

ASSESSING HISTORICAL REALIBILITY OF THE AGENT-BASED MODEL OF THE GLOBAL ENERGY SYSTEM

Anna Shchiptsova^{1,5} Jiangjiang Zhao² Arnulf Grubler^{1,3} Arkady Kryazhimskiy^{1,4}
Tiejun Ma^{1,2}

¹ *International Institute for Applied Systems Analysis, Laxenburg, Austria*

² *East China University of Science and Technology, Shanghai, China*

³ *School of Forestry and Environmental Studies Yale University, New Haven, USA*

⁴ *Steklov Mathematical Institute, Russian Academy of Sciences, Moscow, Russia*

⁵ *shchipts@iiasa.ac.at* (✉)

Abstract

This study looks at the historical reliability of the agent-based model of the global energy system. We present a mathematical framework for the agent-based model calibration and sensitivity analysis based on historical observations. Simulation consistency with the historical record is measured as a distance between two vectors of data points and inference on parameter values is done from the probability distribution of this stochastic estimate. Proposed methodology is applied to the model of the global energy system. Some model properties and limitations followed from calibration results are discussed.

Keywords: agent-based modeling, calibration, energy system

1. Introduction

Traditional energy system planning and climate policy analysis models have their relative strength in the analysis of normative policy goals under conditions of foresight based on a rational actor (social planner) decision paradigm and under limited uncertainties. Major transitions, in terms of energy systems structure or climate change implications, are driven either by exogenous discontinuities (e.g. assumed technological breakthroughs), or by global, perfect implementation of modeled policies, typically regulating prices or quantities (or both). In real life however, actor decisions that drive

transitions are heterogeneous, interdependent, myopic and are better characterized by “bounded rationality”. There is also no global coordination and perfect cooperation among actors, making universal implementation of policy measures impossible to achieve. This is among the reasons why to date energy systems and climate policy models are used exclusively in a forecasting mode, projecting out into the future, and have to date not been able to replicate major past transitions (Grubler 2012). This also severely limits any endogenous validation of future scenarios against the historical record.

An alternative (and complementary)

modeling approach was developed in Ma et al. (2008) and Ma and Grubler (2008) based on the stochastic agent-based simulation framework. Agent-based modeling is viewed as a powerful tool for analysis of complex systems driven by heterogeneous interacting actors and yields new insights by generating multiple system development paths for different model instances (Bonabeau 2002). It has been applied in various research fields such as stock market (Palmer et al. 1994), dissemination of culture (Axelrod 1997), electricity trading (Bunn and Oliveira 2001), drivers' route choice (Dia 2002) and pedestrian walking behavior (Antonini et al. 2006), co-evolution of individual behaviors and social institutions (Bowles et al. 2003), formation of virtual organizations (Norman et al. 2004), co-evolution of parochial altruism and war (Choi and Bowles 2007), evacuation from buildings under fire (Shi et al. 2009), diffusion of epidemic diseases (Beyrer et al. 2012) and technologies (Delre 2007), military trainings (Cioppa et al. 2004), industry transformation (Isley et al. 2013), etc.

A comprehensive survey of validation techniques associated with simulation models can be found in Balci (1994) and Sargent (2013), which describe various formal and informal methods used during model development. Studies devoted to measuring simulations consistency with the empirical evidence are diverse and base on different, often alternative key principles. Social science approaches to empirical validation include indirect calibration, Werker-Brenner calibration and history-friendly validation. Detailed description and critical review of these methods are given in Fagiolo et al. (2007). Alternative approaches involve

statistical methods (Kleijnen 1995) and companion modeling (Barreteau 2003, Moss 2008). Use of statistical techniques is common in studies of complex systems, but restricted by data availability and the nature of the studied phenomenon. Agent-based modeling often serves to describe a system that undergoes several structural changes. This requirement demands adaptation of classical statistical tests, and sometimes makes them impossible to use. Companion modeling is based on expert judgement and engages stakeholders in the modeling and validation process (Moss 2008), and thus, it can be criticized as being subjective, hard to replicate and limited in the studies of emergent phenomena.

Our approach to assessing historical reliability of the agent-based model is in line with the indirect calibration method. Proposed calibration criterion and the choice of methods used in sensitivity analysis of calibration results facilitate interpretation of the obtained "plausible" subset in the parameter domains and thus, serve as an attempt to resolve the second problem of indirect calibration, which was mentioned in Fagiolo et al. (2007).

The agent-based model is a tool for *forward modeling*, i.e. it serves to make prediction on the possible trajectories of system development. Calibration of the agent-based model fundamentally represents the problem of *inverse modeling*. The classical approaches to inverse problem involve regularization techniques (Tikhonov and Arsenin 1977) and provide the "best estimate" solution. However, validation of the agent-based model against the historical record requires not just point estimates of the best-fit parameters but also complete statistical

information about consistency between model simulations and historical observations. The Bayesian framework (Tarantola 2005, Kaipio and Somersalo 2005) does this by conjunction of information from the theoretical forward model and prior information on the observations and model parameters. The solution of inverse problem is a posterior distribution in the space of model parameters. This approach becomes computationally expensive in the case of the high dimensional historical data, which is essentially present in the model of the global energy system, where the historical record is given by the time series over more than 200 years. In this study we propose to combine information from the agent-based model and historical observations to reduce the dimensionality of the problem. After that, inverse modeling on the obtained quantity can be performed.

