Improvement of model-based energy systems analysis through systematic model experiments

Hans Christian Gils[a,b,*], Jochen Linßen[c], Dominik Möst[d], Christoph Weber[e]

[a] *German Aerospace Center (DLR), Institute of Networked Energy Systems, Curiestr. 4, 70563 Stuttgart, Germany*
[b] *Stuttgart Research Initiative on Integrated Systems Analysis for Energy (STRise), Keplerstraße 7, 70174 Stuttgart, Germany*
[c] *Jülich Research Center – Institute of Technoeconomic Systems Analysis (IEK-3), Wilhem-Johnen Straße, 52425 Jülich, Germany*
[d] *TU Dresden, Chair of Energy Economics, Münchner Platz 3, 01062 Dresden, Germany*
[e] *University of Duisburg-Essen, Chair for Management Science and Energy Economics (EWL), Universitätsstr. 11, 45117 Essen, Germany*

Abstract

This Virtual Special Issue collects the results of six research projects conducted under the umbrella of the MODEX network. MODEX stands for Model Experiments and refers to the combination of qualitative and quantitative model comparisons, the latter based on a harmonized model application. While all projects focused on the comparison of energy system models, the experiments covered a broad range of different areas. The experiments address different model types, such as vehicle diffusion models, power system models, distribution grid models, and transmission grid models, but also different thematic scopes, such as policy instruments, decentralized flexibility, and sector coupling. In addition, the methods and tools for the systematic comparison of energy system models were further developed in all projects. Another essential aspect of all projects was the harmonization of models and data, which was realized to the degree possible and useful in the respective project context. In total, the experiments performed within MODEX included 40 energy system models operated by 39 institutions. The results of the MODEX projects significantly expand the spectrum of systematic comparisons of energy system models. The added value compared to previous model experiments consists particularly in the profound characterization of the models involved, the consistent linking of model differences with outcome differences, and the transparent description of appropriate methods and challenges of comparing energy system models. Due to a high degree of transparency, the project findings can be used directly by developers and users of other models.

Keywords: model comparison, energy system modeling, model experiments,

*Corresponding author
Email address: hans-christian.gils@dlr.de (Hans Christian Gils)

energy systems analysis

Highlights

- MODEX is a milestone for structured model experiments of energy system models
- Model experiments enhance methodological knowledge of models and their comparison
- Model understanding and data harmonization are key challenges in model experiments
- Despite some similarities, all models examined have dedicated strengths

List of abbreviations

BEV battery electric vehicle

CHP combined heat and power

CO_2 carbon dioxide

HP heat pump

PV photovoltaics

VRE variable renewable energy

1. Introduction

1.1. Background and motivation

As the effects of climate change become increasingly tangible and the will to take political and societal action for climate protection grows, research into future energy supply systems and the development of strategies for transforming existing systems are becoming ever more important. In the scientific as well as in the business environment, complex energy system models of different focus are used. Especially in the last 10 years, the number, variety and complexity of such models has developed very dynamically. Even with similar research scope, energy system models often show very different modeling approaches, model formulations and model features. This, together with different model assumptions, results in sometimes highly divergent model results and thus policy recommendations. How the manifold differences in data, methods, system boundaries affect the results can only be determined by comprehensive model experiments. In this respect, a model experiment is the combination of a qualitative model comparison and the analysis of quantitative indicators resulting from a harmonized application of the models.

1.2. State of research

The literature offers a large number of qualitative model comparisons. There, sometimes more than a hundred models are compared with respect to their scope and properties. Among the most recent of these comparisons are, for example, those by Prina et al. [1], Fattahi et al. [2], and Klemm et al. [3]. In these as well as in comparable works, however, no application of the models is performed. As a result, the effect of model differences remains unclear. In contrast, other works focus explicitly on the comparative application of models within clearly defined experiments. Among the earlier such analyses is the FORUM project, in which differently oriented comparisons of models frequently used in Germany were made. The series of model experiments focused on the economic impact of climate protection policies [4], the implications of a nuclear phase-out [5], the contribution of the German energy industry to the achievement of European emission reduction goals [6], renewable energy technologies in liberalized energy markets [7], and effects of innovation on the energy system and economic development [8]. The Energy Modeling Forum Stanford (EMF) looks back on 37 model experiments in the meantime. These also focused on different topics, but also on different countries. Most recently, the Japanese energy system [9] and the North American energy trade and integration [10, 11, 12] were examined. The model experiments performed in FORUM and the EMF have mostly focused on spatially and temporally aggregated energy system and overall economic models, whereas detailed grid or sector models have not been considered. Furthermore, mostly only a few framework assumptions and story-lines but no complete models and input data sets were harmonized.

Instead, two recent model experiment projects focusing on the German power system considered a much more comprehensive harmonization of data and models. They particularly focused on the effect of different modeling approaches [13] and the consideration of sector coupling [14], respectively. Still, these works only cover a small range of important aspects in comparing power and energy system models. Furthermore, they do not address the roles of open source models and open data.

1.3. Formation of the MODEX project network

The initiative to establish the MODEX project network origins in the *Forschungsnetzwerk Energiesystemanalyse*, a network of experts in energy systems analysis established and funded by the German Federal Ministry for Economic Affairs and Climate Action [15]. The scientists expected this to lead to a better understanding of the differences and similarities between the energy system models used in Germany, which had recently grown considerably in number and diversity. This initiative was taken up by the Ministry and used as the basis for a call for research funding. The idea was to bring together thematically different consortia to cooperate with each other in model comparison projects under the umbrella of the MODEX network. Given the large number of models and the current developments, the projects should have different methodological focus. An important aspect was the systematic evaluation of differences in the modeling approaches as well as a comparison of the strengths and weaknesses of different methods. In doing so, the assessment of the informative value of certain methods for specific research questions should be sharpened and potential for improvement should be identified.

