

From the development of an open-source energy modelling tool to its application and the creation of communities of practice: the example of OSeMOSYS

Francesco Gardumi^{a*}, Abhishek Shivakumar^a, Robbie Morrison^b, Constantinos Taliotis^a, Oliver Broad^c, Agnese Beltramo^a, Vignesh Sridharan^a, Mark Howells^a, Jonas Hörsch^d, Taco Niet^e, Youssef Almulla^a, Eunice Ramos^a, Thorsten Burandt^f, Gabriela Peña Balderrama^a, Gustavo Nikolaus Pinto de Moura^g, Eduardo Zepeda^h, Thomas Alfstad^h

^a Department of Energy Technology, KTH Royal Institute of Technology, Brinellvägen 68, Stockholm, Sweden

^b Schillerstraße 85, 10627 Berlin, Germany

^c Institute for Sustainable Resources, University College London (UCL), London, United Kingdom

^d Frankfurt Institute for Advanced Studies, Frankfurt, Germany

^e University of Victoria / British Columbia Institute of Technology, Canada

^f Workgroup for Economic and Infrastructure Policy (WIP), Technical University of Berlin, Berlin, Germany

^g Universidade Federal de Ouro Preto, Minas Gerais, Brazil

^h United Nations Department of Economic and Social Affairs, Development Policy and Analysis Division, New York, United States of America

Emails: gardumi@kth.se, ashi@kth.se, robbie.morrison@postdeo.de, taliotis@kth.se, o.broad@ucl.ac.uk, beltramo@kth.se, vsri@kth.se, mark.howells@energy.kth.se, hoersch@fias.uni-frankfurt.de, Taco_Niet@bcit.ca, almulla@kth.se, epramos@kth.se, thb@wip.tu-berlin.de, gabrie@kth.se, gustavomoura.ufop@gmail.com, ezepedam@gmail.com, talfstad@gmail.com

Corresponding author: Francesco Gardumi, email: gardumi@kth.se

Keywords

Energy system modelling tool; Open-source software; Model-based public policy; Software development practice; Outreach practice;

Abstract

In the last decades, energy modelling has supported energy planning by offering insights into the dynamics between energy access, resource use, and sustainable development. Especially in recent years, there has been an attempt to strengthen the science-policy interface and increase the involvement of society in energy planning processes. This has, both in the EU and worldwide, led to the development of open-source and transparent energy modelling practices.

This paper describes the role of an open-source energy modelling tool in the energy planning process and highlights its importance for society. Specifically, it describes the existence and characteristics of the relationship between developing an open-source, freely available tool and its application, dissemination and use for policy making. Using the example of the Open Source energy Modelling System (OSeMOSYS), this work focuses on practices that were established within the community and that made the framework's development and application both relevant and scientifically grounded.

Acronyms

CLEWs	Climate, Land, Energy and Water strategies
EMP-E	Energy Modelling Platform for Europe
GAMS	General Algebraic Modelling System
GENeSYS-MOD	Global Energy System Model
IEW	International Energy Workshop
INDCs	Intended Nationally Derived Contributions
LEAP	Long-range Energy Alternatives Planning
MoManI	Model Management Infrastructure
O&M	Operation and Maintenance
OSeMOSYS	Open Source energy MOdelling SYStem
POTEnCIA	Policy Oriented Tool for Energy and Climate Impact Assessment
RES	Reference Energy System
SAMBA	South America Model Base
TEMBA	The Electricity Model Base for Africa
TIMES	The Integrated Markal EFOM System
WEAP	Water Evaluation and Planning system

1. Introduction

The link between sustainable development and access to energy is well discussed in the literature [1] and the United Nations have affirmed their commitment to the challenge of granting universal access to sustainable, affordable, reliable and modern energy by 2030 [2]. Society as a whole has a role to play in planning and developing better energy infrastructure that supports wider access to energy. Policy makers guide such developments through funding and policy frameworks, suppliers and consumers drive the energy markets and representatives of civil society in academia, institutions and non-governmental organizations present their views on the impacts of energy access.

Energy modelling can be useful for different components of society, offering insights into the dynamics between energy access, resource use and sustainable development [3]. In recent years, the attempt to reach out to and engage with society at different levels has, both in the EU and worldwide, led to the practice of open-source and transparent energy modelling and the availability of open energy system data. The Open Energy Modelling Initiative, for instance, categorises a variety of open tools and applications [4] in a space that offers energy modellers the opportunity to collaborate around open-source tools and practices [5]. At the EU level, the Energy Modelling Platform for Europe (EMP-E), established in 2017, provides a digest of both closed source and transparent energy modelling frameworks as well as a meeting point for decision makers and scientists [6]. The European Commission is, itself, committed to developing open-source tools (such as POTEnCIA – Policy Oriented Tool for Energy and Climate Impact Assessment [7]) and encourages the use of open-source and transparent practices in funded research projects, including those under the Horizon 2020 programme.

Open-source initiatives need to be maintained, improved, disseminated and communicated. Unlike their closed source counterparts, an entire community of developers and users is available to contribute to this effort, in a decentralised and independent fashion. On the one hand, this provides abundant resources for continuous development and a wide review of every application; on the other, it requires well defined structures and standardised processes that guarantee quality and coordination in the development. In addition, contributions need to be recorded and authors acknowledged.

These challenges have been widely debated in international modelling communities such as the Open Energy Modelling Initiative, the EMP-E, the European Geosciences Union (EGU) and the International Energy Workshop (IEW). However, not all of them have been systematically and comprehensively

addressed in peer-reviewed literature. This paper contributes to such debates by discussing benefits and challenges of the development and exploitation of an open-source energy modelling tool with a large and growing community of users. Specifically, through the example of the Open Source energy Modelling System (OSeMOSYS), it argues the relationship between the development of an open-source, freely available tool and its application, dissemination and use for policy making. This discussion focuses especially on aspects related to the need for transparency, the quality control and the open-source governance. Section 2 introduces the tool, describing its design principles, its core features, its applicability and its interface; it then gives an overview of key code developments by the community; Section 3 shows applications of the original and the modified versions of OSeMOSYS, drawing conclusions on the potential and limitations of the tool; from there Section 4 moves, describing established open-source governance practices; it then zooms into how the developments and applications were employed in outreach and stakeholder engagement activities, pointing out also licensing issues. Section 5 concludes the paper.

2. Tools and methodologies

2.1. Overview of OSeMOSYS

OSeMOSYS (the Open Source energy MOdelling SYStem) is an open-source and freely available bottom-up modelling framework for the long-range optimisation of the energy system and energy mix of user-defined regions [3]. The term ‘modelling framework’ in this context designates software that generates specific models by populating them with user-defined data. OSeMOSYS has been employed in the scientific literature, in academic teaching and in capacity building for energy planners, to provide insights on possible transformation pathways of large energy systems and their impacts on the economy, the society and the environment. It is one of the open-source tools featured by the OptIMUS Community, a community of practice led by United Nations Development Program and United Nations Department of Economic and Social Affairs and aimed at fostering the use of open-source modelling tools for capacity building on sustainable development.

In many respects, the structure and the insights OSeMOSYS provides are similar to those offered by tools like MESSAGE [8] and TIMES [9,10]. The first functional version of OSeMOSYS (initially named Soft-MESSAGE) was presented and made available at the 2008 International Energy Workshop (IEW) in Paris [11], while the first academic paper describing its ethos and structure was published by Howells et al. in 2011 [3]. OSeMOSYS was born from an effort of multiple institutions, including co-authors from UCL, UNIDO, UCT, Stanford, PSI and other institutions. It has been funded through in-kind contribution of the involved partners and through projects that allowed a testing ground for improvement.

2.1.1. Design principles

OSeMOSYS is a linear program, seeking the energy mix (in terms of technology portfolio and generation) which simultaneously meets exogenously defined – elastic or inelastic – energy demands and minimises the total cost of the system, discounted to an initial reference year, under predefined constraints, and from the perspective of a single decision-maker. Constraints can include conversion efficiencies of supply and end-use technologies, user-defined relations between different types of energy inputs and outputs, upper limits on emissions, lower limits on renewable generation, upper and lower limits on investments and energy and power capacity balances.

OSeMOSYS is structured in blocks of functionality, each one consisting of a stand-alone set of equations and inequalities which can be plugged-in to the core code to add specific insights for the case-study of interest. Three levels of abstraction characterise each block: a plain English description, an algebraic formulation and its code implementation. This makes the tool accessible to a wide audience, with an easy learning curve. Examples from academia have shown that novice students are able to learn to use

the tool, develop a country model and gain relevant insights in the course of a six-month master's thesis [12–20].

The reference version of OSeMOSYS is written in GNU MathProg, an open-source modelling language where coded equations resemble their algebraic form. It is available as a text file on GitHub [21]. To create and run an energy system model, the user needs this text file, together with an input data text file formatted as compatible with GNU MathProg and a solver (for instance the freely available GNU Linear Programming Kit – GLPK). The advantage of GNU MathProg is the fast learning curve, which makes the code of OSeMOSYS accessible also to users without prior experience of linear programming. Versions in GAMS and Python were later developed and made available on GitHub [21]. These aim to engage with the large community of users of the two modelling languages and make OSeMOSYS directly linkable with libraries and models written in GAMS and Python. The difference in the code formulation between GAMS and GNU MathProg is limited, while the Python version diverges more significantly. A comparative example of the code formulation in the three modelling languages, referring to the objective function of OSeMOSYS, is presented in Appendix A.

2.1.2. Core code

The first version of the OSeMOSYS code was released in 2008 [11]. Howells et al. detailed the original structure in [3]. They presented it in three levels of abstraction - qualitative description, algebraic formulation and code formulation - establishing what would become a standard in the development of the tool. The original structure consisted in seven blocks of functionality, focusing on:

- *The objective function*, which estimates the lowest Net Present Cost (NPC) of the energy system to meet exogenously defined energy demands;
- *Costs*, defined by a set of equations making an account of the capital and O&M technology costs and discounting them to the base year;
- *Storage*, which defines balances and limitations for stored energy.
- *Capacity adequacy*, which ensures the model estimates enough capacity to meet the required energy demand in each of the user-defined time slices;
- *Energy balance*, which ensures the annual balance of energy vector production and consumption along the entire energy chain;
- *Emissions*, which accounts for the emissions released within the modelled time frame, considering user-defined emission limits and/or emission penalties;
- *Constraints*, that allows the user to impose limits on the total installed capacity or production of technologies.

A number of additional blocks of functionality were released later on and relate to:

- An improved representation of storage;
- The provision of reserve capacity [22];
- The cost of cyclic operation of fossil fuel-fired power plants [23];
- Smart-grids and demand-side flexibility [24].

Note that subsequent code releases also contained minor modifications to existing blocks of functionality, mainly to fix bugs, refine the constraints to impose a minimum renewable generation target and simplify trade between regions.

Appendix B lists the key inputs and outputs that the user can feed into and obtain from an OSeMOSYS model, when using the version of the code currently published on GitHub.

2.1.3. Applicability

As mentioned, OSeMOSYS seeks an energy mix which meets exogenously defined energy demands and minimises the total discounted system cost under predefined constraints. The key elements of the

energy mix are distinguished in the code as fuels and technologies. A technology is intended as a black-box with user-defined transfer functions and characteristics. For instance, the user may represent an ultra-supercritical coal power plant by defining a technology with coal as input, electricity as output, 45% efficiency, 2000 \$/kW investment cost and 88 \$/kW/a operation and maintenance (O&M) cost (sample data from [25]); equivalently, he/she may represent a car by assigning diesel as input and passenger-km as output, or a transmission line assigning electricity as input and electricity as output. A fuel in OSeMOSYS is any energy vector, entering or exiting a technology. For example, imported oil, domestic oil, refined oil, coal, electricity for transmission, electricity for distribution, high temperature heat, diesel, passenger-km are all named fuels. The final energy demands are named fuels, too. With such flexible definitions for technologies and fuels, OSeMOSYS can be applied to any sector within the energy system. Figure 1 shows the Reference Energy System (RES) of an OSeMOSYS model of Cyprus developed by Taliotis et al. [26]. The RES is a schematic representation of the chain of technologies and fuels as described in the model, from the energy supply (on the left) to the exogenously defined demands (on the right).

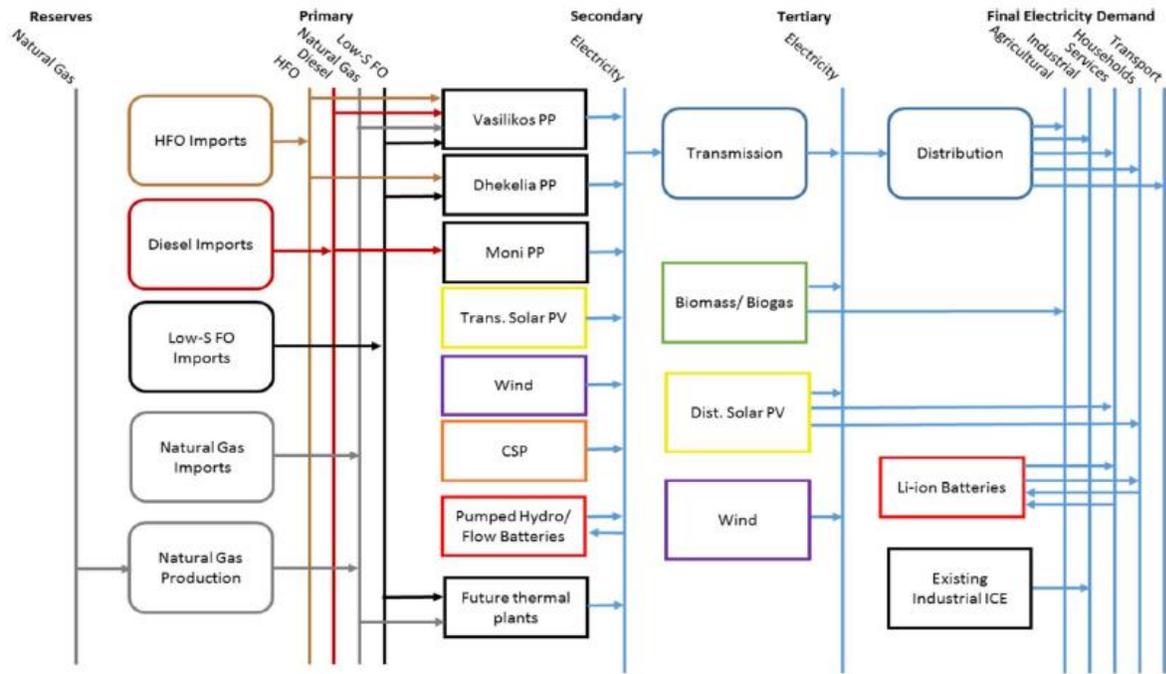


Figure 1. Reference Energy System (RES) of the Cyprus model [26].