The rest of the paper is organized as follows. In Section 2 we introduce the agent-based model of the global energy system, model simulations and historical data. Section 3 outlines a mathematical framework, which addresses calibration of the agent-based model and sensitivity analysis application on the output results of the calibration procedure. Section 4 presents examples of calibration in the case of the agent-based model of the global energy system. We discuss results and model limitations in section 5. Some remarks are given in the final section.

2. Agent-based model

2.1 Description

A novel feature of the agent based model of

the global energy system is that it treats technologies as "agents", which are defined at the level of a facility/plant or a device that transforms resources or energy flows following both the tradition of activity or process analysis (Ayres and Kneese 1969) as well as that of "bottom-up" energy models (e.g. Messner and Strubegger 1994, Riahi et al. 2007, Riahi et al. 2012). Technologies have characteristics defined by their resource/energy inputs, outputs, resulting efficiency with associated emissions and costs. For simplicity we do not differentiate between capital and operating costs of technologies and use levelized costs. It is the characteristics of technologies that govern their long-term survival under the selection environment of our technology system (and not their mere existence).

The model starts in 1800 with several basic primary energy technologies (e.g. biomass burning for providing heat). New energy technologies come into being little by little. Existing energy technologies and their combinations form energy chains which connect primary resources or energy sources/forms to the energy service demands of consumers. Energy chains are either new combinations of primary energy technologies or re-combinations of previously existing chains. Alternative technological combinations or chains can provide the same energy services, and hence they compete as in the real world (Halsnaes et al. 2007, Grubler 2012). The model assumes that the cheaper technological chains to the given service demand will prevail over time.

The emergence of new technologies as well as their (re-)combination into new energy chains is essentially conceptualized via a random walk model (reflecting the unpredictability, often

serendipity, of technological innovation) subject however to resource constraints and economic incentives. New technologies cannot survive indefinitely once formed. If a technology or a combination is not used to providing energy service over a certain time period, it will be removed from the system. The model uses parameter called "retention time" for newly emergent technologies that cannot be integrated immediately into new technological combinations but nonetheless "stay around" for a while awaiting potential integration into the technology system.

Initially energy service demands are assumed to be given. A constant annual growth rate is associated with each kind of energy service demand. While the energy system unfolds, each demand will be adjusted with a given price elasticity depending on the price/cost of energy service. New technological combinations can also create new demands. Modern service demand (e.g. electricity for running computers) is not available at the early stage of the energy system development, and is triggered with the emergence of power generation technologies and modern devices such as telephones and computers.

There is a technological learning effect for an emergent new technology in the model. The learning rate follows a lognormal distribution. The more technology is tried, the higher the probability that its cost will decrease. Consequently new technologies with highly uncertain characteristics and small market volume will be adopted only very cautiously, yielding the classical slow take-off pattern of technological diffusion (Grubler 1991).

The first version of the agent-based model

with 62 energy technologies was reported in Ma et al. (2008). Here we use a version of the model with 133 energy technologies to include more end-use technologies. Mathematical details of the agent-based model can be found in Ma et al. (2008), Ma and Grubler (2008) and in the appendix.

2.2 Model simulations and historical observations

We do not provide a detailed description of the historical trajectory of the global energy system in terms of the emergence of hundreds combinations of individual energy technologies, which are to a high degree substitutable and may be overredundant from a structural perspective of system development. Aiming to register the occurrence of major historical transition events, we define the historical record in terms of the dynamics of highly aggregate characteristics of the energy system. For this purpose we use data from De Stercke (2014). In particular, we describe development of the energy system by the values of 6 response indicators:

- heat demand
- mobility demand
- modern service demand (specific services provided by modern devices such as computers, consuming electricity or hydrogen in addition to other energy services)
- non-fuel demand (industry feedstock, i.e. energy used for non-energy purposes)
- total energy demand (the aggregate value of four energy service demands)
- primary carbon (carbon emissions in primary energy).

The latter response indicator is a structural

variable in the model, which illustrates environmental climate change, while other indicators are primary variables related to energy demand.

Model simulations are generated for the past development of the energy system starting from 1800 and ending at 2010 with the step equal to one year. We are interested in differentiation between model instances, which vary in the values of the limited number of parameters. Other variable characteristics of the agent-based model remain constant during simulations and define a context, in which calibration procedure is performed. In this study we reduce the efficiency of boiler technologies as described in Table 1 and set the growth rates of energy service demands to the values shown in Table 2. The values of these characteristics in the agent-based model were minor revised to be consistent with the initial state of the historical trajectory of primary carbon.

Table 1 Efficiency of boiler technologies

Technology	Efficiency
Boiler coal	0.27
Boiler oil	0.28
Boiler gas	0.28
Boiler biomass	0.26
Boiler h2	0.30
Boiler methanol	0.22
Boiler ethanol	0.22

Besides, we postpone the entrance of technologies, which form energy chains to non-fuel demand. These technologies become available from the year 1875.

Calibration procedure focuses on 4 input parameters in the agent-based model. Namely,

- initial cost of technologies (initial

investment to be made to adopt a new technology, used as a proportional coefficient for INV values in the Tables A.2 and A.3), values from 1 to 10 (dollars)

- learning rate (rate at which technological costs reduce with experience accumulation from its usage), values from 0.05 to 0.45 (dimensionless)
- innovation rate (rate at which new technologies appear and become available to form energy chains), values from 0.015 to 0.135 (dimensionless)
- retention time of technologies (represents innovation impatience, time for which a new technology stays available to form energy chains), values from 20 to 60 (years).

These parameters are significant factors which influence the evolution process and complexification of the simulated global energy system (Ma and Grubler 2008).