In preparation for the project definition, scientists in Germany involved in the modeling of energy systems were called upon to enter their models in the database of the Open Energy Platform (OEP) [16]. For this purpose, standardized model templates were used, which were based on those from previous model experiments, including [14], [17], and [13]. The collected information on all models was later subjected to a qualitative comparison, which was published in [18].

Based on the overview of existing models on the OEP, the research community was invited to form consortia and define model experiments. Six projects of different focus were finally selected for funding within the framework of the MODEX project network. The project open_MODEX, for example, was dedicated to the comparison of open models, whereas the project MODEX-Net focused on power transmission models. The MODEX-POLINS project, meanwhile, compared models with a focus on policy instruments, while the FlexMex project concentrated on power system flexibility and sector coupling. The MODEX-EnSAVes project not only compared differently oriented models in various subgroups, but also partially coupled them in the analysis of decentralized flexibility. The MEO project was the only one not dedicated to large-scale systems, but to operational energy system analysis at the distribution grid level. All in all, the MODEX projects comprised 40 models from 39 institutions.

The Virtual Special Issue "MODEX: energy system model comparisons through harmonized applications" collects the key results and findings of the six projects. The following Section 2 provides a more detailed insight into the projects and their results. Building on this, a synthesis of all projects is presented in Section 3. The results of a survey conducted among all modelers involved in the MODEX projects on their most important experiences are presented and discussed in Section 4. Finally, the scientific contribution of the MODEX projects is placed in the context of previous model experiments in Section 5.

2. The MODEX model experiments

Based on the articles published in the Special Issue, this section summarizes the core findings of the six MODEX projects.

2.1. MODEX-EnSAVes

In the MODEX-EnSAVes project, various methodical approaches for analyzing the energy system were applied and compared using a specific use case. The focus of the application was on issues relating to sector coupling to compare the market ramp-up of new electricity applications in the various models. The researchers investigated how critical situations (e.g., "cold dark doldrums") can be managed in the future by coupling system-analytical models for demand development with electricity market models. The latter market models analyzed generation adequacy in future power systems and whether the generation and other backup capacities are sufficient to cope with periods of high residual load. The team derived expected future load profiles for the new electricity applications from the demand-side models that served as input for the electricity market models. The general project focused on battery electric vehicles (BEVs) and heat pumps (HPs) in residential buildings. Various influencing factors have been reflected differently in the individual model approaches to consider the decision-making of different actors. The consortium applied various sector-specific and detailed models and coupled the modeling approaches within the modeling experiment. Two papers prepared within the MODEX-EnSAVes project have been published in this Special Issue. The first paper provides a systematic comparison of four electricity market models based on detailed and harmonized scenario input parameters [19]. Selected output parameters, such as investment and dispatch decisions for flexible power plants, storage dispatch, wholesale electricity prices, carbon dioxide (CO_2) emissions, and generation adequacy in hours with critical supply have been compared for the German energy system until 2030. The experiment shows that the results of the various models differ only slightly from each other. Differences in the results are traced back to conceptual differences, as the models can be distinguished concerning their mathematical approaches and level of detail. The second paper focused on how the decision of buying a car – which is not entirely objective and is only partly based on costs – could be considered in models [20]. The paper analyzes the integration of user behavior into three market diffusion models and compares these models with a focus on three steps: first, it compares the modeling approaches, then uses a harmonized data set to model the future market diffusion of alternative fuel vehicles, and third compares results with and without behavioral aspects. Behavioral aspects included in the three models are the use of charging infrastructure, the limited vehicle availability, the consideration of range anxiety as a hampering factor, the willingness-to-pay-more for alternative drive-trains as a supporting factor, and a distinction of users' driving distances.

2.2. MODEX-Net

The main objective of the MODEX-Net project was the comparison of existing transmission grid models with focus on Germany and Europe. Based on defined model experiments with harmonized data, the project partners identified and described similarities and differences in terms of methodology, structure, and data. Based on this, they derived proposals on how transmission grid models can be further improved. Within the MODEX-Net project, an open framework to compare a variety of transmission grid models focusing on the German and pan-European power system was developed [21]. The comparison is performed in both a qualitative and quantitative manner, depending on the investigated modeling aspect including input data, methods, system boundaries, and results. The core elements of the comparison framework include fact sheets for qualitative comparison, harmonization of input data such that differences are reduced without losing core aspects of the models, indicators and visualization tools to quantify the differences between the models and identify similarity groups. Results and methods for this framework are made publicly available. Another paper focuses on the time series of power demand and variable renewable energy (VRE) feed-in required for transmission grid modeling in high spatial resolution [22]. It compares top-down, bottom-up and mixed approaches that are being used to aggregate or disaggregate input data. Furthermore, it evaluates various approaches to assign regionalized profiles to the model's grid connection points. The authors find that the variety of approaches leads to significant differences on a regional scope, even if global values are the same. They develop a methodology allowing for a comparison of different regionalization and grid node assignment approaches using simple parameters and without explicit knowledge of grid topology. The results show that the input data resolution and the use of a top-down or a bottom-up approach are the most determinant factors in the regionalization process. Additional analyses aimed to identify the main differences that exist regarding grid simulation models for Germany and methodologies for studying congestion management in the European context [23]. The effects of model parametrization and formulation on congestion management results are investigated based on three case studies focusing on outage simulation, line constraint relaxation and the modeling of cross-border measures. Results indicate that data parametrization can have significant impacts on model results with respect to volumes and geographic distribution of congestion management measures. The analysis also highlights the need to thoroughly calibrate key model parameters. The research findings assist the grid modeling community and power system planners in simulating congestion management and increase the validity and informative value of grid simulation models.

