste at consumption level		
Technology and fuels	Transport	Freight distance
		Passenger efficiency
		Passenger technology
		Freight efficiency
		Freight technology share
		Freight mode
		Freight utilization rate
	Buildings	Fuel mix
		Building envelope
		District heating share
		Technology and fuel share
		Heating and cooling efficiency
	Manufacturing	Appliances efficiency
		Material efficiency
		Material switch
		Technology efficiency
		Energy efficiency
		Fuel mix
		Carbon Capture in manufacturing
	Power	Carbon Capture to fuel
Coal phase out		
Carbon Capture ratio in power		
Nuclear		
Wind		
Solar		

		Hydro, geo & tidal
		Balancing strategies
		Charging profiles
Resources and land use	Land and food	Climate smart crop production
		Climate smart livestock
		Bioenergy capacity
		Alternative protein source
		Forestry practices
		Land management
		Hierarchy for biomass end-uses
		Water and biodiversity
		Land prioritisation
Boundary conditions	Demographics & long-term	Population
		Urban population
		EU emissions after 2050
	Domestic production	Domestic food production
		Domestic product output
		Domestic material production
	Constraints	Global mitigation effort
		Discount factor

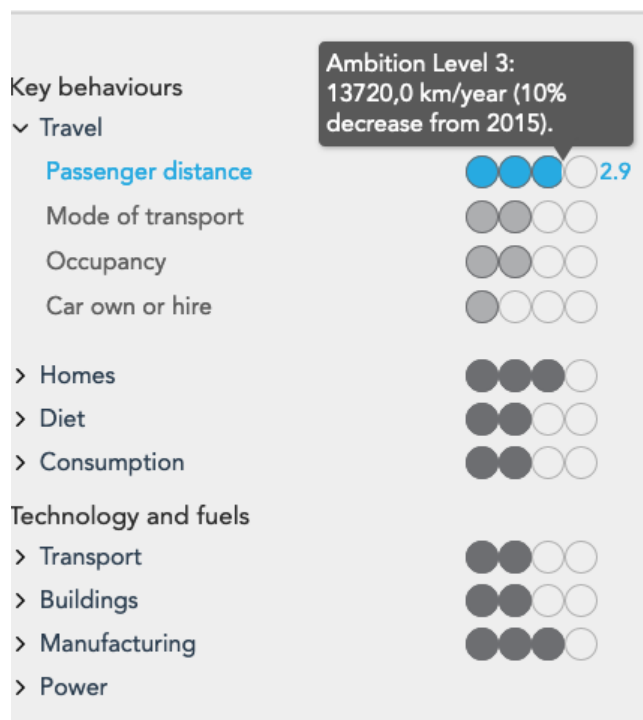


Figure 32. Current status of the TPE: Screenshot of a part of the lever list for user interaction in the web application (Transition Pathways Explorer). When opening the web application, only the lever group names are visible. Changing the ambition level of a lever group changes the ambition levels of all levers belonging to that group accordingly. Clicking on the name of the group opens a drawer to show the individual levers. The “group ambition” is computed as the average of the individual lever’s ambitions. A click on the lever name opens a pop-up with the description of the rationale of the lever, some background information and the time evolution of the ambition levels one to four. Hovering on the ambition levels prompt short descriptions.

## VI. REFERENCES

[Butkus et al., 2018] Butkus, M., Cibulskiene, D., Maciulyte-Sniukiene, A., & Matuzeviciute, K. (2018). What Is the Evolution of Convergence in the EU? Decomposing EU Disparities up to NUTS 3 Level. *Sustainability*, 10(5), 1552.

[European Parliament, 2018] M.Dolls, C.Fuest, C.Krolage, F.Neumeier, D.Stöhlker (2018). Convergence in EMU: What and How?. Available at : [http://www.europarl.europa.eu/RegData/etudes/IDAN/2018/614502/IPOL\\_IDA\(2018\)614502\\_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/IDAN/2018/614502/IPOL_IDA(2018)614502_EN.pdf)

[Felton 2008] Felton, N. (2008). Consumption spreads faster today. *New York Times*, 10.

[IPCC, 2001] Intergovernmental Panel on Climate Change (2001). IPCC Good Practices Guidance and Uncertainty Management in National Greenhouse Gas Inventories, accepted by the IPCC Plenary at its 16th session held in Montreal, 1-8 May, 2000, Corrigendum, June 15, 2001.

[Rogers, 1995] Rogers, E. M. (1995). Diffusion of Innovation. 4th. *New York: The Free*.

[Rousseeuw, 1993] Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424), 1273-1283.

[Science Based Target, 2015] CDP, UN Global Compact, WRI, WWF (2015). Science Based Target – How businesses are aligning their goals with climate science.