As a useful summary, Figure 1 shows the range of system development paths generated from the agent-based model. Simulations were run for 10000 combinations of parameters, whose values vary in the domains described above. In total, Figure 1 displays 1 million model trajectories.

In this paper we concentrate on the cases where each parameter takes the extreme and average values from its domain. There are 81 model instances, for which simulations are repeated 100 times. This data is taken as an input for the calibration procedure described below. Thus, the objective of the study is to test the ability of the agent-based model to replicate historical development in the limited number of general scenarios.

Table 2 Dynamics of energy service demands

Energy service demand	Initial volume (kwyr)	Annual growth rate	Price elasticity
Heat	5.71×10^7	6.13%	0.2
Mobility	1.77×10^5	2.12%	0.5
Modern service	1.66×10^4	4.03%	0.3
Non-fuel	1.38×10^6	5.49%	0.3

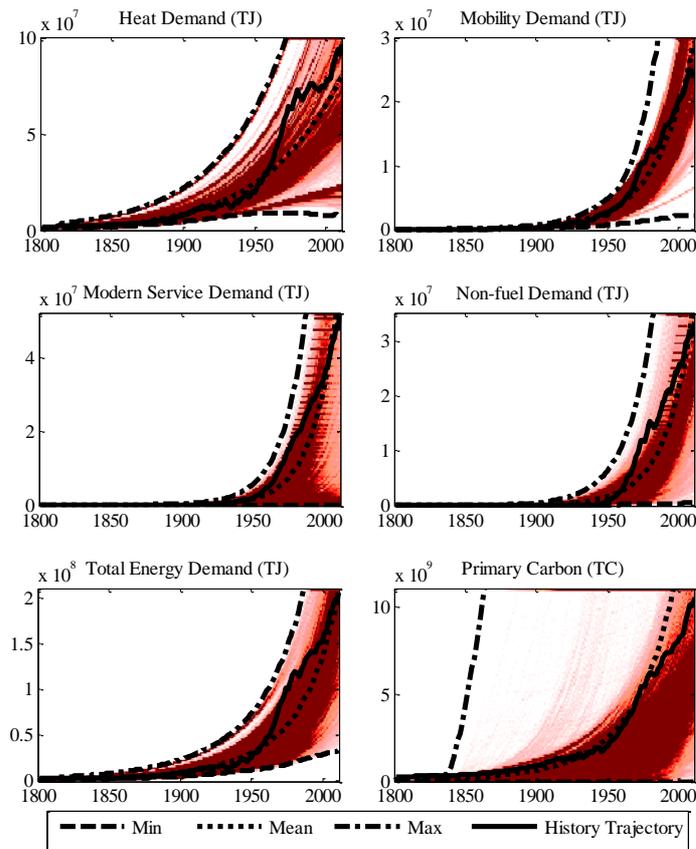


Figure 1 Model simulations and historical data for individual response indicators

3. Methodological framework

3.1 Notations

Let's consider the agent-based model as a black box, which provides a point-by-point probabilistic mapping from the input parameters

of the model to its response indicators. Each vector of parameters defines conditions, in which the model simulates development trajectories of the system. Simulation outcomes are represented by m time series $s_1^i(p), s_2^i(p), \dots, s_T^i(p)$ over T years, where p is a vector of input parameters

and i is an index of the simulation run. We measure the fit of the simulation outcome to the historical time series h_1, h_2, \dots, h_T by distance between two vectors of data points:

$$d_i(p) = \sqrt{\sum_{t=1}^T (s_t^i(p) - h_t)^2}, \quad (1)$$

where $p \in P$ and $i = 1 \dots m$. P is a finite set of possible parameter combinations. Note that for now we consider one-dimensional trajectories.

As the agent-based model has stochastic nature, distance $d(p)$ is a random value. Here we make no assumptions on the distribution of this random variable, but proceed with the empirical cumulative distribution function $\hat{F}_p(d)$ as its non-parametric estimator. Additionally, we discretize data and put a uniform grid on the distance values. Below it is assumed that grid approximation of the empirical distribution function can be neglected with respect to the proposed solution of the inverse problem.

3.2 Calibration

Distance $d(p)$ measures misfit between prediction from the model p and the observed trajectory from the historical record. Inference on the parameter values is based on the distribution of this random variable. We define the process of calibration of the agent-based model as ordering of model instances with different input parameter combinations by the value of the selected calibration criterion. In fact, model ordering means the ordering of the elements in the set P of the input parameters.

Our choice of the calibration criterion is deduced from geometrical considerations. Without any prior conditions on the shape of the

distance probability distribution we say that the model with parameter vector p is better in terms of its consistency with the historical record than the model with parameter vector p' , if the area under the empirical distribution function $\hat{F}_p(d)$ is bigger than the area under $\hat{F}_{p'}(d)$. This definition is equivalent to the fact that we seek to minimize the expected value of distance distribution. Moreover, we can simply associate each parameter combination with the normalized value $S(p)$, which we call a *relative volume* of the distance distribution for parameter p

$$S(p) = \frac{1}{n} \sum_{k=1}^n \hat{F}_p(d^k), \quad (2)$$

where d^k denotes the right border of the grid cell k and n is the number of grid cells. Hence, the relative volume $S(p)$ measures models' fit to the historical record on average.