2.3. FlexMex

FlexMex focused on the use of flexibility options to balance VRE power generation. The central question was how different approaches of optimization and technology modeling affect the system dispatch in hourly resolved power sector models with sector coupling. In addition, the effect of different model scopes on the deployment of flexibility options was investigated. To consistently separate data-related from model-related differences in results, the input data of the nine participating models were fully harmonized. The application of the models was then divided into two main experiments. In the first experiment, individual flexibility options were investigated based on a comprehensive qualitative analysis of the models and their differences [24]. Considering highly reduced test cases, it was possible to isolate and quantify model-specific effects. Substantial differences in technology modeling were identified primarily for BEVs, reservoir hydro power, power transmission, and demand response. Furthermore, the paper finds that the consideration of simplified test cases has high potential for the evaluation of individual differences in technology modeling and model features, whereas it is less suited for the analysis of different optimization approaches. Complementary analyses dealt with the model endogenous expansion of electricity storage, power transmission lines and controllable power plants in reduced test cases [25]. These identify deviating assumptions regarding technology availability factors, consideration of predefined levels and energy-to-power-ratios of storage, and representation of power transmission as most influential technology modeling differences. In addition, it is shown that a fixed dispatch order leads to substantially higher capacity investments compared to perfect foresight linear optimization. Building on the technology-specific analyses, more complex scenarios were considered in the second model experiment [26]. There, sixteen stylized scenarios differing in VRE supply share, technology scope, and optimization scope were considered. Despite the high number of models and interacting model differences, result deviations can still be related to model characteristics. The paper highlights that differences in modeling approach and the representation of specific technologies lead to comparatively small deviations, whereas a heterogeneous model scope can have a much more substantial impact. This particularly applies to differences in the consideration of sector coupling. Still, both the exploitation of sector coupling flexibility and the trends in the use of individual technologies are mostly robust across the models. In summary, the FlexMex project results can provide a better understanding of the effect of different modeling approaches and thus contribute to the interpretation of model results. The input and result data [27], as well as the evaluation tools of the project [28] were also published.

2.4. MODEX-POLINS

In the MODEX-POLINS project, the assessment of the impacts of policy instruments was in the focus. Given that energy models are also widely used to evaluate the effects of various policy instruments, the ambition has been to provide a systematic investigation of similarities and differences in model outcomes when energy policies are investigated in a realistic setting. The models participating in this experiment have been harmonized regarding key input parameters like fuel prices, existing power plant capacities or VRE feed-in time series, yet a core issue has been to investigate whether differences in the modeling approaches lead to differences in the policy impacts. In order to separate the policy impacts from general model differences, both baseline developments and developments under modified policy instruments have been investigated. In the MODEX Special Issue, results regarding two types of policy instruments are reported: Ruhnau et al. [29] focus on carbon pricing while Bucksteeg et al. [30] investigate different policies regarding integrated electricity and heat systems, i.e., district heating systems with combined heat and power (CHP) plants. In both papers, there are substantial variations in policy impacts across the models. As one major driver for differences in results, the treatment of investments has been investigated. Pure dispatch models, where the power generation fleet is given exogenously, tend to show lower impacts than models with endogenous investments. In the case of high carbon prices, [29] notably identify market-based investments in VRE technologies as a major source of higher emission reductions. As another key driver for deviations between model outcomes, the modeling of CHP plants and the corresponding district heating grids has emerged. An aggregated modeling of these generation technologies tends to overstate the flexible response of energy systems to changes in the policy instruments. For the future development of the integrated electricity and heat systems themselves, the modeling of power-to-heat technologies and of the corresponding regulations has turned out to be an important aspect.

2.5. MEO

MEO stands for Model Experiments in Operational energy system analysis. Correspondingly, the project focused on the modeling of detailed system operations. In the paper [31], an exemplary low voltage grid with high photovoltaics (PV) penetration is considered. In one scenario, line expansion is analyzed, while the other scenario focuses on the installation of a voltage regulation distribution transformer (VRDT). Other case studies not included in the Special Issue assessed increased penetration of HPs, BEVs or small-size CHP units in the same grid. The focus of the project has been on investigating both the solution approaches and the capabilities of the different models as well as on comparing the computational results. Four models with different backgrounds participated in the reported model experiment. These include power system analysis tools which include automated grid planning approaches along with various types of power flow calculations. But also a more reliability oriented analysis tool is included and another one that applies an agent-based framework that in principle may not only be used in offline simulations but also as part of an on-site implementation with real-time interactions of agents.

The model experiments described in [31] suggest that the actual power flow calculations do not differ much between the tools. Yet when it comes to proposing measures for grid expansion, substantial differences arise. These mainly stem from different (heuristic) algorithms being used to identify suitable grid extension measures. In the scenario investigating the installation of an VRDT, two different control strategies are applied. Thereby the models provide similar results, with the main differences being related to the level of network operation points (i.e. time steps) considered. The reliability-centered tool provides a rather detailed approach when it comes to analyzing the impact of potential faults on the system. Yet this comes at the expense of less time steps being considered, which in turn implies somewhat diverging results for the normal operation stage. On the other hand, the tool in question is the only one, which is also capable of investigating faulty network states. In these states, some voltage violations are observable. For the dynamic voltage control strategy, the results show that all models are capable of simulating such control strategies, even though some noticeable differences in the outcomes occur. Overall the outcomes of this model experiment demonstrate that representations of the single grid elements are rather similar. Yet more crucial is the question how these element-wise models are combined to an overall system and how temporal aspects are handled in the simulations.