The absence of an embedded set of technologies and fuels in OSeMOSYS and the flexibility in their definition also allows the user to vary the scale of a study, from a village to a continent. The model of Suro Craic, a village in Timor Leste was published in [27], while a model of the African continent was presented in [28].

The concept behind OSeMOSYS can be extended not only to the modelling of energy systems, but also to the modelling of the nexus between Climate, Land-use, Energy and Water (CLEW, discussed in [29]). A very simplified structure of an extended 'RES' including links between such sectors is provided in Figure 2.

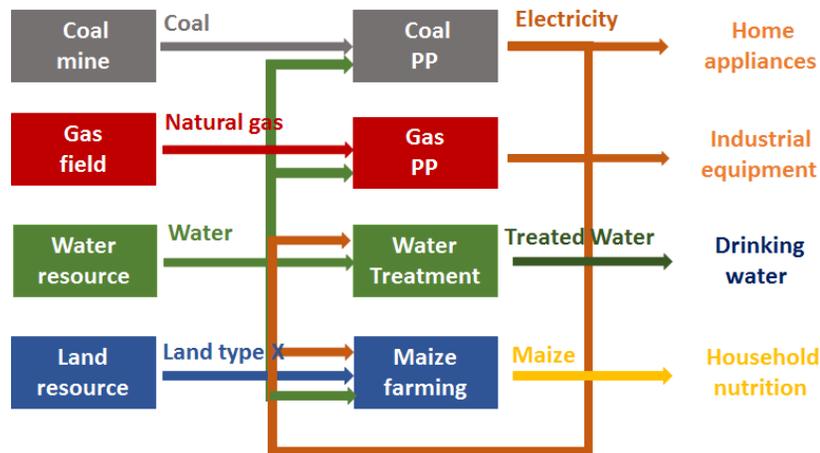


Figure 2. Conceptual schematic of a CLEWs model. Source: *OpTIMUS.community*.

Complex CLEWs models have been run through composite modelling frameworks soft-linking OSeMOSYS with hydrological and land-use models. Details on how the soft-linking is developed in practice are available in [29].

Summarising, the flexibility in the definition of the fundamental elements of OSeMOSYS, such as technologies and fuels, makes it suitable for a variety of applications covering a wide range of spatial and temporal scales. The computational barriers which would prevent high spatial and temporal resolutions have recently been greatly reduced with a reformulation of the code, as described in Section 2.2. Limitations emerge, nonetheless, when the field of application is extended:

- OSeMOSYS was originally designed as an energy modelling system. The naming of input parameters and output variables was therefore found to be inappropriate when applied to integrated systems modelling including land and water resources. New model versions with more comprehensive naming structures are thus being developed with support from the community;
- Thanks to a recent code reformulation, OSeMOSYS can be used for short-term modelling and be applied, for instance, to the study of security aspects on electricity grids. Nonetheless, formulations to account for short-term characteristics of unit dispatch, such as e.g. minimum and maximum downtime, are yet to be introduced.

2.1.4. Interface

The process of developing an energy system model in a framework such as OSeMOSYS is non-trivial. It requires the mapping of primary resources and their conversion chains down to their final uses. Therefore, for the tool to be fully accessible, it requires an open-source, user friendly, yet explanatory and comprehensive interface. The Model Management Infrastructure (MoManI) was recently developed as one interface to create, manage, run, store inputs and outputs of linear programming models such as those built in OSeMOSYS. It is freely available both as a browser-based and as a desktop version. Using MoManI as an interface for OSeMOSYS, the user can:

1. Modify the structure of OSeMOSYS (i.e. change the equations and inequalities in its core code);
2. Develop the energy system model of a user-define region;
3. Develop a set of scenarios to explore different technology options and/or to address various policy questions;
4. Run the optimisation, currently performed through the GLPK solver;
5. Visualise the results.

Implementing and operating OSeMOSYS in MoManI shortens the learning curve and reduces barriers associated with the complexity of computer modelling, thereby supporting effective teaching and capacity building activities. MoManI has been employed in higher education courses at KTH Royal Institute of Technology and in several capacity building activities (for instance, [30,31]). The structure of MoManI allows different teams to collaborate simultaneously from around the globe.

2.2. Code developments

The open-source nature and the accessibility of the tool have allowed the community of developers to both modify existing code blocks and create new ones independently over the years. A few examples are listed and briefly described below. They range from structural modifications of existing blocks of code, to extending the applicability of the tool or developing new plugins adapted to case-specific insights.

Reduction of the translation time.

One of the highest barriers to the creation of large continental models with detail to the country level, or highly detailed national applications, is the time and memory requirement to translate a model into a solvable matrix. In order to increase the applicability of the tool, a revision of the OSeMOSYS code was completed in late 2016, as the result of a public call launched and led by UNite Ideas.

Much of the flexibility of OSeMOSYS arises from the broad definition of all energy components as technologies and all energy carriers as fuels. The connections between them are specified in the data file by the user in the form of region- and time-specific tensors named *Input-* and *OutputActivityRatio*, that define the rate at which each fuel goes in and comes out of an active technology. Since in the majority of cases each technology converts one fuel into another – e.g. a gas turbine turning gas into electricity – the expected density of these tensors is very low, on the order of 1 over $|\text{regions}| \times |\text{technologies}| \times |\text{modes}| \times |\text{fuels}|$. While previous versions of OSeMOSYS already filtered out technology-fuel relations with a zero rate and presented a straightforward linear problem to the solver, the amount of time used by having to apply this filter was repeatedly underestimated.

The updated version speeds this process up by generating intermediate parametric sets from a single scan over the sparse connection tensors after reading in the data file. These sets hold all the combinations of technology and modes of operation which consume or produce a specific fuel, and the rest of the model definition refers to them exclusively when considering technologies and fuels without any loss of generality. The updated version of OSeMOSYS including these changes is available on GitHub [21].

All measures combined have been shown to reduce the time for translating the MathProg language of a reference country case-study from 526 seconds to 38 seconds on an Intel Core i5-2520M processor, a factor of 13; other cases show a similar speed-up. The subsequent time for solving the model is not affected. Further, when the code adjustments were tested on one of the largest OSeMOSYS models, the TEMBA model, considering the electricity supply system of 47 African countries, the sum of the translation and the calculation time was reduced from 18 hours to under 2 hours.

Development of an OSeMOSYS-GAMS version.

Due to the high distribution of GAMS (General Algebraic Modelling System) in academic and economic environments, the provision of an OSeMOSYS version in this modelling language is important to allow for further dissemination. The first available GAMS version of OSeMOSYS was provided by Ken Noble [32] and updated to the newest version by Löffler et al. [33] while developing the Global Energy System Model (GENeSYS-MOD). This application of OSeMOSYS, as well as its modular additions, will be made publicly available and updated regularly going forward.

The GAMS version of OSeMOSYS, similar to the one in GNU MathProg, engages with the large community of GAMS users. Its major drawback is that the language is not open-source and, since OSeMOSYS exceeds the GAMS demo limitations, expensive licenses are needed. On the other hand, its strength lies in the code formulation and data file structure being very similar to the one in GNU MathProg, allowing users to easily shift from one version to the other.

OSeMOSYS-Python version.

In recent years, Python has emerged as among the most popular and powerful programming languages available to the research community. This has been in part due to the availability of specialised libraries to perform functions such as data handling, optimisation, visualisation and parallel processing. A version of OSeMOSYS was developed in Python in order to gain access to this rich library of additional functionalities and an active programming community. OSeMOSYS-Python is written in Pyomo, a Python-based open-source optimisation modelling language. It is available on GitHub [21] and its structure follows that of the original OSeMOSYS version written in GNU MathProg. The main advantages of using the OSeMOSYS-Python version are: access to vast and growing number of Python libraries for handling large datasets (pandas), sharing code online (Ipython/Jupyter notebooks) and visualisation (matplotlib, seaborn); large and active community of Python users (stackoverflow, reddit); ability to use data files in AMPL format, allowing backwards compatibility with data files created for use with the OSeMOSYS-GNU MathProg and OSeMOSYS-GAMS versions.

Short-term operational constraints of power plants.

The global push towards increasing the share of Renewable Energy Sources (RES) in the energy supply poses challenges in terms of security and adequacy of electricity networks. Intermittent generation from e.g. wind and solar power may require back up. Two of the widely discussed options include controllable fossil fuel-fired generation (such as Open Cycle or Combined Cycle Gas Turbines) and storage. In order to assess the costs and benefits of using the first as a back-up option for peaking generation, new blocks of functionality computing short-term costs and operational constraints of dispatchable generation were designed. These include:

- Computing the reserve capacity dispatch to meet an exogenously given demand, under constraints on ramp rates and minimum duty [22]: the enhanced version of OSeMOSYS was compared to a modelling framework coupling TIMES and PLEXOS, through a case-study analysing optimal energy infrastructure investments in Ireland in 2020 [34]. While avoiding the high computational burden of the TIMES-PLEXOS model (the time resolution of the latter is 700 times higher), the OSeMOSYS model provides similar results. For instance, the investments diverge by 5%.
The new block of functionality was further modified to make the reserve capacity demand an endogenous variable, namely a function of the penetration of intermittent renewables [13].
- Costs related to the flexible operation of power plants, specifically: increased specific fuel consumption at lower load, wear and tear costs associated to the number of ramp-up and ramp-down cycles and costs for refurbishing existing units [23].

Revision of the storage block of functionality.

This block of functionality was significantly revised and particular effort put into ensuring that the intra-day exchange of stored energy respects the storage capacity even when the time resolution of the study is coarser. The revised version of the storage block of functionality is embedded in the latest published version of the code.

OSeMOSYS for short-term planning.

This version of OSeMOSYS was developed to further evaluate the short-term performance characteristics of systems with a high penetration of variable RES. It stems from the original code, enhanced by both the short-term operational constraints and the storage block of functionality described above. A number of additional modifications were introduced in order to improve the applicability of OSeMOSYS to finer time resolutions. Their focus was to preserve the temporal sequence of renewable energy availability and to evaluate the reaction of storage and other system management techniques to these dynamics. Specific changes include:

- Revised storage equations that are more computationally efficient for short-term modelling. Specifically, the intra-time slice storage equations in the base OSeMOSYS code were replaced with inter-time slice equations. This allows for much faster computation of the storage levels and allows for a larger number of scenarios to be computed in a shorter amount of time.
- Equations that model the ramping constraints of conventional generators. With large penetrations of variable renewables, the ramping demand in the system is significantly increased. The ability to constrain the ramping capabilities of generators in the system allows for a more accurate representation of the system dynamics and associated costs.
- Equations that incorporate the cost of curtailment into the model. This is not usually accounted for in a long-term model due to the averaging imposed by the time slice definitions.

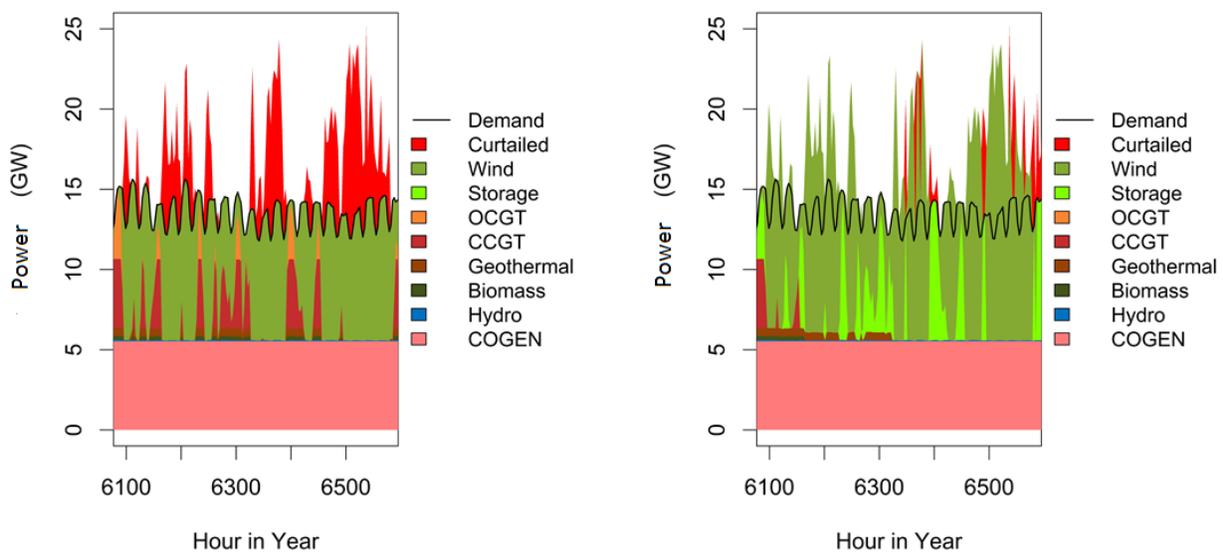


Figure 3: Comparison between the power generation profile without and with storage.