After we calculated the relative volume $S(p)$, parameter combinations can be sorted by this value. The result of this procedure is a permutation $\sigma(p_1, p_2, \dots, p_{|P|})$ of the sequence $(p_1, p_2, \dots, p_{|P|})$ of the enumerated parameter vectors in P . For this permutation inequalities

$$S(p_{\sigma^{-1}(1)}) \geq S(p_{\sigma^{-1}(2)}) \geq \dots \geq S(p_{\sigma^{-1}(|P|)}) \quad (3)$$

hold. Here $|P|$ is a number of parameter combinations in P and $\sigma^{-1}(i)$ denotes an index in the original sequence for the i -th parameter vector in the permutation. We call this permutation an *optimal permutation* of parameter combinations. For the sake of simplicity, we do not consider a case when any of the inequalities, which define the optimal permutation, holds as equality, and therefore, the optimal permutation of parameters is unique.

Before now we expected that the system trajectory is one-dimensional. Generally output data of the agent-based model is given in multiple response indicators. Thereby the fit of the model simulations to the historical record requires collective measurement in several dimensions. In fact, we should derive a joint probability of events that each one-dimensional model trajectory is closer to the respective historical trajectory than some threshold value. The average measurement of these probabilities taken over all possible combinations of the grid cells determines the value of the calibration criterion. Let's assume that indicator trajectories are independent. Then the joint probability represents a product of probabilities of such events. Moreover, it can be easily shown that a *joint relative volume* $S^N(p)$ is equal to the product of relative volumes associated with each individual indicator:

$$S^N(p) = S_1(p) \cdot S_2(p) \cdot \dots \cdot S_N(p), \quad (4)$$

where $S_i(p)$ is a relative volume of the i -th output indicator and N is the number of output indicators. The value of this calibration criterion represents a normalized $(N + 1)$ -dimensional volume under the joint empirical distribution function. In case of $N = 1$ the joint relative volume matches the value of the relative volume connected with the individual response indicator and corresponds to the area under the empirical distribution function of the distance random variable.

3.3 Sensitivity analysis

Basically, proposed calibration process is a mapping of every parameter combination p to its relative volume $S(p)$ and its position in the

optimal permutation. We can perform sensitivity analysis to enhance insight on the relationships between input and output in the agent-based model. Our objective here is not connected with the development of a model with some prespecified relationship (e.g. using regression analysis), but rather it consists in testing whether some dependence is present in the optimal permutation.

Let's say that we picked a parameter combination in the observed optimal permutation of parameters. In this case a question of interest is to estimate that the change in a value of some selected parameter in this combination will improve the agent-based model fit to the historical trajectory. In fact, we explore whether some values locally dominate another ones from the same parameter domain.

Note that calibration results for one parameter can be treated as outcomes from the random experiment, whose sample space contains all possible ordered sequences of parameter values. We observe repetitions of this experiment and the number of repetitions is determined from the number of possible combinations of other parameter values in the model. As the result, we can estimate the probability distribution of the outcomes in this random experiment and consequently, probabilities of events that there is a preference in the individual parameter values in the calibration model (e.g. the probability that the maximum value from the domain of selected parameter improves models' fit to the historical record for an arbitrary parameter combination). If we observe some preference (local dominance) in the value of selected parameter, this means that with the high probability we cannot refine calibration results by changing this parameter

from the dominant values and this holds for an arbitrary combination of other parameters.

On the other hand, the optimal permutation produces a sequence of the ordered values for each individual parameter. Therefore, our calibration model can be treated as an information source. In its widest sense, information source is just an object that emits a sequence of symbols from a finite value domain according to some statistical rules (Shannon 1948, Welsh 1988). Our goal is to explore a structure of the source observing a finite stream of data. From the practical perspective this is equivalent to the question at that extent we can predict values of the single parameter independently from other parameters in the calibration results and whether there are globally dominant values in the parameter domain.

Basically, the source uncertainty is quantified by its entropy rate, which gives the average entropy per symbol of the source. By its definition the information source represents a stochastic process. The type of the process gives us a context, in which randomness in the calibration model can be measured using the entropy rate. In particular, we assume that there is a statistical dependence of the next symbol in the sequence on the values, which were observed in the past. For this purpose we use a first order Markov approximation built from the data. To ensure that the Markov source is irreducible, we transform the optimal permutation of parameter values into a circular sequence by adding the first two values of the sequence to its end as proposed in Rukhin (2000).

The entropy rate of the first order Markov source X with K states equals

$$H(X) = \sum_{k=1}^K w_i H_i, \quad (5)$$

where H_i is an entropy of the state i and w_i is a stationary probability associated with this state. In entropy computations we take the logarithm to base 2. Consequently, the entropy rate is measured in bits per symbol. The value of the entropy rate lays in the range from 0 to $\log_2 L$, where L is the number of values in parameter domain. The low value of the entropy rate indicates that assumed statistical dependence is present in the optimal permutation and therefore, we can distinguish a subset of globally dominant (or dominated) values in the parameter domain.

4. Numerical examples

In this section we present calibration results for the agent-based model of the global energy system. All subsequent analysis was done on the grid with 50 cells, which covers changes in the distance values of the individual response indicator.

At first, we carried out proposed calibration procedure for each response indicator independently. The results are summarized in Table 3, which shows the range of solutions in terms of relative volume and corresponding changes in the expected value of distance. The “best estimate” solution indicates a limit of how well the agent-based model can replicate the historical record on average. On the other hand the “worst estimate” identifies the upper bound of an average model deviation from the historical trajectory.

Figure 2 demonstrates the dynamics of the models’ fit to the historical record, when we switch from one parameter combination to

another in the optimal permutation. The models' fit is measured relative to the "perfect fit", which corresponds to the case when the agent-based model repeats historical observations in every simulation run (with the precision defined by the lower grid border). It is a point, at which $S(p)$ reaches its maximum.