2.6. open_MODEX

The open_MODEX project focused on a systematic and profound comparison of open source energy system models. These were referred to as frameworks in the open_MODEX project, but this is mostly analogous to the term model in the other projects. The restriction to open models allowed the consortium to gain detailed insights into the code and application of the models. For a systematic assessment of system quality, information quality and interface quality, a novel method for evaluating the usability of energy system models was developed [32]. This includes standardized steps and a comprehensive evaluation form. The method can be applied to individual or multiple models as part of a cross-evaluation approach. In this approach, developers of one model evaluate the usability of other models. Such a cross-evaluation approach was applied in the case study, in which five temporally and spatially resolved power sector models were evaluated in their usability. Thereby, the correct provision of input data and unclear error messages were identified as the main challenges. Complementing this, fundamental differences in the formulation of the models involved in open_MODEX were identified and analyzed [33]. To facilitate a direct mathematical comparison, a common terminology was defined first. The most relevant differences concern time representation, cost accounting, power sector technology modeling, intertemporality features, and multi-sector features. A case study was conducted to quantitatively evaluate some of these differences. In doing so, the investment cost calculation, storage modeling, and the use of a heuristic rather than an optimized technology deployment were identified as key drivers of result differences. If limited to dispatch optimization, there was mostly a high agreement in the annual power generation of individual technologies. In contrast, endogenous capacity optimization resulted in substantially higher deviations. Another case study under harmonized application of the five models participating in the project focused on the simplified analysis of possible power system designs for Germany in 2030 [34]. Assumptions on the CO_2 budget, electricity demand, coal phase-out, and VRE share were varied. The results show that the models make significantly different choices for endogenous investments and technology deployment, and the effect of the model differences increases with the defined VRE share. In addition, the model experiment analyzed and quantified the installation rates of key technologies needed to achieve the new federal government's goals. Here, the results show that the planned VRE capacities for 2030 do not guarantee that the targeted 80% supply share can be realized. Also in open_MODEX, the data templates as well as the data sets of the case studies were published in full [35, 36].

3. Overarching findings

The overview of the project results of the MODEX network reflects the wide range and diversity of the model experiments carried out (Section 2). Nevertheless, some common features and emphases become apparent. For example, methodological aspects of qualitative model comparison were addressed in all projects. While in some projects these covered only the model scope and its basic approach, in others smaller model features were also systematically included. In one project, even the model formulation was explicitly compared. Across all projects, it became apparent that a fundamental qualitative comparison of the models is a necessary prerequisite for understanding the differences in quantitative indicators derived from the model results. Not all papers include a qualitative model comparison. This results from the fact that often the same models were used in several experiments and a repeated presentation of the qualitative differences is omitted.

Most of the model experiments were limited to, or focused heavily on, the power sector. However, in some cases, the flexibility of coupling to other sectors was at least partially considered. In the MODEX-EnSaVes project, demand models and vehicle diffusion models were also compared. There was a clear focus not only thematically, but also in the model properties. Thus, in almost all projects temporally and spatially resolved power system models were analyzed. The exception here is the MEO project in which the operation of electricity, gas and heat distribution networks was addressed. Despite the fundamental similarity of the models considered, the results of the other projects can also only partially be compared directly. This is related to the different focus in the models and experiments, which is reflected in a wide range of different spatial and technological scope and detail. However, some common observations in the results can be identified. This includes the insight that differences in the consideration of endogenous capacity expansions are a key driver of divergent model results. In addition, it was found that the type and scope of modeling of sector coupling technologies are of great importance for the model results.

Most of the experiments dealt with optimization models. Deviating from this, data models for disaggregating model input data for high-resolution models were compared in [22], simulation models in [20], and a combination of optimization and simulation models in [31, 21].

Two main approaches were taken in the quantitative comparisons of optimization models (Table 1). Either simplified or even highly stylized scenarios were considered under full harmonization of the models and their input data, or realistic scenarios of the energy system transformation were analyzed under partial model and data harmonization. The reason for this is the experience of these and previous model experiments that harmonization is associated with a very high effort. This is due to different model and data structures as well as different nomenclatures used in the models. The use of simplified test cases is a vehicle here to reduce complexity and enable more extensive harmonization. In return, the possibility of also deriving policy-relevant insights from the modeling work decreases. There is a similar correlation with the number of scenarios, which can be increased more easily with simplified test cases. Across all projects, harmonized models show mostly similar results that can be explained in their differences. With increasing heterogeneity of input data and model scopes, however, the divergences of results and the challenges to comprehensively attribute them to model differences increase.

A cross-project data management group was established to ensure a close exchange regarding the approach to data harmonization and data publication within the MODEX network. Among other things, this group provided systematic coverage of the barriers and requirements of comprehensive data harmonization in model experiments [37]. In addition, methods for data harmonization and documentation were collected and compared. For this purpose, the required input data of all model experiments were gathered in a first step. On this basis, common data requirements and possibilities for harmonization were identified. It was found that due to the wide range of models and experiments, very few assumptions, such as fuel prices, could be harmonized across all projects. In addition, the importance of documenting and licensing data for reuse by other scientists was highlighted.

The MODEX projects strongly contributed to higher transparency regarding models and data, yet also some challenges regarding open data and open source approaches became apparent. Many models are deliberately designed as open source models and also the availability of open data has considerably increased over the last years. This is a key enabler for independent scientific review and cross-model validations like the MODEX model experiments. Additionally, open science approaches meet the increasing requirement of public funding bodies for open provision of research outputs. However, the success of collaborative and continued open source development is dependent on the steady availability of funding for model maintenance. Some research groups by contrast continue to operate undisclosed models along with proprietary data. This often comes along with closer collaborations with industry partners (electric utilities and/or grid operators), who are willing to share data only under non-disclosure agreements. Due to the greater proximity to application compared to pure research projects, these activities often also result in additional requirements for model quality. In this sense, frequent application in contract research projects for companies or public institutions can have a positive impact on the quality of models and their input data. This in turn can lead to an improved competitive position for further collaborative research projects and studies. Using standard economic concepts, this be may be stated as follows: the possibility to generate additional revenues (research contracts) through proprietary knowledge incentivizes the research group to invest more in that knowledge, and thus to continuously improve models and data. In contrast, the disclosure of models and data can help to increase the visibility of the group in the research community and serve as evidence of high-quality work. To summarize, open and closed approaches have partially different drivers and limitations. However, a general judgment on the transparency and quality of the data, models and results cannot be derived from this. Over the last years, hybrid strategies have emerged, with research groups sharing parts of their data and models. This may be a way forward as it enables cross-model comparisons like the MODEX model experiments while maintaining at the same time additional incentives for investments in better data.