Figure 3 shows results obtained when using OSeMOSYS for short-term planning. The curtailed energy is marked in red above the demand line. Energy stored for future use is shown in light green.

Cascaded water storage.

This addition to the basic OSeMOSYS storage equations allows for cascaded facilities to be included with an upper reservoir and generation station feeding water into a lower reservoir for a second generation station. Further, it is designed to track the storage levels and water flows, in and between each reservoir. Constraints representing minimum and maximum output flows from each dam are included to model both flood control and fish habitat management. Specific focus was put on maintaining a similar storage structure as the base OSeMOSYS code while providing a more user friendly formulation for modelling hydroelectric generation.

This application has proven important in the study of integrated approaches for the management of water and energy resources [35].

The cascaded hydro storage equations have been uploaded to GitHub where anyone interested can download, use and modify them for their own purposes [36].

3. Key applications

All the features listed above were developed to support a number of applications, varying in both target audience and scale. In line with common energy modelling practice, these applications are designed to provide insights – rather than numbers – into the impacts of energy transition pathways and the causality linkages between elements of the system. In turn, the insights inform outreach activities and address the concerns of a wide variety of stakeholders.

The main OSeMOSYS applications can be categorised based on their sectoral coverage and on the scale of the study. The sectoral coverage can be divided into Energy and CLEWs (Climate, Land, Energy and Water strategies). The scale of the study can be clustered into national, regional or global.

In the following, some of the main applications of OSeMOSYS are listed, divided in the abovementioned categories. More details about each of these applications are given in Appendix C.

Energy modelling

- **National scale:**

Natural gas outlook for Cyprus. A model of the energy system of Cyprus, including energy, heating, cooling and transportation demand, with detail to the power unit level. Employed to provide insights to the Government on the benefits of exploiting gas resources in a high-renewables scenario [26,37].

Meeting the INDCs in Bolivia. A single-node model of the power generation sector of Bolivia is developed in OSeMOSYS and linked with a detailed energy demand model developed in the Long-Range Energy Alternatives Planning (LEAP) software. The models provided insights for the Ministry of Development Planning and the Division of Economic and Policy Analysis (UDAPE) to write policy notes [38] and submit the Intended Nationally Determined Contribution (INDC) to the Conference of Parties (COP21) in Paris [39].

- **Regional scale:**

Resilience of the African electricity infrastructure. Energy models developed for each of the four sub-Saharan African power pools with detailed hydropower from seven major river basins. The model, featuring high temporal resolution (48 annual time slices) and spatial resolution (every major hydro power plant modelled separately), aims to provide insights on the resilience of the electricity infrastructure to climate change scenarios [40].

TEMBA (The Electricity Model Base for Africa). Electricity system model of the African continent, including each of the 47 continental countries and the transmission links between them. Scenarios up to 2040 provide insights into the impact of investments in transboundary transmission links: they have a potential to alter the electricity generation mix of the countries and reduce the electricity generation costs [28].

SAMBA (South America Model Base). Model of the electricity supply infrastructure of South America. It includes 11 countries, represents transmission links between them and breaks Brazil down into four interconnected regions. It has a high temporal resolution (48 annual time slices) and includes detailed and county-specific representation of generation technologies in all countries. Insights are derived on the role of hydropower for electricity supply and trade, but also on the relative importance of specific input parameters as key drivers for the overall system costs [41–43].

- **Global scale:**

GENeSYS-MOD Global Energy System Model. Model of the global energy supply system, aimed to assess the feasibility of global decarbonisation pathways, under GHG emission reduction targets and decreasing prices of renewable energy technologies [33].

CLEWs modelling

- **National scale:**
Water-energy-food nexus in Uganda. An energy system model of Uganda developed in OSeMOSYS is soft-linked with a detailed hydrological water balance model developed in the Water Evaluation and Planning tool (WEAP) [44]. The composite modelling framework is used to analyse the impact of a number of future climate scenarios (defined exogenously) on the use of water resources and the competition between demand sectors.
- **Regional scale:**
Transboundary water-energy resources assessment. Three models investigating the nexus between the availability of water resources from river basins and their use for irrigation and energy generation are developed. The first two models focus on the Western Balkans, respectively on the Sava [45] and on the Drina river basin [46]. The third model focuses on the Syr Darya river basin, in central Asia.
- **Global scale:**
GLUCOSE – The Global CLEWs Model. Demonstrative model aimed at supporting strategies for sustainable development [47]. It has high spatial aggregation and includes the six most material-intensive industries. It provides insights on the dynamics caused by constraints such as greenhouse gases taxes, land cost, limitation of agricultural land and global emission caps.

3.1. Insights and limitations

From all these applications, OSeMOSYS emerges particularly as an engagement tool.

First of all, it is useful to engage with policy makers. The straightforward comprehension, jointly with the free and open-source nature, prove to abate barriers between these stakeholders and the technicians working on the models. Specifically, the documentation on the code and the models provide material for trainings (as in the case of Cyprus and Bolivia, for instance); the terminology used, deliberately descriptive and as close as possible to the common sense, facilitates the communication among modellers and between the modellers and the policy makers (one example is the OSeMOSYS Pan European engagement model under development within the Horizon 2020 REEEM project [48]); the ease of use allows staff from the institutions to be directly involved in the modelling, providing the ground for capacity transfer.

Secondly, large engagement with the research community is achieved, once again due to the free and open-source nature, but also due to the availability of the tool in three modelling languages. The latter has facilitated the communication with researchers of diverse expertise (e.g. economists used to modelling in GAMS), enlarging the range of developments and applications.

This all benefits greatly from a synergy with the use – as far as feasible – of openly available input datasets and the open-access publication of inputs and results of the models. Where the inputs are covered by copyright, they can be published after aggregation, thus overcoming what is acknowledged as a huge barrier to transparency.

However, also limitations emerge from the described applications.

Firstly, every bottom-up tool requires a number of techno-economic inputs. In many cases, the scarcity of sources for the input data forces the modellers to make assumptions. This may affect the usefulness of the insights provided by the models. Iterations to allow for future refinements of the inputs, upon increased availability, must always be foreseen.

Secondly, as the range of application is wide, both the modellers and the recipients might at times lose the view on what the focus and utility of the tool is. This paper and others – already published or yet to

come – on methodologies and applications are intended to establish terminology, methodological frameworks, literature background and examples to help address this limitation.

Lastly and more generally, the openness and the availability of the OSeMOSYS code in three modelling languages leads to the multiplication of the documentation and of the code versions used for different applications, calling for the creation of a community management structure. This is nothing but a set of rules, agreed by representatives of the community, to be followed when developing and applying the tool. Such structure is presented in the following Section.

4. Open-source governance of the tool

The experience from OSeMOSYS shows that the work for proposing, developing, reviewing and distributing code updates and new functionalities lies at the core of the organisation of a large open-source community. Introducing structured processes for this facilitates not only the work of the code developers, but also the engagement of modellers and other families of users. Recently, a formalised process was put in place to deal with the rapid growth of both software complexity and OSeMOSYS user/developer base. It was structured in such a way as to:

- Involve the whole community;
- Be transparent in every phase;
- Be scientifically rigorous;
- Rely solely on shared documentation;
- Acknowledge the contributors;
- Be resource-efficient.

In some of its phases, the newly defined process resembles the peer-review process for scientific journals. It involves four main figures: the Community Manager, acting as an Editor; the OSeMOSYS Steering Committee, acting as the Editorial Board; Reviewers, chosen among a pool of experts contributing to the OSeMOSYS Community; and the Developers, who are equivalent to the Authors of a paper.

The main tools of the development and review process are freely and openly available online platforms, which the users have to register to: a Reddit forum [49], a GitHub page [21] and a Read the Docs page, yet to be completed [50].

A flow diagram and a step-by-step description of the process are provided in Appendix D.

These processes are structured so as to meet the requirements listed above, particularly resource-efficiency. When tools are open-source and freely available, specific sources of funding may be necessary. Therefore, it is crucial that the contributors be involved based on their specific expertise and on the contingent needs of the research they are already carrying on within funded research programmes, doctoral studies and Bachelor's and Master's curricula.

5. Outreach

The Steering Committee manages, among others, the process of developing and sharing OSeMOSYS versions and functionalities, thus ensuring that up-to-date peer-reviewed modelling material is continuously and openly made available online. This material is then exploited by two main groups of users: academia and policy makers. In the following subsections, a brief description is given of how OSeMOSYS is used in these two cases and what considerations are required to ensure coordination between the developers and the families of users.

5.1. Academia

One place where the scientific effort on the development of OSeMOSYS naturally finds application is academia. In particular, three fields of application are equally important:

- Higher Education Teaching;

- Master's and Doctoral thesis studies;
- Funded research projects.

Within each of these fields, the core of the scientific production to be disseminated to and exploited by the whole Community comes to life. At Bachelor's and Master's teaching level, teaching modules, including theory and application, are created in different institutions. A standard format for the teaching material is in the course of being agreed between several institutions within the OpTIMUS.community [51], with the aim of making the modules self-standing, shareable, useable and modifiable by any trainer. The minimum requirements for this to happen are the acknowledgement of the original author of the material, notes for the speakers and guidelines for laboratory sessions. On the other hand, both within Master's and Doctoral thesis studies and within funded research projects, a number of models is produced, together with peer-reviewed publications explaining their structure, hypotheses, results and key insights. Thus, such research activities indirectly provide funding for high quality, peer-reviewed developments and applications of the tool, which could otherwise lack funding by itself.

5.2. Policy makers

In many cases, the scientific material is produced in academia jointly with policy makers. In such instances, the work focuses mainly on developments and applications that will help to derive model-based insights into the optimal design of energy and resource supply systems. Starting from expressions of interest of specific governments to international organisations (such as United Nations or World Bank), national models are developed in academia in tight collaboration with local experts and are transferred back to local institutions using a structured capacity building process. A sample structure for this is detailed in Figure 4. It shows a capacity building process for CLEWs activities.

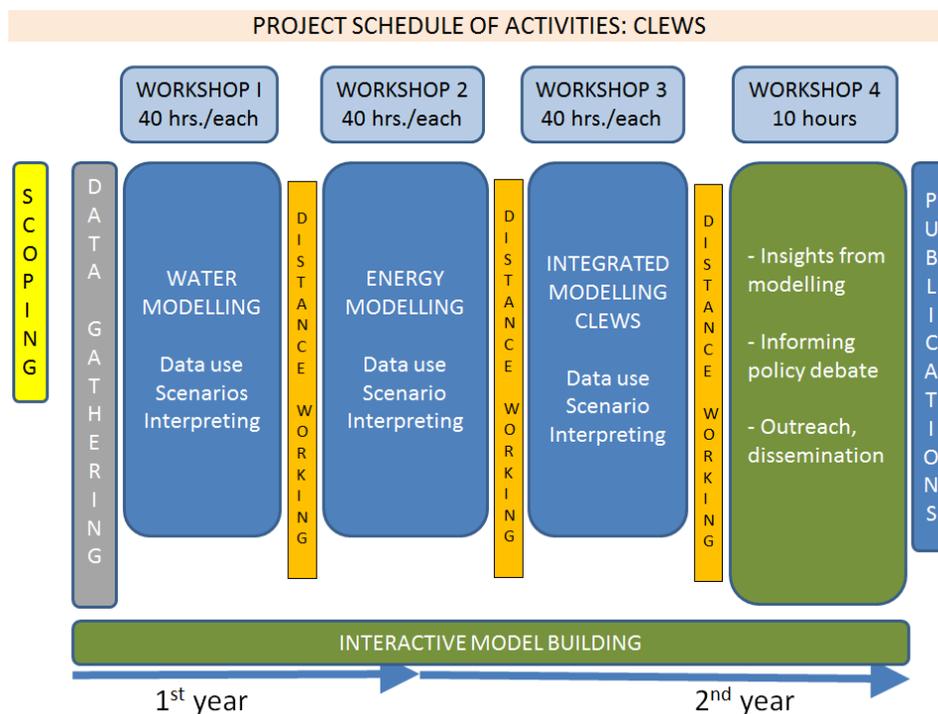


Figure 4. Sample scheme of a capacity building process involving a CLEWs assessment. Source: OpTIMUS.community.

Instructors are selected from a list of expert OSeMOSYS Reviewers who have experience in coding and problem solving ('debugging'), teaching in Higher Education Institutions, and academic application of OSeMOSYS.

5.3. Licensing

The diffusion of open-source software and its related applications or teaching material has to be regulated by open licences. These add binding conditions to a copyright that grant users free use of a piece of work. Users can also modify and redistribute the work, given that they embed the original license (or a less permissive but compatible alternative license) and comply with all other terms. Open licensing is therefore an essential feature of open projects. Owing to their specialist nature, licences have evolved specifically for software, data, and documentation.

In 2009 an Apache 2 software license was added to the OSeMOSYS codebase. Software licenses vary in terms of permissiveness, but models written in an algebraic modelling language (such as MathProg and GAMS) do not benefit from using the more restrictive GNU GPL licences, for instance. Under an Apache 2 license, users are free to study, use, improve, and redistribute the associated source code, the only restriction being that, on redistribution, the contributors are acknowledged. The Python version of OSeMOSYS also carries an Apache 2 licence.