Sensitivity analysis results indicate that there exists a local dominance in the calibration results. Table 4 summarizes all cases of such kind. These results illustrate, for instance, that the maximum value from the domain of the initial cost of technologies parameter is almost surely

dominated by other values in the heat demand dimension, when we pick an arbitrary parameter combination.

A first order Markov source has a modeling descriptive power in three cases as shown in Table 5. The value of the entropy rate lays in the range from 0 to 1.58 bits per symbol. A small value of it indicates that the calibration model possesses Markov property and can be approximated by a first order source. Inference on the parameter values is obtained from the optimal permutation and from a Markov chain associated with the source. Thus, we can

Table 3 Calibration results for the case of individual response indicators

Response indicator.	"Best estimate" solution		"Worst estimate" solution	
	Relative volume $S(p)$	Expected value $d(p)$	Relative volume $S(p)$	Expected value $d(p)$
Heat demand	0.92	44.31 (kwyr ·10 ⁻⁸)	0.77	99.84 (kwyr ·10 ⁻⁸)
Mobility demand	0.97	5.26 (kwyr ·10 ⁻⁸)	0.87	14.74 (kwyr ·10 ⁻⁸)
Modern service demand	0.88	27.91 (kwyr ·10 ⁻⁸)	0.74	50.77 (kwyr ·10 ⁻⁸)
Non-fuel demand	0.92	14.98 (kwyr ·10 ⁻⁸)	0.84	24.59 (kwyr ·10 ⁻⁸)
Total energy demand	0.91	69.19 (kwyr ·10 ⁻⁸)	0.70	167.32 (kwyr ·10 ⁻⁸)
Primary carbon	1	38.18 (tC ·10 ⁻⁸)	0.97	1420.49 (tC ·10 ⁻⁸)

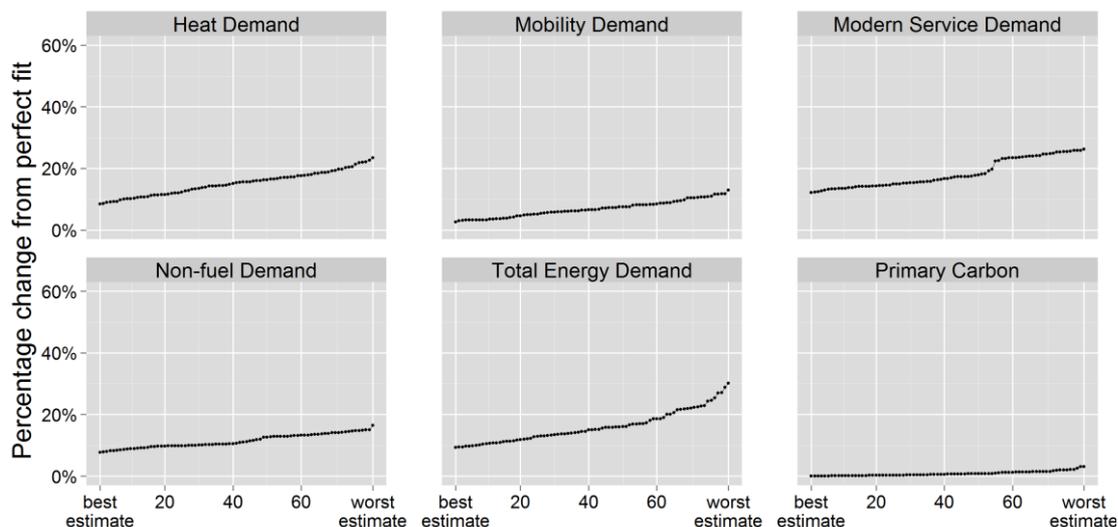


Figure 2 Dynamics of calibration results for the case of individual response indicators

Table 4 Locally dominant values in parameter domains. Case of individual response indicators

Response indicator	Initial cost of technologies $p^{(1)}$	Learning rate $p^{(2)}$	Innovation rate $p^{(3)}$	Retention time of technologies $p^{(4)}$
Heat demand	1, 5 (min. and avr.) (Prob.=0.96)	—	0.075, 0.135 (avr. and max.) (Prob.=0.92)	40, 60 (avr. and max.) (Prob.=1)
Mobility demand	5 (avr.) (Prob.=1)	—	—	—
Modern service demand	—	—	0.075, 0.135 (avr. and max.) (Prob.=1)	40, 60 (avr. and max.) (Prob.=1)
Non-fuel demand	1, 5 (min. and avr.) (Prob.=0.93)	—	0.075, 0.135 (avr. and max.) (Prob.=1)	—
Total energy demand	1, 5 (min. and avr.) (Prob.=0.92)	—	0.075, 0.135 (avr. and max.) (Prob.=0.93)	40, 60 (avr. and max.) (Prob.=0.96)
Primary carbon	—	—	—	40, 60 (avr. and max.) (Prob.=0.92)

Table 5 Dominance in the first order Markov approximation for the case of individual response indicators

Response indicator	Parameter	Entropy rate $H(X)$	Dominated value	Dominant value
Mobility demand	initial cost of technologies	0.89	—	5 (avr.)
Modern service demand	innovation rate	0.75	0.015 (min.)	—
Non-fuel demand	innovation rate	0.93	0.015 (min.)	—

differentiate model instances by the global dominance relation in an individual parameter in

these three response indicators with the uncertainty measured by the entropy rate.

The output of the agent-based model of the global energy system is described at different levels of aggregation. Joint measurement of models' fit to the historical record is done at the level of response indicators related to the structure of energy demand and at the level of general response indicators. The structure of energy demand is described by four energy service demands: heat, mobility, modern service and non-fuel demands. The general indicator group includes total energy demand (as a primary variable associated with energy) and primary carbon (as a structural variable associated with carbon emissions). We assume independence of indicator trajectories in both cases. The results

are summarized in Table 6 and illustrated in Figure 3.