Table 1: Characteristics of the model experiments performed in the MODEX projects. In the columns on harmonization (harm.) of data and models, an "x" stands for a complete, an "(x)" for a partial harmonization. Model harmonization here refers to an identical technology scope.

Project	**Ref.**	**Technology focus**	**Data harm.**	**Model harm.**	**Number of scenarios**	**Number of models**	**Stylized test cases**	**Real-world scenarios**	**Capacity expansion**	**Qualitative comparison**
MODEX-EnSAVes	[19]	Technologically detailed power system	(x)		1	4		x	x	x
MODEX-EnSAVes	[20]	Passenger vehicle market diffusion	(x)		2	3		x		x
MODEX-Net	[21]	Power system and transmission grid	(x)		2	8		x		x
MODEX-Net	[22]	Power demand and VRE generation	(x)		2	8		x		
MODEX-Net	[23]	Power network congestion management	(x)		2	8		x		
FlexMex	[24]	Power system flexibility options	x	x	18	9	x			x
FlexMex	[25]	Power plant, grid, storage expansion	x	x	4	6	x		x	
FlexMex	[26]	Power system flexibility options	x	(x)	16	8	x		x	
MODEX-POLINS	[29]	Impact of carbon pricing	(x)		2	5		x	x	
MODEX-POLINS	[30]	Flexible power-heat coupling	(x)		2	5		x	x	
MEO	[31]	Power distribution grid	x	x	2	4	x			x
open_MODEX	[32]	Systematic model usability testing	x	x	1	5	x			x
open_MODEX	[33]	Power system modeling techniques	x	x	2	5		x	x	x
open_MODEX	[34]	Power system emission reductions	x	x	8	5		x	x	x

4. Lessons learned

After submission of their papers for the Special Issue, the authors of all contributions were asked about their experiences in the respective MODEX projects. Questions covered their main areas of activity, challenges encountered and learning progress. First of all, the high response rate in the survey is remarkable: 41 completed questionnaires with approx. 80 contacted authors correspond to a noteworthy response of over 50%. In addition, the authors' experience in energy system modeling is also remarkable: 25% of the respondents indicated an experience of more than 10 years, another 30% between 5 and 10 years and 40% of the authors still have an experience of more than one to 5 years. The focus of the authors' work was relatively evenly distributed over data harmonization, model preparation, benchmarking and further model development.

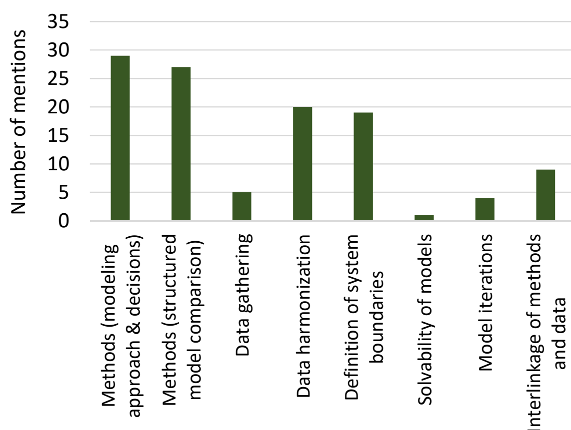

Figure 1: Feedback from the survey of modelers on the aspects of model comparison in which a high level of learning could be achieved (multiple choice).

The purpose and the benefits of model experiments are, of course, manifold: authors primarily report lessons learned concerning modeling approaches and model comparison methods (Figure 1). A high level of learning was also reached regarding data harmonization and the definition of system boundaries. Instead, data gathering, solvability of models as well as model iterations were considered less important in this regard. On average, substantial learning was achieved in three of the proposed areas. The range and number of responses on domains of learning clearly underscore the added value of model experiments.

Major challenges in the MODEX projects resulted especially from data harmonization, but model understanding, result interpretation, and model preparation were also perceived challenging by several participants (Figure 2). The

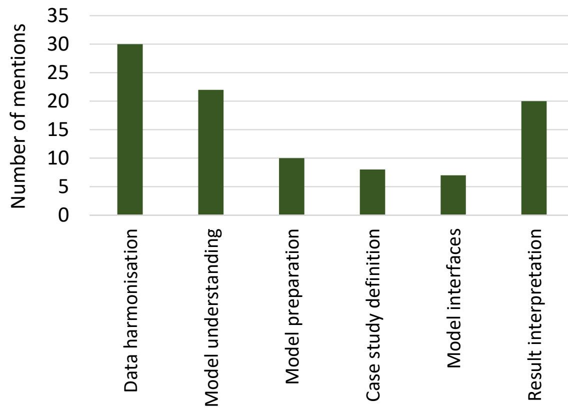

Figure 2: Feedback from the survey of modelers on the main challenges during the model experiments (multiple choice).

identification of data harmonization as the greatest challenge reflects the importance of establishing uniform data formats, parameter labels, and interfaces.

The central role of methodology and data handling in the lessons learned and challenges is also reflected in the evaluation of the project experience. Thus, the majority of modelers indicate the exchange among each other as the most appreciated aspect (Figure 3). This is in line with previous model experiments and also the experience of the editors of this Special Issue. The understanding of different models was named as second most important benefit from the MODEX experiments. In summary, the value of model experiments for the modelers can be emphasized again, as well as the wide range of experiences made. This finally led to the authors expressing a considerable interest in conducting further model experiments.