Unlike open licensing for software, open data licensing is a new field with limited case law and legal analysis. OSeMOSYS as a modelling tool does not collect and disseminate data explicitly, but rather modelling projects using OSeMOSYS are a consumer of open data from appropriate sources, some tailored to energy modelling (such as OpenEI [52] and Open Power System Data [53]) and some more general (such as official statistics offices).

The law on open documentation, and written content more generally, is in place. The use of OSeMOSYS favours the Creative Commons Attribution (CC BY 4.0) license for documentation and teaching material. The OSeMOSYS community prefers that scientific papers be published under open access as funding allows.

6. Discussion and conclusions

This paper describes the existence and characteristics of the relationship between the development of an open-source, freely available tool and its application, dissemination and use. This flow can be channelled and made sustainable by establishing a set of clear practices, structures and processes. The authors provide an example from their experience with the Open Source energy Modelling System (OSeMOSYS). The prime enabler for the engagement of a wide community of users is the open-source, free and straightforward nature of the tool. However, as the community grows and diversifies, this is not sufficient anymore. The structures proposed in this work for the development of new functionalities serve the purpose of enabling the wide potential of independent contributions by the community, while ensuring high quality standards, peer-review and scientific merit. If the following conditions are verified:

- the developments and test-applications of the tool suggested by the Community and planned by the Steering Committee are addressed by scholars according to their research questions,
- the developments, test-applications and real applications are documented according to agreed standards along the whole process, allowing for the whole work to be published,

the resources for the development and maintenance of the tool can be ensured.

The outcome of the tool development phase is a number of solid, documented and reviewed functionalities, ready to be applied in academia and in capacity transfer activities with decision makers. In academia, they are used as illustrative examples for tertiary education courses, to introduce students from different backgrounds to energy systems modelling, as well as platforms for research. In capacity

building activities, models of villages, countries, regions and continents are created, in collaboration with staff from utilities, Ministries and academy, and are transferred to the staff of all these institutions through series of trainings. As discussed in Section 3, the accessibility of the structure of the models, together with the openness of the datasets employed, facilitates the knowledge transfer. The engagement of local experts, together with the peer-review the related academic production is subject to, provides the ground for a strong validation of the assumptions and the scientific relevance of the work.

The multiplicity of developments and applications described above, framed into the outlined governance structure, reach out to and engage with all of the following communities of practice:

- **Open software.** Model developers and experts in different programming languages contribute to the creation of new functionalities and related documentation, which are then reviewed within the community and submitted to peer-reviewed journals.
- **Higher Education.** Teaching staff, research staff and students employ the different available versions of the code and additional functionalities for developing applications, teaching material and scientific production.
- **Capacity building.** Officials from governments and international organisations, high level decision makers and technical staff from ministries are the beneficiaries of 'trainings of trainers' programs, where they are transferred the knowledge collected by the two other communities of practice and contribute to the creation of new knowledge.

First of all, this intertwined open-source research and application framework fulfils the scientific scope to provide transparent knowledge. Transparency is one of the pillars of research and development strategies in developing and developed countries and it lies not only in openness, but also in accessibility and stakeholder engagement.

More broadly, such framework meets the social scope to empower communities with the development of solutions for a better access to energy.

References

- [1] Abeeku Brew-Hammond, Energy access in Africa: Challenges ahead, *Energy Policy*. 38 (2010) 2291–2301. doi:<https://doi.org/10.1016/j.enpol.2009.12.016>.
- [2] United Nations Statistical Commission, United Nations Sustainable Development Knowledge Platform, (2015). <https://sustainabledevelopment.un.org/post2015/transformingourworld> (accessed July 3, 2017).
- [3] M. Howells, H. Rogner, N. Strachan, C. Heaps, H. Huntigton, S. Kypreos, A. Hughes, S. Silveira, J. DeCarolis, M. Bazilian, A. Roehrl, OSeMOSYS: the open source energy modeling system: an introduction to its ethos, structure and development, *Energy Policy*. 39 (2011) 5850–5870. doi:0.1016/j.enpol.2011.06.033.
- [4] Open Energy Modelling Initiative, OpenEnergy Platform, (n.d.). <http://oep.iks.cs.ovgu.de/> (accessed July 8, 2017).
- [5] Open Energy Modelling Initiative, Open Energy Modelling Initiative, (n.d.). https://wiki.openmod-initiative.org/wiki/Main_Page (accessed October 2, 2017).
- [6] Horizon 2020 LCE21 Projects, Energy Modelling Platform for Europe, (2017). <http://www.energymodellingplatform.eu/> (accessed January 3, 2018).
- [7] L. Mantzos, T. Wiesenthal, POTEnCIA model description: Version 9.0, 2016. doi:10.2791/416465.
- [8] L. Schrattenholzer, The energy supply model MESSAGE, Laxenburg, Austria, 1981. <http://adsabs.harvard.edu/abs/1981STIN...8225632S> (accessed January 11, 2016).
- [9] R. Loulou, U. Remme, A. Kanudia, A. Lehtila, G. Goldstein, Documentation for the TIMES Model Part I, 2005.

- [10] R. Loulou, U. Remme, A. Kanudia, A. Lehtila, G. Goldstein, Documentation for the TIMES Model Part II, 2005. <http://www.etsap.org/Docs/TIMESDoc-Details.pdf> (accessed January 11, 2016).
- [11] International Energy Agency, International Energy Workshop 2008, (2008). http://internationalenergyworkshop.org/old/Workshop_2008.html (accessed October 7, 2017).
- [12] C. Flood, Hydropower in Sweden: An investigation of the implications of adding detail to the modelling of hydropower in OSeMOSYS, KTH Royal Institute of Technology, 2015.
- [13] C. Maggi, Accounting for the long term impact of high renewable shares through energy system models : a novel formulation and case study, Politecnico di Milano, 2016. <https://www.politesi.polimi.it/handle/10589/125684>.
- [14] A. Beltramo, Active consumers at the centre of the energy system: Towards modelling consumer behaviour in OSeMOSYS, KTH Royal Institute of Technology, 2016.
- [15] H. Henke, The Open Source Energy Model Base for the European Union (OSEMBE), KTH Royal Institute of Technology, 2017.
- [16] I. Pappis, Electrified Africa – Associated investments and costs, KTH Royal Institute of Technology, 2016.
- [17] C. Sundin, Exploring the water-energy nexus in the Omo river basin: A first step toward the development of an integrated hydrological-OSeMOSYS energy model, KTH Royal Institute of Technology, 2017.
- [18] N. Moksnes, UN Sustainable development goals from a Climate Land Energy and Water perspective for Kenya, KTH Royal Institute of Technology, 2016.
- [19] J. Anjo, Development of the Portuguese Electricity System model in OSeMOSYS and study the impact of demand response implementation in the future., Instituto Superior Tecnico, 2017.
- [20] A. Tognollo, Energy demand analysis within long-term rural energy planning : a critical review and simulation for a more appropriate modelling approach, Politecnico di Milano, 2016.
- [21] Optimus.community, OSeMOSYS GitHub, (2017). <https://github.com/KTH-dESA/OSeMOSYS> (accessed October 3, 2017).
- [22] M. Welsch, M. Howells, M. Hesamzadeh, B. O Gallachoir, P. Deane, N. Strachan, M. Bazilian, D. Kammen, L. Jones, G. Strbac, H. Rogner, Supporting security and adequacy in future energy systems: The need to enhance long-term energy system models to better treat issues related to variability, *Int. J. Energy Res.* 39 (2015) 377–396. doi:10.1002/er.3250.
- [23] F. Gardumi, A multi-dimensional approach to the modelling of power plant flexibility, Politecnico di Milano, 2016.
- [24] M. Welsch, M. Howells, M. Bazilian, J. DeCarolis, S. Hermann, H. Rogner, Modelling elements of smart grids—enhancing the OSeMOSYS (open source energy modelling system) code, *Energy*. 46 (2012) 337–350. doi:10.1016/j.energy.2012.08.017.
- [25] IEA-ETSAP, Energy Supply Technologies Data, (2015). <http://109.73.233.125/~ieaetsap/index.php/energy-technology-data/energy-supply-technologies-data> (accessed January 11, 2016).
- [26] C. Taliotis, H. Rogner, S. Ressler, M. Howells, F. Gardumi, Natural gas in Cyprus: The need for consolidated planning, *Energy Policy*. 107 (2017) 197–209. doi:10.1016/J.ENPOL.2017.04.047.
- [27] F. Nerini, R. Dargaville, M. Howells, M. Bazilian, Estimating the cost of energy access: The case of the village of Suro Craic in Timor Leste, *Energy*. 79 (2015) 385–397. doi:10.1016/j.energy.2014.11.025.
- [28] C. Taliotis, A. Shivakumar, E.P. Ramos, M. Howells, D. Mentis, V. Sridharan, O. Broad, L. Mofor, An indicative analysis of investment opportunities in the African electricity supply sector — Using TEMBA (The Electricity Model Base for Africa), *Energy Sustain. Dev.* 31 (2016) 50–66. doi:10.1016/J.ESD.2015.12.001.
- [29] M. Howells, S. Hermann, M. Welsch, M. Bazilian, R. Segerström, T. Alfstad, D. Gielen, H. Rogner,

- G. Fischer, H. van Velthuizen, D. Wiberg, C. Young, R.A. Roehrl, A. Mueller, P. Steduto, I. Ramma, Integrated analysis of climate change, land-use, energy and water strategies, *Nat. Clim. Chang.* 3 (2013) 621–626. doi:10.1038/nclimate1789.
- [30] International Centre for Theoretical Physics, CLEWS Summer school, (2017). <http://indico.ictp.it/event/8008/overview> (accessed January 3, 2018).
- [31] United Nations Department of Economic and Social Affairs, Regional workshop: “Modelling Tools to Inform National Policies for Sustainable Development and the 2030 Agenda,” (2017). <https://www.un.org/development/desa/dpad/2017/regional-workshop-modelling-tools-to-inform-national-policies-for-sustainable-development-and-the-2030-agenda/> (accessed January 3, 2018).
- [32] K. Noble, OSeMOSYS: The Open Source Energy Modeling System-A translation into the General Algebraic Modeling System (GAMS), Stockholm, 2012. http://www.osemosys.org/uploads/1/8/5/0/18504136/osemosys_in_gams_wp_desa_2012.pdf.
- [33] K. Löffler, K. Hainsch, T. Burandt, P.-Y. Oei, C. Kemfert, C. von Hirschhausen, Designing a Model for the Global Energy System—GENeSYS-MOD: An Application of the Open-Source Energy Modeling System (OSeMOSYS), *Energies.* 10 (2017). doi:10.3390/en10101468.
- [34] M. Welsch, P. Deane, M. Howells, B. O Gallachoir, F. Rogan, M. Bazilian, Rogn, Incorporating flexibility requirements into long-term energy system models—A case study on high levels of renewable electricity penetration in Ireland, *Appl. Energy.* 135 (2014) 600–615. doi:10.1016/j.apenergy.2014.08.072.
- [35] L. de Strasser, A. Lipponen, M. Howells, S. Stec, C. Bréthaut, A Methodology to Assess the Water Energy Food Ecosystems Nexus in Transboundary River Basins, *Water.* 8 (2016) 59. doi:10.3390/w8020059.
- [36] T. Niet, GitHub - tniet/OSeMOSYS, (2017). <https://github.com/tniet/OSeMOSYS> (accessed October 7, 2017).
- [37] C. Taliotis, Large scale renewable energy deployment - Insights offered by long-term energy models from selected case studies, KTH Royal Institute of Technology, 2017.
- [38] Ministerio de Planificación del desarrollo Estado Plurinacional de Bolivia, Economic and social development plan 2016-2020 within the framework of integrated development for living well, 2016. http://www.planificacion.gob.bo/uploads/PDES_INGLES.pdf.
- [39] Ministry of Development Planning, Division of Economic and Policy Analysis, Intended Nationally Determined Contribution from the Plurinational State of Bolivia, n.d. <https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/post/INDC-Bolivia-english.pdf>.
- [40] R. Cervigni, R. Liden, J.E. Neumann, K.M. Strzepek, Enhancing the Climate Resilience of Africa’s Infrastructure: The Power and Water Sectors, World Bank, Washington DC, 2015.
- [41] G. Moura, Long-term power systems integration using SAMBA - South America Model Base - and the bargaining power of countries: a cooperative games approach, Universidade Federal do Rio de Janeiro, 2017.
- [42] G. Moura, M. Howells, L. Legey, SAMBA: the open source South American Model Base. A Brazilian perspective on long-term power systems investment and integration, Stockholm, 2015.
- [43] N. Moksnes, C. Taliotis, O. Broad, G. de Moura, M. Howells, A scenario elicitation methodology to identify the drivers of electricity infrastructure cost in South America, in: 19th EGU Gen. Assem. EGU2017, Proc. from Conf. Held 23-28 April. 2017 Vienna, Austria., p.14881, 2017: p. 14881. <http://adsabs.harvard.edu/abs/2017EGUGA..1914881M> (accessed October 3, 2017).
- [44] J. Sieber, WEAP Water Evaluation and Planning System, in: 3rd Int. Congr. Environ. Model. Softw., Brigham Young University ScholarsArchive, Burlington, Vermont, 2006. <http://scholarsarchive.byu.edu/iemssconference> (accessed October 3, 2017).