Table 6 Calibration results for the case of multiple response indicators

Response indicators	"Best estimate" solution	"Worst estimate" solution
	Joint relative volume $S^n(p)$	Joint relative volume $S^n(p)$
Energy service demands ($n=4$)	0.69	0.44
Total energy demand + primary carbon ($n=2$)	0.91	0.69

Table 7 Locally dominant values in parameter domains. Case of multiple response indicators

Response indicators	Initial cost of technologies $p^{(1)}$	Learning rate $p^{(2)}$	Innovation rate $p^{(3)}$	Retention time of technologies $p^{(4)}$
Energy service demands	1, 5 (min. and avr.) (Prob.=1)	—	0.075, 0.135 (avr. and max.) (Prob.=1)	40, 60 (avr. and max.) (Prob.=0.96)
Total energy demand + primary carbon	1, 5 (min. and avr.) (Prob.=0.93)	—	0.075, 0.135 (avr. and max.) (Prob.=0.93)	40, 60 (avr. and max.) (Prob.=1)

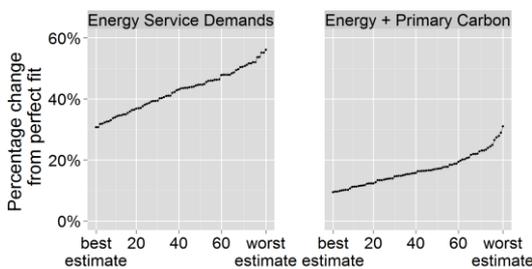


Figure 3 Dynamics of calibration results for the case of multiple response indicators

Sensitivity analysis reveals the same dominance relation in the individual parameter values in both cases of joint measurement. Results are included in Table 7.

5. Discussion

5.1 Variability and models' fit

Under model variability we distinguish two opposite cases for simulation trajectories in the particular model instance (predictions on system development paths under the same initial conditions). The model instance has low variability, when the range and heterogeneity of trajectories are limited. On the other hand, variability is high, when simulation trajectories are expanding and (or) have different functional structure. Calibration of the global energy system reveals the fact that variability in system

development paths of the particular agent-based model does not make a linear impact on the ability of this model instance to replicate historical record. This empirical observation is illustrated in Figure 4 for the heat demand response indicator. Simply speaking, it is not

necessary that two adjacent parameter combinations in the calibration results (with small deviation in $S(p)$) correspond to the models with similar structure in development paths.

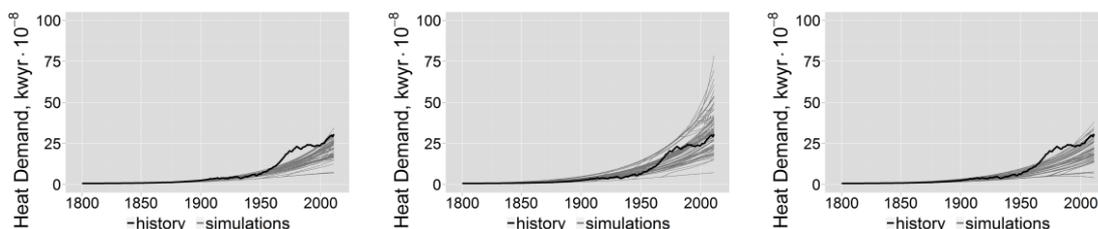


Figure 4 Simulation trajectories for successive parameter combinations in calibration results.

Case of the heat demand indicator

5.2 Energy service demands

The historical fit for energy service demands is characterized by the low costs of technologies, which have a moderate threshold to being created and adopted in the existing energy chains. But at the same time technologies stay available for a long period after they were developed. This scenario (in terms of parameter combinations) is preferable to replicate the historical record in independent measurements for heat and mobility demands and for joint measurement of energy demand structure, which reconciles effects in each of energy demand dimensions.

Additionally, results of sensitivity analysis suggest that technological learning rate has a highly irregular and nonlinear impact on the system ability to replicate observations in energy demands. This result is also supported by the fact that the value of learning rate differs significantly in the “best estimate” solutions of independent measurements. It should also be noted from performed analysis that a low threshold for technologies creation (minimum value of

innovation rate parameter) almost surely limits model performance in the dimensions of modern service and non-fuel demands. In fact, we observe from the model, that energy chains to these demands require creation of sophisticated technologies to fit the historical rate (which is plausible in terms of historical experience).

5.3 Predictions on carbon emissions

We observe the almost constant performance of the agent-based model in the primary carbon dimension. This result is illustrated by small deviations in the values of $S(p)$ shown in Figure 2. Thereby, the rate of carbon emissions is modeled independently from the changes in the input parameters. This observation leads us to the conclusion that all model instances have similar historical fit measured relative to primary carbon. Note that at the same time absolute value of the expected distance changes significantly because of the presence of outlier trajectories in the generated simulations. Empirical results also suggest that the number of outliers is proportional

to the position of model instance in the calibration results.

5.4 Joint measurement in dimensions of energy and carbon emissions

The joint assessment of historical fit in several response indicators is a product of average model performance in each of the dimensions. Measurements in individual indicator contribute equally to the value of calibration metric. But we cannot conclude how model performs relative to one indicator based on joint assessment, which is nonlinear. That's why it is important to explore model performance in each dimension simultaneously with joint measurement. In particular, joint assessment for general indicator group (total energy demand and primary carbon) is independent on one of the indicators. Almost constant behavior of calibration results in the primary carbon dimension does not influence joint model performance. So, we cannot draw a conclusion how model behaves in both energy demand and carbon emissions relative to historical observations.