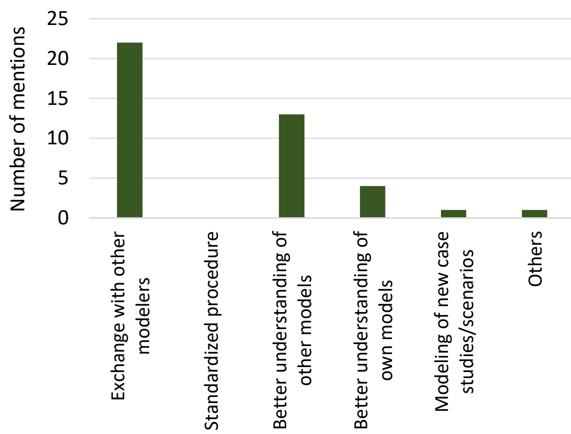

Figure 3: Feedback from the survey of modelers on the key benefit of the model experiments (single choice).

5. Conclusions and Outlook

The MODEX projects were able to significantly expand the spectrum of systematic comparisons of energy system models. This was ensured by the wide range of projects and the different subject areas and model types they focused on. The added value compared to previous model experiments consists particularly in the profound characterization of the models involved, the consistent linking of model differences with outcome differences, and the transparent description of appropriate methods and challenges of comparing energy system models. The implementation of a project network favored the achievement of complementary insights by combining the different perspectives. Not only by their breadth, but also by their depth, the MODEX model experiments offer many new findings. Thus, across all projects, much time was invested in the detailed qualitative analysis of models. In addition, models and input data were harmonized wherever possible and useful in the respective project context. Finally, the harmonized model application to differently designed scenarios and test cases allowed the profound quantitative analysis of the effect of model differences. Due to a high degree of transparency, these findings can be used directly by developers and users of other models.

Model experiments such as those carried out within the framework of MODEX not only strengthen the understanding of the modelers involved and external modelers, they also ensure a strengthening of the model quality and thus the informative value of the models. This is underlined by the conclusions of the publications produced in the projects, as well as by the survey among the participating modelers.

Notwithstanding the different model types and objects of investigation, it became apparent across all projects that each model was developed with certain focal points in mind. As a result, the level of detail is particularly high for these focal points, while simplifications are made elsewhere. This in turn can lead to extensive differences in the results of the models, even with harmonized input data. Depending on the project, however, it was also possible to find large similarities in the results for individual model pairs or groups with a comparable focus.

Despite the wide range of models compared and the extensive further development of methodologies for the comparison of energy system models, the MODEX projects also reveal complementary research needs. This concerns, for example, the systematic comparison of solution times of different models, the investigation of security of supply or the extension to sectors and model types that were not considered in the MODEX projects. Such follow-up projects can build in many respects upon the model comparison methods and tools developed in the MODEX projects as well as on the data provided.

Acknowledgements

The MODEX research projects received funding from the German Federal Ministry of Economic Affairs and Climate Action under grant numbers 03ET4074 (MODEX-Net), 03ET4075 (MODEX-POLINS), 03ET4076 (open_MODEX), 03ET4077 (FlexMex), 03ET4078 (MEO), and 03ET4079 (MODEX-EnSAVes). The authors thank the members of the Forschungsnetzwerk Energiesystemanalyse for initiating a project network on model experiments, the German Federal Ministry of Economic Affairs and Climate Action for financing the projects and the Projektträger Jülich for coordinating the activities of the MODEX network. Furthermore, we thank Aoife Foley for enabling and managing the MODEX Special Issue.

References

[1] Matteo Giacomo Prina, Giampaolo Manzolini, David Moser, Benedetto Nastasi, Wolfram Sparber, Classification and challenges of bottom-up energy system models - A review, Renewable and Sustainable Energy Reviews 129 (2020). doi:10.1016/j.rser.2020.109917.
URL http://www.sciencedirect.com/science/article/pii/S1364032120302082

[2] Amirhossein Fattahi, Jos Sijm, André Faaij, A systemic approach to analyze integrated energy system modeling tools: A review of national models, Renewable and Sustainable Energy Reviews 133 (2020). doi:10.1016/j.rser.2020.110195.
URL https://www.sciencedirect.com/science/article/pii/S1364032120304858

[3] Christian Klemm, Peter Vennemann, Modeling and optimization of multi-energy systems in mixed-use districts: A review of existing methods and approaches, Renewable and Sustainable Energy Reviews 135 (110206) (2021). doi:10.1016/j.rser.2020.110206.
URL http://www.sciencedirect.com/science/article/pii/S1364032120304950

[4] Forum für Energiemodelle und energiewirtschaftliche Systemanalysen in Deutschland (Ed.), Energiemodelle zum Klimaschutz in Deutschland - Strukturelle und gesamtwirtschaftliche Auswirkungen aus nationaler Perspektive, Springer, 1999.

[5] Forum für Energiemodelle und energiewirtschaftliche Systemanalysen in Deutschland (Ed.), Energiemodelle zum Kernenergieausstieg in Deutschland - Effekte und Wirkungen eines Verzichts auf Strom aus Kernkraftwerken, Springer, 2002.

[6] Forum für Energiemodelle und energiewirtschaftliche Systemanalysen in Deutschland (Ed.), Energiemodelle zum Klimaschutz in liberalisierten Energiemärkten - Die Rolle erneuerbarer Energieträger, LIT, 2004.

[7] Forum für Energiemodelle und energiewirtschaftliche Systemanalysen in Deutschland (Ed.), Energiemodelle zum europäischen Klimaschutz - Der Beitrag der deutschen Energiewirtschaft, LIT, 2004.

[8] Forum für Energiemodelle und energiewirtschaftliche Systemanalysen in Deutschland (Ed.), Energiemodelle zu Innovation und moderner Energietechnik - Analyse exogenen und endogenen technischen Fortschritts in der Energiewirtschaft, LIT, 2007.