- [45] United Nations Economic Commission for Europe (UNECE), Reconciling resource uses in transboundary basins: assessment of the water-food-energy-ecosystem nexus, New York and Geneva, 2015.
- [46] United Nations Economic Commission for Europe (UNECE), Assessment of the water-food-energyecosystems nexus and benefits of transboundary cooperation in the Drina River Basin, New York & Geneva, 2017.
http://www.unece.org/fileadmin/DAM/env/water/publications/WAT_Nexus/ECE_MP.WAT_NON_E_9/Drina-EN-for_Web_final.pdf.
- [47] United Nations, Prototype Global Sustainable Development Report, New York, 2014.
- [48] REEEM Consortium - grant agreement No 691739, REEEM Horizon 2020 Project, (2016).
<http://www.reeem.org/> (accessed January 3, 2018).
- [49] Optimus.community, Optimus.community Reddit Forum, (2016).
<https://www.reddit.com/r/optimuscommunity/> (accessed October 3, 2017).
- [50] Optimus.community, OSeMOSYS Read the Docs, (2017).
<https://readthedocs.org/projects/osemosys/>.
- [51] Optimus.community, OpTIMUS.community, (2017). <http://www.optimus.community/> (accessed October 3, 2017).
- [52] National Renewable Energy Laboratory, Open EI, (n.d.). https://openei.org/wiki/Main_Page (accessed January 3, 2018).
- [53] Europa-Universität Flensburg, DIW Berlin, Technical University of Berlin, Neon Neue Energieökonomik, Open Power System Data, (n.d.). <https://open-power-system-data.org/> (accessed January 3, 2018).
- [54] United Nations Department of Economic and Social Affairs, Economy, energy and development in Bolivia, (n.d.). <https://www.un.org/development/desa/dpad/2016/economy-energy-and-development-in-bolivia/> (accessed October 3, 2017).
- [55] S&P Global Platts, World Electric Power Plants Database, (2016).
<http://www.platts.com/Products/worldelectricpowerplantsdatabase>.
- [56] Second National Development Plan (NDPII)-Vision 2040, (2015).
- [57] United Nations Department of Economic and Social Affairs Population Division, World Urbanization Prospects, (2014). <https://esa.un.org/unpd/wup/> (accessed October 3, 2017).
- [58] Uganda Bureau of Statistics, 2016 Statistical Abstract, Kampala, Uganda, 2016.
[http://www.ubos.org/onlinefiles/uploads/ubos/statistical_abstracts/2016 Statistical Abstract.pdf](http://www.ubos.org/onlinefiles/uploads/ubos/statistical_abstracts/2016%20Statistical%20Abstract.pdf).
- [59] Prototype Global Sustainable Development Report, United Nations, New York, 2014.
- [60] M. Welsch, Enhancing the Treatment of Systems Integration in Long-term Energy Models, KTH Royal Institute of Technology, 2013. <http://www.diva-portal.org/smash/record.jsf?pid=diva2:664679> (accessed January 11, 2016).
- [61] K. Ram, Git can facilitate greater reproducibility and increased transparency in science, *Source Code Biol. Med.* 8 (2013) 7. doi:10.1186/1751-0473-8-7.
- [62] K. Hunter, S. Sreepathi, J.F. DeCarolis, Modeling for insight using Tools for Energy Model Optimization and Analysis (Temoa), *Energy Econ.* 40 (2013) 339–349. doi:10.1016/J.ENERCO.2013.07.014.

Appendix A

The discounted-cost-minimisation objective function of OSeMOSYS is formulated in the three modelling languages as below. The differences are marked in bold characters.

- GAMS:
minimize cost: sum(YEAR, REGION) TotalDiscountedCost[y,r];
- GNU MathProg:
minimize cost: sum{**r in REGION, y in YEAR**} TotalDiscountedCost[r,y];
- Python:
def ObjectiveFunction_rule(model):
return sum(model.ModelPeriodCostByRegion[r] for r in model.REGION)
model.OBJ = Objective(rule=ObjectiveFunction_rule, sense=minimize)

Appendix B

B.1 – Sets

The ‘sets’ define the physical structure of a model, usually independent from the specific scenarios which will be run. They define the time domain and time split, the spatial coverage, the technologies and energy vectors to be considered, etc. For instance, when a variable is defined as a function of the set ‘YEAR’ it will be indicated as ***variablename*[y]** at it will be computed for every year listed in the set.

Name	Description	Index
YEAR	It represents the time frame of the model, it contains all the years to be considered in the study.	y
TECHNOLOGY	It includes any element of the energy system that changes a commodity from one form to another, uses it or supplies it. All system components are set up as a ‘technology’ in OSeMOSYS. As the model is an abstraction, the modeller is free to interpret the role of a technology at will, where relevant. It may for example represent a single real technology (such as a power plant) or can represent a heavily aggregated collection of technologies (such as the stock of several million light bulbs), or may even simply be a ‘dummy technology’, perhaps used for accounting purposes.	t
TIMESLICE	It represents the time split of each modelled year, therefore the time resolution of the model. Common to several energy systems modelling tools (incl. MESSAGE / MARKAL / TIMES), the annual demand is ‘sliced’ into representative fractions of the year. It is necessary to assess times of the year when demand is high separately from times when demand is low, for fuels that are expensive to store. In order to reduce the computation time, these ‘slices’ are often grouped. Thus, the annual demand may be split into aggregate seasons where demand levels are similar (such as ‘summer, winter and intermediate’). Those seasons may be subdivided into aggregate ‘day types’ (such as workdays and weekends), and the day further sub divided (such as into day and night) depending on the level of demand.	l
FUEL	It includes any energy vector, energy service or proxies entering or exiting technologies. These can be aggregate groups, individual flows or artificially separated, depending on the requirements of the analysis.	f
EMISSION	It includes any kind of emission potentially deriving from the operation of the defined technologies. Typical examples would include atmospheric emissions of greenhouse gasses, such as CO ₂ .	e
MODE_OF_OPERATION	It defines the number of modes of operation that the technologies can have. If a technology can have various input or output fuels and it can choose the mix (i.e. any linear combination) of these input or output fuels, each mix can be accounted as a separate mode of operation. For example, a CHP plant may produce heat in one mode of operation and electricity in another.	m

REGION	It sets the regions to be modelled, e.g. different countries. For each of them, the supply-demand balances for all the energy vectors are ensured, including trades with other regions. In some occasions it might be computationally more convenient to model different countries within the same region and differentiate them simply by creating ad hoc fuels and technologies for each of them.	r
SEASON	It gives indication (by successive numerical values) of how many seasons (e.g. winter, intermediate, summer) are accounted for and in which order. This set is needed if storage facilities are included in the model.	ls
DAYTYPE	It gives indication (by successive numerical values) of how many day types (e.g. workday, weekend) are accounted for and in which order. This set is needed if storage facilities are included in the model.	ld
DAILYTIMEBRACKET	It gives indication (by successive numerical values) of how many parts the day is split into (e.g. night, morning, afternoon, evening) and in which order these parts are sorted. This set is needed if storage facilities are included in the model.	lh
STORAGE	It includes storage facilities in the model.	s

B.2 – Parameters

The parameters are the user-defined numerical inputs to the model. While usually the structure of a model, therefore the sets, remains fixed across scenarios, it is common practice to change the values of some parameters when running different scenarios and/or sensitivity analyses. As will be clear in the following, each parameter is a function of the elements in one or more sets. For instance, **CapitalCost[r,t,y]** indicates that the capital cost is a function of the region (r), the technology (t) and the year (y)¹.

Global parameters	
YearSplit[l,y]	Duration of a modelled time slice, expressed as a fraction of the year. The sum of each entry over one modelled year should equal 1.
DiscountRate[r]	Region specific value for the discount rate, expressed in decimals (e.g. 0.05 to indicate 5%)
DaySplit[lh,y]	Length of one DailyTimeBracket in one specific day as a fraction of the year (e.g., when distinguishing between days and night: $12h/(24h*365d)$).
Conversionls[l,ls]	Binary parameter linking one TimeSlice to a certain Season. It has value 0 if the TimeSlice does not pertain to the specific season, 1 if it does.

¹ Please note that the order of the indexes of the parameters presented here and available in the current version of OSeMOSYS is different from the one in Howells et al. []. For more information refer to the Change Log of the OSeMOSYS versions on <http://www.osemosys.org/get-started.html>.

ConversionId[l,d,l]	Binary parameter linking one TimeSlice to a certain DayType. It has value 0 if the TimeSlice does not pertain to the specific DayType, 1 if it does.
Conversionlh[lh,l]	Binary parameter linking one TimeSlice to a certain DaylyTimeBracket. It has value 0 if the TimeSlice does not pertain to the specific DaylyTimeBracket, 1 if it does.
DaysInDayType[ls,ld,y]	Number of days for each day type, within one week (natural number, ranging from 1 to 7)
TradeRoute[r,rr,f,y]	Binary parameter defining the links between region r and region rr, to enable or disable trading of a specific commodity. It has value 1 when two regions are linked, 0 otherwise
DepreciationMethod[r]	Binary parameter defining the type of depreciation to be applied. It has value 1 for sinking fund depreciation, value 2 for straight-line depreciation.
Demands	
SpecifiedAnnualDemand[r,f,y]	Total specified demand for the year.
SpecifiedDemandProfile[r,f,l,y]	Annual fraction of energy-service or commodity demand that is required in each time slice. For each year, all the defined SpecifiedDemandProfile input values should sum up to 1.
AccumulatedAnnualDemand[r,f,y]	Accumulated Demand for a certain commodity in one specific year. It cannot be defined for a commodity if its SpecifiedAnnualDemand for the same year is already defined and vice versa.
Performance	
CapacityToActivityUnit[r,t]	Conversion factor expressing the energy that would be produced when one unit of capacity is fully used in one year.
CapacityFactor[r,t,l,y]	Capacity available each TimeSlice expressed as a fraction of the total installed capacity, ranging from 0 to 1. It gives the possibility to account for the forced outages.
AvailabilityFactor[r,t,y]	Maximum time a technology can run in the whole year, as a fraction of the year, ranging from 0 to 1. It gives the possibility to account for planned outages.
OperationalLife[r,t]	Useful lifetime of a technology, expressed in years.
ResidualCapacity[r,t,y]	Capacity available from before the modelling period.

InputActivityRatio[r,t,f,m,y]	Rate of fuel input (use) to a technology as a ratio of the rate of activity.
OutputActivityRatio[r,t,f,m,y]	Rate of fuel output of a technology as a ratio of the rate of activity.
Technology costs	
CapitalCost[r,t,y]	Capital investment cost of a technology, per unit of capacity.
VariableCost[r,t,m,y]	Cost of a technology for a given mode of operation (Variable O&M cost), per unit of activity.
FixedCost[r,t,y]	Fixed O&M cost of a technology, per unit of capacity.
Storage	
TechnologyToStorage[r,t,s,m]	Binary parameter linking a technology to the storage facility it charges. It has value 1 if the technology and the storage facility are linked, 0 otherwise.
TechnologyFromStorage[r,t,s,m]	Binary parameter linking a storage facility to the technology it feeds. It has value 1 if the technology and the storage facility are linked, 0 otherwise.
StorageLevelStart[r,s]	Level of storage at the beginning of the first modelled year, in units of activity.
StorageMaxChargeRate[r,s]	Maximum charging rate for a storage facility, in units of activity per year.
StorageMaxDischargeRate[r,s]	Maximum discharging rate for a storage facility, in units of activity per year.
MinStorageCharge[r,s,y]	Lower bound to the amount of energy stored, as a fraction of the maximum, ranging between 0 and 1. The storage facility cannot be emptied below this level.
OperationalLifeStorage[r,s]	Useful lifetime of a storage facility.
CapitalCostStorage[r,s,y]	Capital investment cost of a storage facility, per unit of energy.
ResidualStorageCapacity[r,s,y]	Storage capacity available from before the modelling period.
Capacity constraints	
CapacityOfOneTechnologyUnit[r,t,y]	Capacity of one new unit of a technology. In case the user sets this parameter, the related technology will be installed only in batches of the specified capacity and the problem will turn into a Mixed Integer Linear Problem.
TotalAnnualMaxCapacity[r,t,y]	Total maximum existing (residual plus cumulatively installed) capacity allowed for a technology in a specified year.

TotalAnnualMinCapacity[r,t,y]	Total minimum existing (residual plus cumulatively installed) capacity allowed for a technology in a specified year.
Investment constraints	
TotalAnnualMaxCapacityInvestment[r,t,y]	Maximum capacity of a technology allowed to be newly installed in the specified year, expressed in power units.
TotalAnnualMinCapacityInvestment[r,t,y]	Minimum capacity of a technology allowed to be newly installed in the specified year, expressed in power units.
Activity constraints	
TotalTechnologyAnnualActivityUpperLimit[r,t,y]	Total maximum level of activity allowed for a technology in one year.
TotalTechnologyAnnualActivityLowerLimit[r,t,y]	Total minimum level of activity allowed for a technology in one year.
TotalTechnologyModelPeriodActivityUpperLimit[r,t]	Total maximum level of activity allowed for a technology in the entire modelled period.
TotalTechnologyModelPeriodActivityLowerLimit[r,t]	Total minimum level of activity allowed for a technology in the entire modelled period.
Reserve margin	
ReserveMarginTagTechnology[r,t,y]	Binary parameter to tag the technologies which are allowed to contribute to the reserve margin. It has value 0 if a technology is not allowed, 1 if it is.
ReserveMarginTagFuel[r,f,y]	Binary parameter to tag the fuels to which the reserve margin applies. It has value 0 if the reserve margin does not apply to the fuel, 1 if it does.
ReserveMargin[r,y]	Minimum level of the reserve margin required to be provided for all the tagged commodities, by the tagged technologies. If no reserve margin is required, the parameter will have value 1; if, for instance, 20% reserve margin is required, the parameter will have value 1.2.
RE Generation target	
RETagTechnology[r,t,y]	Binary parameter to tag the renewable technologies which must contribute to reaching the indicated minimum renewable production target set in the model. It has value 1 for the tagged technologies, 0 otherwise.
RETagFuel[r,f,y]	Binary parameter to tag the fuels contributing to reaching the renewable target. It has value 1 for the tagged commodities, 0 otherwise.
REMinProductionTarget[r,y]	Minimum ratio of all renewable fuels tagged in the RETagFuel parameter, to be produced by the technologies tagged with the RETagTechnology parameter.