6. Conclusions

Our results reveal irregular and nonlinear performance of the agent-based model of the global energy system measured relative to historical observations in energy demands. However, we observe some local properties of the model input parameters, which are defined via a restriction of parameter domains on the sets of dominant values. Calibration results indicate that the agent-based model fails to predict the past behavior of the energy system in the context of energy demand and carbon emissions. This

limitation should be taken with caution because of exploratory character of our study. Extensive simulation runs with increased parameter domains should provide solid evidence whether model predictions on the past are limited or not.

We should mention here that proposed mathematical framework allows validating an agent-based model on the bigger dataset and in principal, to study model sensitivity in the scope of parameter interdependencies. The computational and data complexification in the model may require advanced techniques from the fields of data analysis and information theory. Application of these advanced mathematical methods to agent-based model validation can be viewed as a possible direction of future research.

7. Appendix A

Most of the mathematical details of the global energy system model can be found in Ma and Grubler (2008). Here we report only changes made in the agent-based model with 133 technologies, which has more end-use technologies.

The share of the i -th chain in the energy service demand equals

$$\tilde{s}_i^{t+1} = \begin{cases} s_i^t \frac{p^t}{c_i^{t+1}}, & c_i^{t+1} \geq p^t \\ \frac{es_i^t}{(e-1)s_i^t + 1}, & c_i^{t+1} < p^t \end{cases} \quad (\text{A.1})$$

where \tilde{s}_i^{t+1} is the non-normalized share at step $t + 1$, s_i^t is the normalized share at step t , p^t is the price for satisfying the demand at step t , c_i^{t+1} is the cost of technology chain at step $t + 1$ and e is defined as

$$e = 2.7182^{\frac{\ln 81}{5+51.64+(p^t/c_i^{t+1})^{-6.95}}} \quad (\text{A.2})$$

After that the chain shares s_i^{t+1} at step $t + 1$ are normalized.

Parameters in equation (A.2) are chosen in such a way that when the chain's cost is close but less than the market price, it will take around 50 years for the chain diffusion from 10% to 90%.

The chain is considered as inactive when its share is less than 10^{-6} .

There is a limit on the annual use of each renewable resource, which cannot be exceeded: biomass - 2.56×10^{10} kwyr, wind - 9.5×10^9 kwyr, hydro - 3.56×10^9 kwyr, geothermal - 10^{10} kwyr. If this limit is reached, then the share of the chains, which have this resource as an input resource, is adjusted proportional to L/E^t , where L is the limit value and E^t is the cumulative extraction of this resource at step t .

After this adjustment some demand is freed and needs to be allocated between existing energy chains. The weights are calculated according to (A.3) and then normalized

$$\tilde{w}_i^t = \max\left(0, 1 - 10^{\frac{c_i^t - c_{min}^t}{c_{min}^t}}\right), \quad (A.3)$$

where \tilde{w}_i^t is the non-normalized weight for additional demand allocation, c_i^t is the chain cost at step t and c_{min}^t is the minimal cost of technology chain at step t .

It is possible that a technology is deployed as a component in several different energy chains for satisfying final energy service demands. We associate cumulative capacity with each component. Initial cumulative capacity of the component is set to 1000 kwyr.

Investment cost for a regular technology includes a constant cost and the cost of learning, which equals the sum of costs of all components. The cost of a non-extraction singular technology

has only the cost of learning. The levelized cost C_l of a non-extraction technology is calculated as follows

$$C_l = \frac{1}{pf} \left(\frac{inv}{pl} + fom \right) + vom, \quad (A.4)$$

where pf is a plant factor, inv is an investment cost, pl is a plant life, fom is a fixed operation and maintenance cost and vom is a variable operation and maintenance cost.

The number of drawing and combinations of technologies and chains M^t depends on the size of the economy Ω^t and is calculated as

$$M^t = \min(5000, 50 + 500 \ln \Omega^t), \quad (A.5)$$

$$\Omega^t = \max\left(1, \frac{\sum_{i=1}^m p_i^t d_i^t}{\sum_{i=1}^m p_i^0 d_i^0}\right),$$

where p_i^t and d_i^t are the price and demand of the energy service i at step t , p_i^0 and d_i^0 are the initial price and demand of the energy service i , m is the number of the energy service demands.

8. Appendix B

Table A.1 Energy forms in the reference energy system

Id	Description	Level
0 1	Coal	Resource
0 2	Crude oil	Resource
0 3	Gas	Resource
0 4	Biomass	Resource
0 5	Uranium	Resource
0 6	Wind	Resource
0 7	Hydro	Resource
0 8	Solar	Resource
0 9	deuterium	Resource
0 A	Geothermal	Resource
1 2	Crude oil	Primary
2 1	Coal as second energy	Secondary
2 2	Methanol as second energy	Secondary
2 3	Oil Products	Secondary
2 4	Gas as second energy	Secondary

2 5	Ethanol as second energy	Secondary
2 6	Biomass as second energy	Secondary
2 7	Nuclear fuel	Secondary
2 8	Electricity	Secondary
2 9	Hydrogen	Secondary
2 A	Heat	Secondary
3 1	Coal	Secondary
3 2	Methanol	Secondary
3 3	Oil Products	Secondary
3 4	Electricity	Final
3 5	Gas	Final
3 6	Ethanol	Final
3 7	Biomass	Final
3 8	Hydrogen as final energy	Final
4 2	Modern Services	Useful
4 3	Heat	Useful
4 4	Mobility	Useful
4 5	Non-Fuel	Useful