[9] M. Sugiyama, S. Fujimori, K. Wada, S. Endo, Y. Fujii, R. Komiyama, E. Kato, A. Kurosawa, Y. Matsuo, K. Oshiro, F. Sano, H. Shiraki, Japan's long-term climate mitigation policy: Multi-model assessment and sectoral challenges, Energy 167 (2019) 1120–1131. doi:https://doi.org/10.1016/j.energy.2018.10.091.
URL https://www.sciencedirect.com/science/article/pii/S0360544218320814

[10] J. E. Bistline, M. Brown, S. A. Siddiqui, K. Vaillancourt, Electric sector impacts of renewable policy coordination: A multi-model study of the north american energy system, Energy Policy 145 (2020) 111707. doi:https://doi.org/10.1016/j.enpol.2020.111707.
URL https://www.sciencedirect.com/science/article/pii/S0301421520304341

[11] H. G. Huntington, A. Bhargava, D. Daniels, J. P. Weyant, C. Avraam, J. Bistline, J. A. Edmonds, S. Giarola, A. Hawkes, M. Hansen, P. Johnston, A. Molar-Cruz, M. Nadew, S. Siddiqui, K. Vaillancourt, N. Victor, Key findings from the core north american scenarios

in the emf34 intermodel comparison, Energy Policy 144 (2020) 111599. doi:https://doi.org/10.1016/j.enpol.2020.111599.
URL https://www.sciencedirect.com/science/article/pii/S0301421520303372

[12] S. Giarola, A. Molar-Cruz, K. Vaillancourt, O. Bahn, L. Sarmiento, A. Hawkes, M. Brown, The role of energy storage in the uptake of renewable energy: A model comparison approach, Energy Policy 151 (2021) 112159. doi:https://doi.org/10.1016/j.enpol.2021.112159.
URL https://www.sciencedirect.com/science/article/pii/S0301421521000288

[13] K. Siala, M. Mier, L. Schmidt, L. Torralba-Díaz, S. Sheykhha, G. Savvidis, Which model features matter? an experimental approach to evaluate power market modeling choices, Energy 245 (2022) 123301. doi:https://doi.org/10.1016/j.energy.2022.123301.
URL https://www.sciencedirect.com/science/article/pii/S0360544222002043

[14] Hans Christian Gils, Thomas Pregger, Franziska Flachsbarth, Mareike Jentsch, Constantin Dierstein, Comparison of spatially and temporally resolved energy system models with a focus on Germany's future power supply, Applied Energy 255 (2019). doi:10.1016/j.apenergy.2019.113889.

[15] Systems analysis research network.
URL https://www.forschungsnetzwerke-energie.de/system-analysis

[16] Open energy platform.
URL https://openenergy-platform.org/

[17] Y. Scholz, B. Fuchs, F. Borggrefe, K.-K. Cao, M. Wetzel, K. von Krbek, F. Cebulla, H. C. Gils, F. Fiand, M. Bussieck, T. Koch, D. Rehfeldt, A. Gleixner, D. Khabi, T. Breuer, D. Rohe, H. Hobbie, D. Schönheit, H. Ü. Yilmaz, E. Panos, S. Jeddi, S. Buchholz, Speeding up energy system models - a best practice guide (Juni 2020).
URL https://gitlab.com/beam-me/bpg

[18] Elias Ridha, Lars Nolting, Aaron Praktiknjo, Complexity profiles: A large-scale review of energy system models in terms of complexity, Energy Strategy Reviews 30 (2020). doi:10.1016/j.esr.2020.100515.
URL http://www.sciencedirect.com/science/article/pii/S2211467X20300687

[19] S. Misconel, R. Leisen, J. Mikurda, F. Zimmermann, C. Fraunholz, W. Fichtner, D. Möst, C. Weber, Systematic comparison of high-resolution electricity system modeling approaches focusing on investment, dispatch and generation adequacy, Renewable and Sustainable Energy Reviews 153 (2022) 111785. doi:https://doi.org/10.1016/j.rser.2021.111785.

URL https://www.sciencedirect.com/science/article/pii/S1364032121010546

[20] T. Gnann, D. Speth, K. Seddig, M. Stich, W. Schade, J. Gómez Vilchez, How to integrate real-world user behavior into models of the market diffusion of alternative fuels in passenger cars - an in-depth comparison of three models for germany, Renewable and Sustainable Energy Reviews 158 (2022) 112103. doi:https://doi.org/10.1016/j.rser.2022.112103.
URL https://www.sciencedirect.com/science/article/pii/S1364032122000326

[21] C. Syranidou, M. Koch, B. Matthes, C. Winger, J. Linßen, C. Rehtanz, D. Stolten, Development of an open framework for a qualitative and quantitative comparison of power system and electricity grid models for europe, Renewable and Sustainable Energy Reviews 159 (2022) 112055. doi:https://doi.org/10.1016/j.rser.2021.112055.
URL https://www.sciencedirect.com/science/article/pii/S1364032121013174

[22] O. Raventós, T. Dengiz, W. Medjroubi, C. Unaichi, A. Bruckmeier, R. Finck, Comparison of different methods of spatial disaggregation of electricity generation and consumption time series, Renewable and Sustainable Energy Reviews 163 (2022) 112186. doi:https://doi.org/10.1016/j.rser.2022.112186.
URL https://www.sciencedirect.com/science/article/pii/S1364032122001101

[23] H. Hobbie, J. Mehlem, C. Wolff, L. Weber, F. Flachsbarth, D. Möst, A. Moser, Impact of model parametrization and formulation on the explorative power of electricity network congestion management models - insights from a grid model comparison experiment, Renewable and Sustainable Energy Reviews 159 (2022) 112163. doi:https://doi.org/10.1016/j.rser.2022.112163.
URL https://www.sciencedirect.com/science/article/pii/S1364032122000910