Emissions	
EmissionActivityRatio[r,t,e,m,y]	Emission factor of a technology per unit of activity, per mode of operation.
EmissionsPenalty[r,e,y]	Penalty per unit of emission.
AnnualExogenousEmission[r,e,y]	It allows the user to account for additional annual emissions, on top of those computed endogenously by the model (e.g. emissions generated outside the region).
AnnualEmissionLimit[r,e,y]	Annual upper limit for a specific emission generated in the whole modelled region.
ModelPeriodExogenousEmission[r,e]	It allows the user to account for additional emissions over the entire modelled period, on top of those computed endogenously by the model (e.g. generated outside the region).
ModelPeriodEmissionLimit[r,e]	Annual upper limit for a specific emission generated in the whole modelled region, over the entire modelled period.

B.3 – Variables

The variables are the outputs computed by the code. As much as the parameters, also the variables are functions of the elements in one or more sets.

Demands		Unit
RateOfDemand[r,l,f,y]>=0	Intermediate variable. It represents the energy that would be demanded in one time slice l if the latter lasted the whole year. It is derived from the parameters SpecifiedAnnualDemand and SpecifiedDemandProfile.	Energy (per year)
Demand[r,l,f,y]>=0	Demand for one fuel in one time slice.	Energy
Storage		
RateOfStorageCharge[r,s,ls,ld,lh,y]	Intermediate variable. It represents the commodity that would be charged to storage facility s in one time slice if the latter lasted the whole year. It is a function of the RateOfActivity and the parameter TechnologyToStorage.	Energy (per year)
RateOfStorageDischarge[r,s,ls,ld,lh,y]	Intermediate variable. It represents the commodity that would be discharged from storage facility s in one time slice if the latter lasted the whole year. It is a function of the RateOfActivity and the parameter TechnologyFromStorage.	Energy (per year)
NetChargeWithinYear[r,s,ls,ld,lh,y]	Net quantity of commodity charged to storage facility s in year y. It is a function of the RateOfStorageCharge and the	Energy

	RateOfStorageDischarge and it can be negative.	
NetChargeWithinDay[r,s,ls,ld,lh,y]	Net quantity of commodity charged to storage facility s in daytype ld. It is a function of the RateOfStorageCharge and the RateOfStorageDischarge and it can be negative.	Energy
StorageLevelYearStart[r,s,y]>=0	Level of stored commodity in storage facility s in the first instance of year y.	Energy
StorageLevelYearFinish[r,s,y]>=0	Level of stored commodity in storage facility s in the last instance time step of year y.	Energy
StorageLevelSeasonStart[r,s,ls,y]>=0	Level of stored commodity in storage facility s in the first instance of season ls.	Energy
StorageLevelDayTypeStart[r,s,ls,ld,y]>=0	Level of stored commodity in storage facility s in the first instance of daytype ld.	Energy
StorageLevelDayTypeFinish[r,s,ls,ld,y]>=0	Level of stored commodity in storage facility s in the last of daytype ld.	Energy
StorageLowerLimit[r,s,y]>=0	Minimum allowed level of stored commodity in storage facility s, as a function of the storage capacity and the user-defined MinStorageCharge ratio.	Energy
StorageUpperLimit[r,s,y]>=0	Maximum allowed level of stored commodity in storage facility s. It corresponds to the total existing capacity of storage facility s (summing newly installed and pre-existing capacities).	Energy
AccumulatedNewStorageCapacity[r,s,y]>=0	Cumulative capacity of newly installed storage from the beginning of the time domain to year y.	Energy
NewStorageCapacity[r,s,y]>=0	Capacity of newly installed storage in year y.	Energy
CapitalInvestmentStorage[r,s,y]>=0	Undiscounted investment in new capacity for storage facility s. Derived from the NewStorageCapacity and the parameter CapitalCostStorage.	Monetary units
DiscountedCapitalInvestmentStorage[r,s,y]>=0	Investment in new capacity for storage facility s, discounted through the parameter DiscountRate.	Monetary units
SalvageValueStorage[r,s,y]>=0	Salvage value of storage facility s in year y, as a function of the parameters OperationalLifeStorage and DepreciationMethod.	Monetary units
DiscountedSalvageValueStorage[r,s,y]>=0	Salvage value of storage facility s, discounted through the parameter DiscountRate.	Monetary units

TotalDiscountedStorageCost[r,s,y]>=0	Difference between the discounted capital investment in new storage facilities and the salvage value in year y.	Monetary units
Capacity variables		
NumberOfNewTechnologyUnits[r,t,y]>=0, integer	Number of newly installed units of technology t in year y, as a function of the parameter CapacityOfOneTechnologyUnit.	No unit
NewCapacity[r,t,y]>=0	Newly installed capacity of technology t in year y.	Power
AccumulatedNewCapacity[r,t,y]>=0	Cumulative newly installed capacity of technology t from the beginning of the time domain to year y.	Power
TotalCapacityAnnual[r,t,y]>=0	Total existing capacity of technology t in year y (sum of cumulative newly installed and pre-existing capacity).	Power
Activity variables		
RateOfActivity[r,l,t,m,y] >=0	Intermediate variable. It represents the activity of technology t in one mode of operation and in time slice l, were the latter to last the whole year.	Energy (per year)
RateOfTotalActivity[r,t,l,y] >=0	Sum of the RateOfActivity of a technology over the modes of operation.	Energy (per year)
TotalTechnologyAnnualActivity[r,t,y] >=0	Total annual activity of technology t.	Energy
TotalAnnualTechnologyActivityByMode[r,t,m,y] >=0	Annual activity of technology t in mode of operation m.	Energy
TotalTechnologyModelPeriodActivity[r,t]	um of the TotalTechnologyAnnualActivity over the years of the modelled period.	Energy
RateOfProductionByTechnologyByMode[r,l,t,m,f,y] >=0	Intermediate variable. It represents the quantity of fuel f which technology t would produce in one mode of operation and in time slice l, if the latter lasted the whole year. It is a function of the variable RateOfActivity and the parameter OutputActivityRatio.	Energy (per year)
RateOfProductionByTechnology[r,l,t,f,y] >=0	Sum of the RateOfProductionByTechnologyByMode over the modes of operation.	Energy (per year)
ProductionByTechnology[r,l,t,f,y] >=0	Production of fuel f by technology t in time slice l.	Energy
ProductionByTechnologyAnnual[r,t,f,y] >=0	Annual production of fuel f by technology t.	Energy
RateOfProduction[r,l,f,y] >=0	Sum of the RateOfProductionByTechnology over all the technologies.	Energy (per year)
Production[r,l,f,y] >=0	Total production of fuel f in time slice l. It is the sum of the	Energy

	ProductionByTechnology over all technologies.	
RateOfUseByTechnologyByMode[r,l,t,m,f,y] >=0	Intermediate variable. It represents the quantity of fuel f which technology t would use in one mode of operation and in time slice l, if the latter lasted the whole year. It is the function of the variable RateOfActivity and the parameter InputActivityRatio.	Energy (per year)
RateOfUseByTechnology[r,l,t,f,y] >=0	Sum of the RateOfUseByTechnologyByMode over the modes of operation.	Energy (per year)
UseByTechnologyAnnual[r,t,f,y] >=0	Annual use of fuel f by technology t.	Energy
RateOfUse[r,l,f,y] >=0	Sum of the RateOfUseByTechnology over all the technologies.	Energy (per year)
UseByTechnology[r,l,t,f,y] >=0	Use of fuel f by technology t in time slice l.	Energy
Use[r,l,f,y] >=0	Total use of fuel f in time slice l. It is the sum of the UseByTechnology over all technologies.	Energy
Trade[r,rr,l,f,y]	Quantity of fuel f traded between region r and rr in time slice l.	Energy
TradeAnnual[r,rr,f,y]	Annual quantity of fuel f traded between region r and rr. It is the sum of the variable Trade over all the time slices.	Energy
ProductionAnnual[r,f,y] >=0	Total annual production of fuel f. It is the sum of the variable Production over all technologies.	Energy
UseAnnual[r,f,y] >=0	Total annual use of fuel f. It is the sum of the variable Use over all technologies.	Energy
Costing variables		
CapitalInvestment[r,t,y] >=0	Undiscounted investment in new capacity of technology t. It is a function of the NewCapacity and the parameter CapitalCost.	Monetary units
DiscountedCapitalInvestment[r,t,y] >=0	Investment in new capacity of technology t, discounted through the parameter DiscountRate.	Monetary units
SalvageValue[r,t,y] >=0	Salvage value of technology t in year y, as a function of the parameters OperationalLife and DepreciationMethod.	Monetary units
DiscountedSalvageValue[r,t,y] >=0	Salvage value of technology t, discounted through the parameter DiscountRate.	Monetary units
OperatingCost[r,t,y] >=0	Undiscounted sum of the annual variable and fixed operating costs of technology t.	Monetary units

DiscountedOperatingCost[r,t,y] >=0	Annual OperatingCost of technology t, discounted through the parameter DiscountRate.	Monetary units
AnnualVariableOperatingCost[r,t,y] >=0	Annual variable operating cost of technology t. Derived from the TotalAnnualTechnologyActivityByMode and the parameter VariableCost.	Monetary units
AnnualFixedOperatingCost[r,t,y] >=0	Annual fixed operating cost of technology t. Derived from the TotalCapacityAnnual and the parameter FixedCost.	Monetary units
TotalDiscountedCostByTechnology[r,t,y] >=0	Difference between the sum of discounted operating cost / capital cost / emission penalties and the salvage value.	Monetary units
TotalDiscountedCost[r,y] >=0	Sum of the TotalDiscountedCostByTechnology over all the technologies.	Monetary units
ModelPeriodCostByRegion[r] >=0	Sum of the TotalDiscountedCost over all the modelled years.	Monetary units
Reserve margin		
TotalCapacityInReserveMargin[r,y] >=0	Total available capacity of the technologies required to provide reserve margin. It is derived from the TotalCapacityAnnual and the parameter ReserveMarginTagTechnology.	Energy
DemandNeedingReserveMargin[r,l,y] >=0	Quantity of produced fuel which is given a target of reserve margin. Derived from the RateOfProduction and the parameter ReserveMarginTagFuel.	Energy
RE Generation target		
TotalREProductionAnnual[r,y]	Annual production by all technologies tagged as renewable in the model. Derived from the ProductionByTechnologyAnnual and the parameter RETagTechnology.	Energy
RETotalProductionOfTargetFuelAnnual[r,y]	Annual production of fuels tagged as renewable in the model. Derived from the RateOfProduction and the parameter RETagFuel.	Energy
Emissions		
AnnualTechnologyEmissionByMode[r,t,e,m,y] >=0	Annual emission of agent e by technology t in mode of operation m. Derived from the RateOfActivity and the parameter EmissionActivityRatio.	Quantity of emission
AnnualTechnologyEmission[r,t,e,y] >=0	Sum of the AnnualTechnologyEmissionByMode over the modes of operation.	Quantity of emission

AnnualTechnologyEmissionPenaltyByEmission [r,t,e,y] >=0	Undiscounted annual cost of emission e by technology t. Product of the AnnualTechnologyEmission by the parameter EmissionPenalty.	Monetary units
AnnualTechnologyEmissionsPenalty[r,t,y] >=0	Total undiscounted annual cost of all emissions generated by technology t. Sum of the AnnualTechnologyEmissionPenaltyByEmission over all the emitted agents.	Monetary units
DiscountedTechnologyEmissionsPenalty[r,t,y] >=0	Annual cost of emissions by technology t, discounted through the DiscountRate.	Monetary units
AnnualEmissions[r,e,y] >=0	Sum of the AnnualTechnologyEmission over all technologies.	Quantity of emission
ModelPeriodEmissions[r,e] >=0	Total system emissions of agent e in the model period, accounting for both the emissions by technologies and the user defined ModelPeriodExogenousEmission.	Quantity of emission

Appendix C

Natural gas outlook in Cyprus

An energy system model of Cyprus, with spatial definition down to the individual power plant and year split in 63 time steps, was developed in collaboration with the Ministry of Energy, Commerce, Industry and Tourism and the Transmission System Operator. Presented in detail in [26], it is among the latest and most elaborated national applications of OSeMOSYS. It includes the additional block of functionality for computing the reserve capacity dispatch as proportional to the penetration of intermittent RES. It has been employed for providing insights regarding aspects of energy security under high renewable penetration and different possible implementation pathways for natural gas infrastructure.

Using four distinct gas availability scenarios that consider gas imports, domestic gas production and a mixture of the two, a comparison is made to identify cost-effective frameworks for natural gas introduction to the island. The impact of these on electricity cost, greenhouse gas emissions and renewable energy technology deployment are presented, while implications on energy security are discussed.

Meeting the INDCs in Bolivia

The United Nations Division of Economic and Social Affairs, UNDESA, implemented capacity building and advisory activities in Bolivia since 2006. The aim of these activities is to enhance the analytical capacity and transfer the knowledge of energy system modelling to government officials to address policy issues related to macro-micro-economics, energy and environment dimensions.

Within this framework, a model of the Bolivian energy system was developed in OSeMOSYS and further transferred to the Bolivian Government. A single-region, single-node model of the power generation sector of Bolivia was developed in OSeMOSYS linked with a detailed energy demand model in the Long-range Energy Alternatives Planning software (LEAP). Scenario analysis was used to answer research questions related to energy security, electricity exports and renewable energy deployment.

As a result of this project, in 2015, the Ministry of Development Planning and the Economic and Policy Analysis Unit (UDAPE) developed an analysis of the energy sector using the aforementioned model and submitted for the first time the Bolivia's Intended Nationally Determined Contribution (INDC) to the Conference of Parties, COP21, in Paris [39]. In April 2016, the conference "Economy, energy and development" was organized by DPAD and UDAPE to communicate the results of the last two years of capacity development projects in the country [54]. Government officials presented policy notes as results of simulations of energy and economy-wide models to address both contributions and limitations of key features of the Bolivia's Economic and Social Development Plan 2016-2020 [38]. The policy notes were prepared collaboratively with cross-ministerial teams and the topics addressed were: 1) Energy and economy-wide impacts of energy and infrastructure investments; 2) Energy efficiency; 3) Energy trade, fossil fuel reserves and electricity generation; and 4) Climate change effects on power generation.

Resilience of the African electricity infrastructure

A changing climate is expected to have an impact on the energy sector. This is especially the case for hydropower, as changes in precipitation patterns can affect the generation output from such facilities. Hydropower constitutes about 17% of Africa's electricity generation capacity [55], while a large number of additional projects are under construction or planned. In order to identify potentially vulnerable projects and design a robust system, the electricity supply system in seven river basins of the African continent is represented in OSeMOSYS and tested with a horizon to 2050 under a large number of climate scenarios [40]. To achieve this, models of each of the four interested Power Pools are developed, with separate country representation, while each hydropower plant is defined separately. To capture seasonal

and daily dynamics of electricity demand, renewable energy resource availability and, especially, hydropower output, each year was divided into 48 time steps.

Results from the scenarios indicate that the electricity generation infrastructure of certain African Power Pools with a generation mix consisting of diverse sources of generation are more resilient than the others. In addition, the impacts of an uncertain climate can be mitigated by investing in transboundary transmission infrastructure between countries where the probable climate-related risks are lower.

TEMBA (The Electricity Model Base for Africa)

TEMBA is an OSeMOSYS model of the electricity supply system of each of the 47 continental Africa's countries and the transmission links between them. Its country-level resolution and the high number of technologies modelled (more than 1000) allow insights to be obtained, on the optimal use of resources and shares between generation expansion and investment in electricity transmission in the various countries. The low time resolution (4 time slices, splitting the year in Day-Summer, Night-Summer, Day-Winter and Night-Winter) does not prevent averaged considerations on the optimal investment in infrastructure. The recently developed new version of OSeMOSYS reducing the translation time may allow the time resolution to be increased. The development of TEMBA was sponsored by World Bank and SIDA. It was presented at the Africa Pavilion at COP21, where its application for national energy planning was discussed by the attending government officials and policy makers.

SAMBA (South America Model Base)

The SAMBA model was developed in OSeMOSYS to represent the electricity supply sector of 11 South American countries [41,42]. Brazil is further divided into four regions, thus creating 14 separate systems. The initial purpose of the model was to examine the potential for and relationship between electricity investments and trade between countries in South America. The developed scenarios can identify countries with the largest potential for export of cost-competitive electricity, as well as the markets in which this electricity will be demanded. The region is rich in both fossil fuel reserves and renewable energy potential.

In a successive study, the SAMBA model was used to conduct a multi-dimensional scenario discovery [43]. Ranges of values for a number of input parameters of the model were identified and 324 scenarios were run. The aim of this exercise was to identify, with a bottom-up perspective, the parameters which the model is most sensitive to. The level of required investments and the cost of electricity were compared for each of these scenarios. It was concluded that the electricity demand growth rate, discount rate and a potential CO₂ emission limit were the greatest driving forces of the system cost.

Global Energy System Model (GENeSYS-MOD)

The Global Energy System Model (GENeSYS-MOD) [33] was developed to assess global decarbonisation pathways, while maintaining high level of disaggregation in the energy and emission analysis. GENeSYS-MOD employs the GAMS version of OSeMOSYS, significantly modified with additional blocks of functionality (e.g. 'Transportation', and 'Trade').

GENeSYS-MOD includes a multitude of supply and transformation technologies to satisfy the different demand needs. Multiple sectors, as well as sector-coupling technologies, are included. Its energy flows, technologies (symbolised by boxes) and demands (shaded boxes) are illustrated in Figure 6.

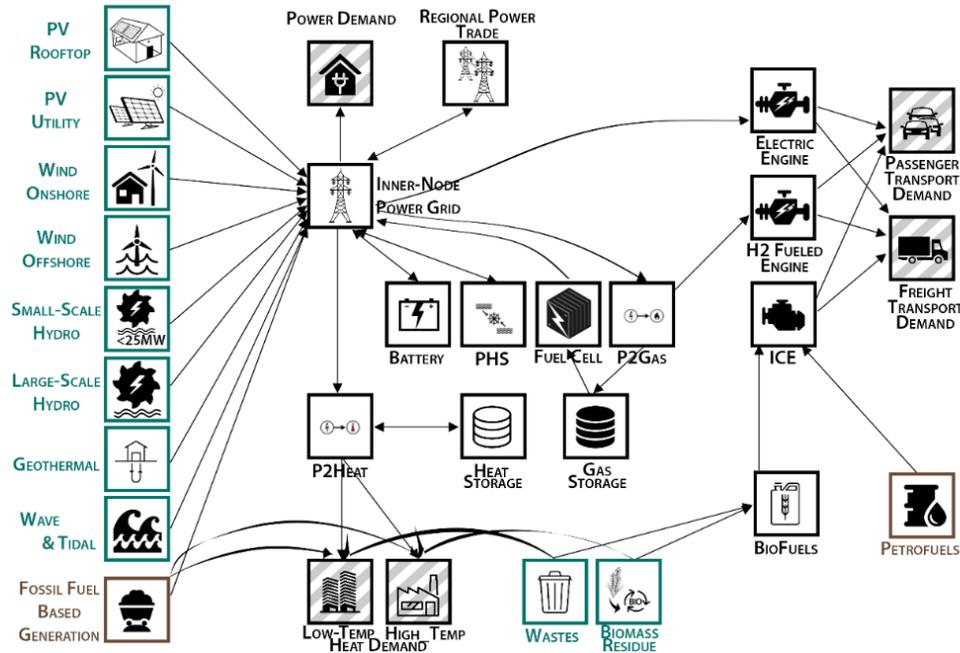


Figure 6. Reference Energy System of GENeSYS-MOD [33].

GENeSYS-MOD was used to analyse decarbonisation scenarios at the global level, broken down into ten regions. The primary goal was to find the cost-optimal energy mix with a global CO₂ emission target of 650 Gt from 2015 to 2050. The model results are driven mainly by the climate constraints and decreasing costs of renewable energy sources, as well as by the availability of cheap storage options [33]. As the carbon emission constraint becomes more binding, fewer fossil fuels are used to supply electricity and a gradual shift towards renewable sources is observed. Such new generation mix benefits also from reduced electricity consumption and new technological trends, such as the introduction of hydrogen in the transportation sector. The final energy mix in 2050 is based on 100% renewable energy sources, including wind, solar power, biomass, and hydropower.

GENeSYS-MOD will be further improved by increase of its temporal and spatial resolution. Case-studies targeting key regions of the global energy transformation such as Europe, China, India, and the USA are currently under development. All changes and additions to the source code, as well as all data, will be publicly available.

Water-energy-food Nexus in Uganda

National policies in the energy, water, land-use and agriculture sectors are often formulated in isolation [29], without taking into consideration the inter-linkages between other interacting systems and the implications of a changing climate. The CLEWs framework was initially developed to address this gap in the assessments: it could help to develop more inclusive policies, where the impacts of decisions in one sector on other co-existing systems can be evaluated. In a transboundary context, it fosters the dialogue between institutions operating in different sectors (such as separate Ministries) across different states or countries [35].

The agriculture sector in Uganda contributed to about 25.3% of the country's GDP in 2013 and provides employment for 72 % of labour force [56]. Most of the country's agricultural land is rain-fed and about 72 % of the farming is subsistence agriculture [56]. For an economy where the agricultural output is decisive and proportional to the availability of water for cultivation, seasonal precipitation patterns and

the overall climate play a crucial role. However, other sectors compete for the same water resource, namely: domestic household consumption and the electricity generation sector. With the population of Uganda growing at 4-5% [57] each year and about 90% of its electricity generated from hydroelectric power plants [58], water availability becomes even more critical.

To identify and quantify potential pressure points, an energy system model of Uganda developed in OSeMOSYS was soft-linked with a detailed hydrological water balance model developed in the Water Evaluation and Planning tool (WEAP). The soft link is developed in three phases: initially, a long-term energy system optimisation is performed with OSeMOSYS and the results concerning the capacity of hydropower plants installed along the time domain are fed to WEAP; in a second phase, a WEAP hydrological model is run with the inputs from OSeMOSYS and from climate scenarios derived from downscaled General Circulation Models, providing water balances for the domain of the study; finally, OSeMOSYS is run again including the water balances as constraints. Further details on the OSeMOSYS-WEAP modelling framework are given in [29].

Such composite modelling framework was employed to analyse the impact of a number of future climate scenarios. Initial results indicate a stronger competition for water in regions away from Victoria Nile River, especially in the North-Eastern states. The major hydropower plants based on the Nile are expected to have lesser climate related impacts compared to the small hydropower plants situated on the smaller tributaries of the Nile.

Transboundary water-energy resources assessments

The storage functionality of OSeMOSYS supported the creation of integrated assessment water-energy models for transboundary river basins in the Western Balkans and Central Asia, in the framework of the UNECE Water Convention.

The first case-study focuses on Sava River Basin and it studies the impact of climate change on the availability of water for hydropower generation and agricultural uses. The Sava River basin is shared between Slovenia, Croatia, Bosnia and Herzegovina, Serbia and Montenegro. The Sava River and its tributaries are essential for the national and regional energy security of the countries. The use of water and dependency from the basin varies across the riparian countries. An OSeMOSYS model of the region was developed and applied to two scenarios: these study the potential implications of Representative Climate Pathway 4.5 for hydropower generation and expansion, assuming two levels of water use for irrigation. The results quantify the decrease in the hydropower generation output in the region due to increased need for irrigation. The hydropower would be replaced by coal-fired generation (as coal resources are present in the region), resulting in turn in increased CO_{2,eq} emissions, as illustrated in Figure 7. The increase in fossil generation would be uneven between the countries in the basin and compensated by electricity trade.

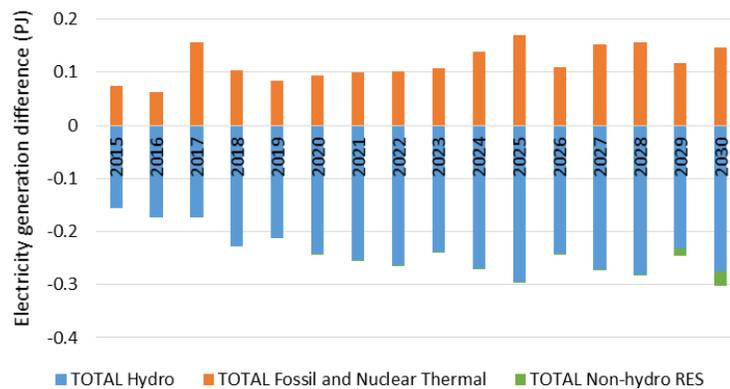


Figure 7. Selected results from the Sava River basin case-study: difference in electricity generation between the RCP4.5 and the maximum irrigation (IRR MAX) scenario in the Sava region.

A second study focused on Drina River Basin, a part of the larger Sava River Basin. This study assesses the benefits of transboundary cooperation in the operation of hydropower plants in the basin [46]. A more detailed OSeMOSYS model of the three countries sharing the basin (Bosnia and Herzegovina, Montenegro and Serbia) was developed to this end. The model represents the cascade of eight hydropower plants located along the Drina river and its main tributaries and it is structured in two parts: the electricity system and the hydrological system, both modelled in OSeMOSYS. The *electricity system* represents the whole power supply chain in each of the three riparian countries, from primary resources to final uses. The *hydrological system* is configured as a set of ‘technologies’ representing sections of the Drina River and its main tributaries and infrastructure dams. In this system, water is the commodity connecting the different sections. For each river, the upstream sections provide water to the downstream ones, according to seasonal availability. When a dam is present in a certain point of the river, it is allowed to fill up the reservoir at any time, taking into account its volume, the seasonal availability of water and the operation constraints of the hydro power plant. The mass balances of water along the rivers and within the dams constrain the model, too.

The insights drawn from this analysis show that improved cooperation has the potential to increase electricity generation in the hydropower plants downstream without compromising the generation upstream. Moreover, it demonstrates the role of cheap hydropower in enhancing electricity trade among the riparian countries and with other neighbouring countries.