[24] H. C. Gils, H. Gardian, M. Kittel, W.-P. Schill, A. Zerrahn, A. Murmann, J. Launer, A. Fehler, F. Gaumnitz, J. van Ouwerkerk, C. Bußar, J. Mikurda, L. Torralba-Díaz, T. Janßen, C. Krüger, Modeling flexibility in energy systems — comparison of power sector models based on simplified test cases, Renewable and Sustainable Energy Reviews 158 (2022) 111995. doi:https://doi.org/10.1016/j.rser.2021.111995.
URL https://www.sciencedirect.com/science/article/pii/S1364032121012582

[25] J. van Ouwerkerk, H. C. Gils, H. Gardian, M. Kittel, W.-P. Schill, A. Zerrahn, A. Murmann, J. Launer, L. Torralba-Díaz, C. Bußar, Impacts of power sector model features on optimal capacity expansion: A comparative

study, Renewable and Sustainable Energy Reviews 157 (2022) 112004. doi:https://doi.org/10.1016/j.rser.2021.112004.
URL https://www.sciencedirect.com/science/article/pii/S1364032121012673

[26] H. C. Gils, H. Gardian, M. Kittel, W.-P. Schill, A. Murmann, J. Launer, F. Gaumnitz, J. van Ouwerkerk, J. Mikurda, L. Torralba-Díaz, Model-related outcome differences in power system models with sector coupling—quantification and drivers, Renewable and Sustainable Energy Reviews 159 (2022) 112177. doi:https://doi.org/10.1016/j.rser.2022.112177.
URL https://www.sciencedirect.com/science/article/pii/S1364032122001022

[27] H. Gardian, H. C. Gils, M. Kittel, A. Murmann, J. Launer, F. Gaumnitz, A. Fehler, J. van Ouwerkerk, J. Mikurda, L. Torralba-Díaz, C. Krüger, T. Janßen, A. Zerrahn, Model input and output data of the FlexMex model comparison (2021). doi:10.5281/zenodo.5802178.
URL https://doi.org/10.5281/zenodo.5802178

[28] H. Gardian, H. C. Gils, Flexmex analysis tool (2022). doi:10.5281/zenodo.6010392.
URL https://doi.org/10.5281/zenodo.6010392

[29] O. Ruhnau, M. Bucksteeg, D. Ritter, R. Schmitz, D. Böttger, M. Koch, A. Pöstges, M. Wiedmann, L. Hirth, Why electricity market models yield different results: Carbon pricing in a model-comparison experiment, Renewable and Sustainable Energy Reviews 153 (2022) 111701. doi:https://doi.org/10.1016/j.rser.2021.111701.
URL https://www.sciencedirect.com/science/article/pii/S1364032121009758

[30] M. Bucksteeg, M. Wiedmann, A. Pöstges, M. Haller, D. Böttger, O. Ruhnau, R. Schmitz, The transformation of integrated electricity and heat systems—assessing mid-term policies using a model comparison approach, Renewable and Sustainable Energy Reviews 160 (2022) 112270. doi:https://doi.org/10.1016/j.rser.2022.112270.

[31] J.-P. Beck, J. Reinhardt, K. Kamps, J. Kupka, C. Derksen, Model experiments in operational energy system analysis: Power grid focused scenario comparisons, Renewable and Sustainable Energy Reviews XXX (2022) XX. doi:XX.
URL XX

[32] S. Berendes, S. Hilpert, S. Günther, C. Muschner, S. Candas, K. Hainsch, J. van Ouwerkerk, S. Buchholz, M. Söthe, Evaluating the usability of open source frameworks in energy system modelling, Renewable and Sustainable Energy Reviews 159 (2022) 112174.

doi:https://doi.org/10.1016/j.rser.2022.112174.
URL https://www.sciencedirect.com/science/article/pii/S1364032122001009

[33] S. Candas, C. Muschner, S. Buchholz, R. Bramstoft, J. van Ouwerkerk, K. Hainsch, K. Löffler, S. Günther, S. Berendes, S. Nguyen, A. Justin, Code exposed: Review of five open-source frameworks for modeling renewable energy systems, Renewable and Sustainable Energy Reviews 161 (2022) 112272. doi:https://doi.org/10.1016/j.rser.2022.112272.

[34] J. van Ouwerkerk, K. Hainsch, S. Candas, C. Muschner, S. Buchholz, S. Günther, H. Huyskens, S. Berendes, K. Löffler, C. Bußar, F. Tardasti, L. von Köckritz, R. Bramstoft, Comparing open source power system models - a case study focusing on fundamental modeling parameters for the german energy transition, Renewable and Sustainable Energy Reviews 161 (2022) 112331. doi:https://doi.org/10.1016/j.rser.2022.112331.

[35] S. Berendes, S. Hilpert, S. Günther, C. Muschner, S. Candas, K. Hainsch, J. van Ouwerkerk, S. Buchholz, M. Söthe, Evaluating the usability of open source frameworks in energy system modelling (supplementary material) (2022). doi:10.5281/zenodo.5578292.
URL https://doi.org/10.5281/zenodo.5578292

[36] C. Muschner, S. Berendes, S. Candas, K. Hainsch, K. Löffler, S. Günther, , S. Buchholz, R. Bramstoft, J. van Ouwerkerk, open$_m odexbasedataset$ (2022). doi:10.5281/zenodo.5854411.
URL https://doi.org/10.5281/zenodo.5854411

[37] H. Gardian, J.-P. Beck, M. Koch, R. Kunze, C. Muschner, L. Hülk, M. Bucksteeg, Data harmonisation for energy system analysis — example of multi-model experiments, Renewable and Sustainable Energy Reviews XXX (2022) XXX. doi:XXX.