The third study looks at Syr Darya River basin. This is shared between Kyrgyzstan, Tajikistan, Kazakhstan, and Uzbekistan. Different interests rule the management of water resources in this basin located in Central Asia, which, along with the Amu Darya river catchment, drains to the Aral Sea. The overexploitation of water resources for the cultivation of cotton is leading to the shrinking and drying up of the sea. Water resources play a different role in different countries of the region: they are the backbone of the electricity generation for the upstream nations, while they are mainly used for irrigation in the downstream nations. These two uses are incompatible: upstream nations need to discharge water from the reservoirs to produce electricity in winter, but this practice hinders water storage for irrigation in the summer. The study investigates options to decrease the dependency on water resources for hydropower generation, by implementing selected energy efficiency measures or increasing the share of non-hydro renewables in the electricity system. As shown in Figure 8, the results provide evidence that the demand for hydroelectricity could be reduced if energy efficiency measures were implemented across the four states or non-hydro renewable energy technologies were exploited.

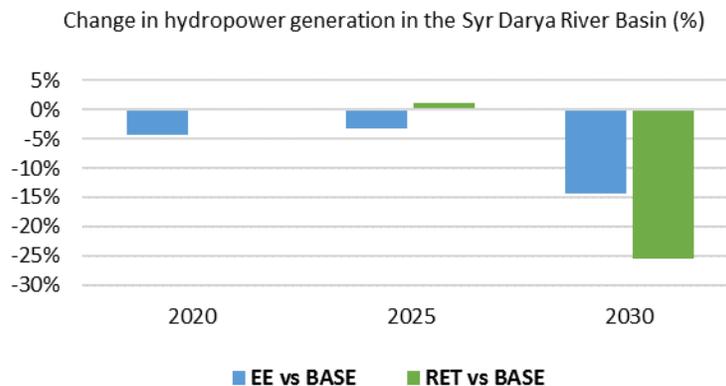


Figure 8. Selected results from Syr Darya River basin case-study: change in hydropower generation in the Energy Efficiency (EE) and Renewable Energy Technologies (RET) scenarios in comparison to the reference case (BASE).

In all the three case-studies presented above, stakeholders were involved during the assessment and their experience and context-based knowledge used to refine the investigation. The dialogue with stakeholders and local experts provided an opportunity to adjust the study to the needs of the country and/or region, increasing its potential impact.

GLUCOSE: a global CLEWs model

To support transparent decision-making processes, an open-source global CLEWs model has been developed. The model attempts a simplistic, globally aggregated approach and includes the six most resource-intensive material industries. From the model results, indicators for sustainable resource strategies can be deduced. These bring forward the discussion on a global resource outlook in relation to social, economic and environmental concerns. Indicative results show that the CLEWs approach reaches significantly different results from single sector modules (Figure 9). Despite its aggregation, the global CLEWs model can provide insights into resource interconnections that are difficult to extract from single-sector models.

As constraints are added to the base model, the optimisation process gradually focuses on the technologies still available. The introduced constraints are a Greenhouse Gas tax, a land cost (related to CO₂ emissions from land use change), a limitation of agricultural land to today's area and three levels of global emission cap. As the constraint parameters are introduced, the energy system shifts towards more gas, hydropower and biomass to reduce carbon emissions. When the total agricultural land is kept constant, the biomass contribution is strongly reduced. Due to the limitation imposed on the rate of investment on renewable technologies, the system becomes more heavily dependent on nuclear as well as on more efficient fossil fuel technologies (i.e. combined cycle gas turbines). On the other hand, Carbon Capture and Storage (CCS) technology is not invested in in any of the scenarios. A more elaborate description of the model and its insights are available in the Prototype Global Sustainable Development Report of the United Nations [59].

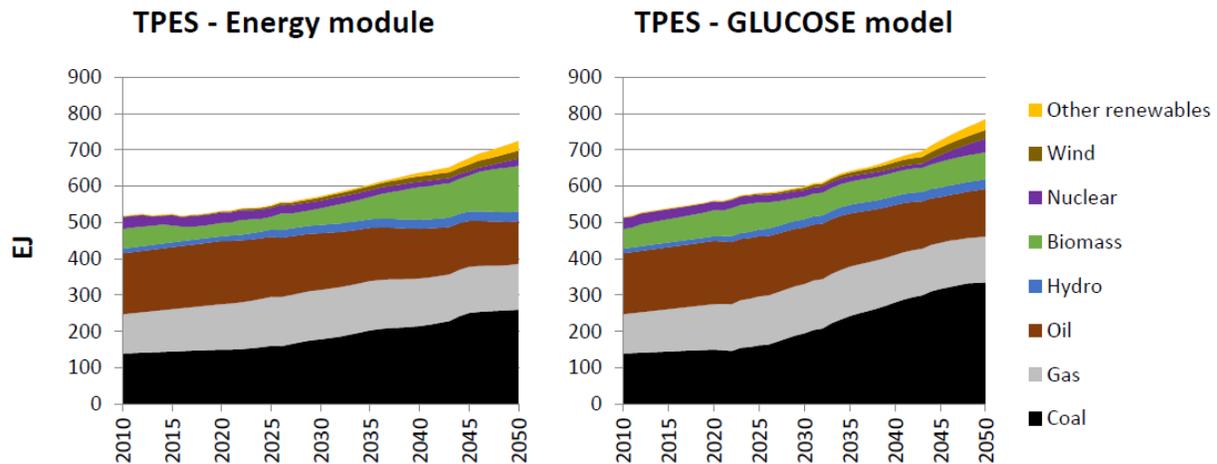


Figure 9 - Total Primary Energy Supply in the baseline scenario of the individual energy module (left) and the GLUCOSE model (right).

Appendix D

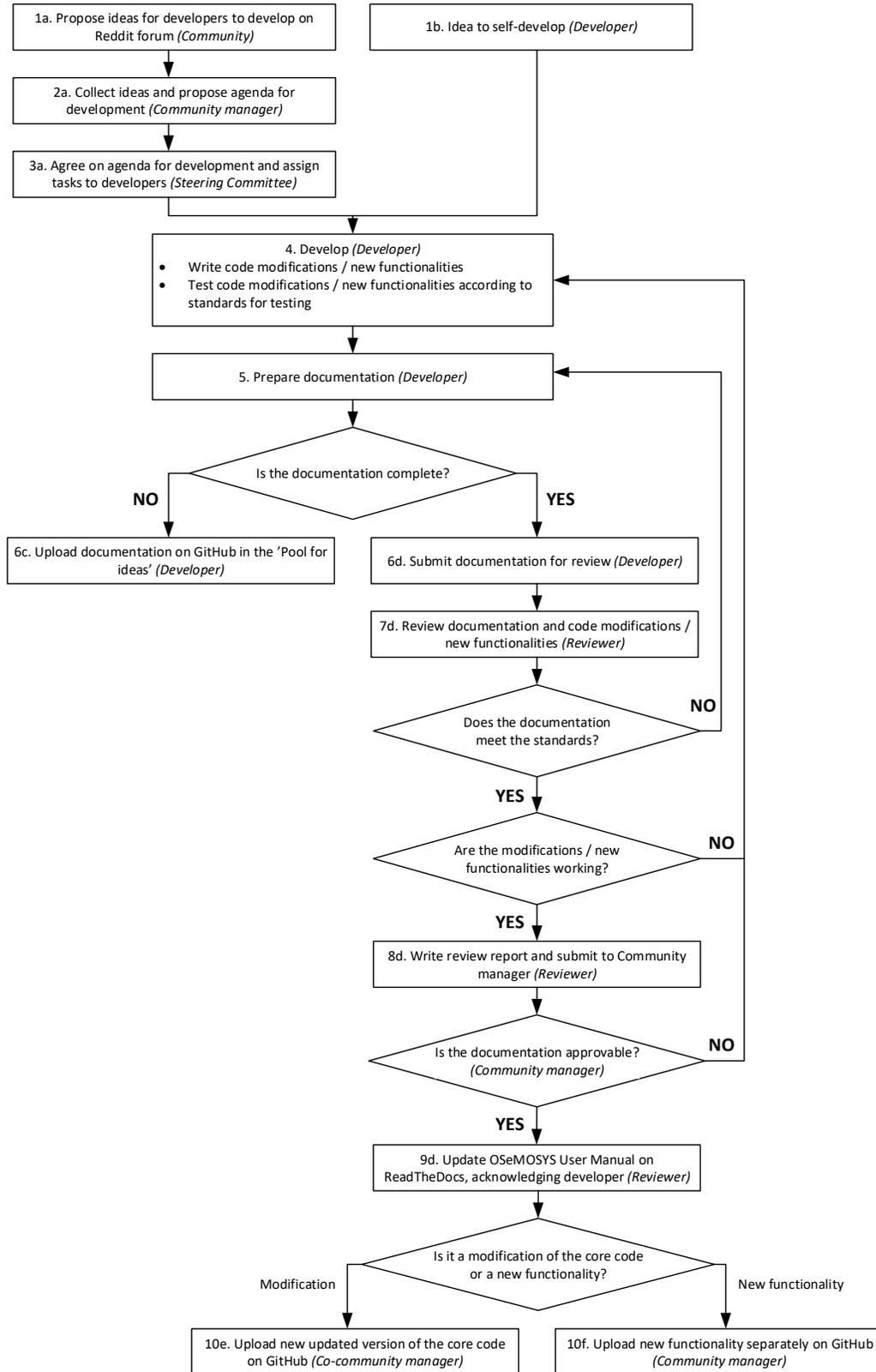


Figure 5. Flow diagram of the process to propose and review functionalities in OSeMOSYS.

1a. Propose ideas for developers to develop on Reddit forum. The process for modifying OSeMOSYS or developing new functionalities starts from the needs of the Community. Those who apply OSeMOSYS but are not coding experts themselves may identify the need to modify the core code of OSeMOSYS due to the presence of bugs or the need for a new specific functionality (e.g. to account for storage losses or for the costs of intermittent generation). In this case, they are given the possibility to propose ideas in a dedicated thread on a Reddit forum [49].

1b. Ideas to self-develop. In parallel to the case considered in 1a, some users who identify the need for modifications to the core code or new functionalities may have the capability to develop them themselves. In such case, they can directly work on the modifications, without proposing them to the Community. This has already occurred a few times since the publication of OSeMOSYS, especially as a part of Master or Doctoral research works [13,23,37,60].

2a. Collect ideas and propose agenda for development. The Community manager shall periodically check the dedicated thread on the Reddit forum, collect all the ideas which nobody has agreed to develop, select those of interest, prioritise them and propose them to the Steering Committee.

3a. Agree on agenda for development and assign tasks to developers. The Steering Committee comments, iterates and finally agrees on the development agenda. Then, the ideas to be developed are shared among the Steering Committee members according to the interests and the available resources.

4. Develop. This step is carried out by the developer offline. It consists in two parts: writing the equations (first in an algebraic form, then in the GNU MathProg, GAMS or Python modelling language) and then testing them. The latter phase requires the creation of a test case-study, defined to highlight how the results change and the ensuing benefits. The test case-study should ideally be derived from a standard case-study provided to all the Community by the Steering Committee.

5. Prepare documentation. A standard set of documentation must be provided by the developer, for his/her work to be accessible, transparent and reproducible. This includes:

- In the case of a modification to the core code of OSeMOSYS, both the original and the modified version of the code;
- In the case of a new functionality, the code formulation of all the new equations;
- A readme file containing: the OSeMOSYS version the modification is compatible with, the English description of each equation, the algebraic formulation of each equation and the code formulation of each equation;
- The test case-study and a description;
- The result files from the application of the test case-study to the original and the modified version of OSeMOSYS. Standard formats are provided for these by the OSeMOSYS Steering Committee.

6c. Upload documentation to GitHub in the 'Pool of ideas'. If, for any reason, the developer cannot provide all the required documentation, he/she has an option of sharing what is available in a space called 'Pool of ideas'. From there, the community can pick contributions and use them or elaborate on them under the terms of the license covering them. The Steering Committee may decide to further develop some of the contributions, as they see fit, and provide the necessary documentation. However, no guarantee of perfect functionality shall be given for what uploaded just in the pool of ideas.

6d. Submit documentation for review. When the developer can provide all the required documentation, he/she may proceed by submitting it for review on GitHub. To this end, the developer shall be assigned a git development branch, separated from the master branch where the reviewed versions of OSeMOSYS are. In this way the developer will be able to work asynchronously, yet being able to include latest changes to the OSeMOSYS version in the master branch at any time. This practice, state of the art in large open-source collaborative efforts [61,62], provides an effective and resource efficient way to coordinate efforts in the community, maintain the authorship trail for all contributions, compare version and retrace errors.

7d. Review documentation and code modifications / new functionalities. This step is carried out by reviewers, who are experts nominated by the Community manager and chosen from a list put together by the Steering Committee. The submission undergoes two types of checks: the first relates to the completeness of the documentation provided by the developer, according to the guidelines, and the second relates to the correct functioning of the code modifications.

8d. Write review report and submit to Community manager. If both revision checks are passed, the reviewers shall write a review report according to a standard similar to the usual one in the peer-review processes of scientific papers.

9d. Update OSeMOSYS User Manual on ReadTheDocs, acknowledging the author. After final control and agreement by the Community manager, the reviewer can proceed to updating the user manual of OSeMOSYS. The manual is available online on ReadTheDocs and modifiable by authorised users. In order to make the process as resource-efficient as possible, the update shall be in the form of adding a paragraph to the properly structured manual, mostly taken from the review report.

10e. Upload new updated version of the core code on Github. When the contribution consists of modifications to the core code of OSeMOSYS, a new updated version must be uploaded to GitHub. This requires knowledge of the modelling language the modification is written with, namely GNU MathProg, GAMS or Python. For this reason, this step is carried out by a co-Community manager, a person with a specific expertise in the relevant language.

10f. Upload new functionality separately on Github. When the contribution is a new functionality working as a plugin to the core code, it is sufficient to upload it to GitHub separately together with its documentation, for users to plug it into OSeMOSYS when needed. This step is carried out in any case by the Community manager.