Table A.2 Definition of singular technologies in the reference energy system*

Technology	In	Out	INV	FOM	VOM	PL	PF	EFF	CO2	Time	LBD
Coal extr	0 1	2 1						1		2	0
Oil extr	0 2	1 2						1		2	0
Gas extr	0 3	2 4						1		2	0
Bio extr	0 4	2 6						1		1	0
Ura extr	0 5	2 7						1		3	0
H2 Elec	2 8	2 9	1000	23	4	30	0.95	0.80	0	2	1
Elec t/d	2 8	3 4	800	55	18	60	0.55	0.86	0	2	1
Animal	3 7	4 4	100	1	0.2	10	0.2	0.01	0.942	1	0
Elec Heating device	3 4	4 3	100	1	0	20	0.5	0.95	0	2	1
Heat pump	3 4	4 3	600	1	0	20	0.5	2.50	0	2	1
Coal NF	3 1	4 5	10	1	0	30	0.6	1	0.814	2	1
Oil NF	3 3	4 5	10	1	0	30	0.6	1	0.631	2	1
Gas NF	3 5	4 5	10	1	0	30	0.6	1	0.482	2	1
Meth NF	3 2	4 5	10	1	0	30	0.6	1	0.549	2	1
Eth NF	3 6	4 5	10	1	0	30	0.6	1	0.549	2	1
H2 NF	3 8	4 5	10	1	0	30	0.6	1	0	2	1

Ele device	3 4	4 2	1000	1	0	30	0.6	1	0	2	1
Biomass t/d	2 6	3 7	10	0	20	30	1	0.93	0	1	0

Table A.3 Definition of regular technologies in the reference energy system*

Technology	In	Out	Description	INV	FOM	VOM	PL	PF	EFF	CO2
Oil refi	1 2	2 3	Refine oil	66	7.5	0	30	0.9	0.9	0.054
Gas bio	2 6	2 4	Produce gas from biomass	670	50	8	25	0.8	0.69	0.942
Met bio	2 6	2 2	Produce methanol from biomass	1580	83	8	25	0.9	0.56	0.636
Eth bio	2 6	2 5	Produce ethanol from biomass	1580	83	8	25	0.9	0.56	0.636
Met coal	2 1	2 2	Produce methanol from coal	1350	76	0	25	0.9	0.65	0.457
Gas coal	2 1	2 4	Produce gas from coal	850	57	0	25	0.8	0.76	0.448
Met gas	2 4	2 2	Produce methanol from gas	700	51	5.4	25	0.90	0.7	0.098
H2 Solid Coal	2 1	2 9	Produce h2 from coal	1250	72	0	25	0.9	0.74	0.814
H2 Solid Bio	2 6	2 9	Produce h2 from biomass	985	63	8	25	0.9	0.67	0.942
H2 Gas	2 4	2 9	Produce h2 from gas	480	41	0	30	0.9	0.77	0.482
H2 Nuclear	2 7	2 9	Produce h2 from nuclear high tempture rector	2000	96	0	30	0.75	2.73	0
H2 Solar	0 8	2 9	Solar thermal power plant for H2 production	4000	87	0	25	0.5	0.6	0
Wind ppl	0 6	2 8	Wind power plant	1400	58	0	30	0.2	0.385	0
PV ppl	0 8	2 8	Solar PV plant	5100	119	0	30	0.25	0.385	0
Solar thermal power	0 8	2 8	Solar thermal power plant	2900	87	0	25	0.5	0.385	0
Hydro ppl	0 7	2 8	Hydro power plant	1000	12.5	0	60	0.5	0.385	0
Geothermal	0 A	2 8	Geothermal power plant	1200	70	40	30	0.7	0.385	0
Fusion	0 9	2 8	Fusion power plant	7200	315	0	30	0.7	1	0
Fission	2 7	2 8	Nuclear fission power plant	1900	93	0	30	0.7	1	0
IGCC H2 Fuel cell coal	2 1	2 8	Coal fuel cell power plant	1790	50	87.6	30	0.8	0.45	0.814
IGCC H2 Fuel cell bio	2 6	2 8	Biomass fuel cell power plant	2000	40	90	30	0.8	0.4	0.942
FC gas	2 4	2 8	Fuel cell powered by gas	1500	52	0	25	0.65	0.6	0.482
FC h2	2 9	2 8	Fuel cell powered by h2	1400	40	0	25	0.65	0.65	0

Technology	In	Out	Description	INV	FOM	VOM	PL	PF	EFF	CO2
FC meth	2.2	2.8	Fuel cell powered by methanol	2000	40	0	25	0.65	0.55	0.549
FC eth	2.5	2.8	Fuel cell powered by ethanol	2000	40	0	25	0.65	0.55	0.549
Gas turbine ppl gas	2.4	2.8	Gas turbine power plant	600	12	21.9	25	0.9	0.33	0.482
Gas turbine ppl h2	2.9	2.8	Gas turbine power plant powered by h2	600	12	20	25	0.9	0.33	0
Single cycle ppl coal	2.1	2.8	Conventional coal power plant	1300	74	0	30	0.65	0.38	0.814
Single cycle ppl oil	2.3	2.8	Conventional oil power plant	730	52	0	30	0.65	0.40	0.631
Single cycle ppl gas	2.4	2.8	Conventional gas power plant	710	51	0	30	0.65	0.40	0.482
Single cycle ppl h2	2.9	2.8	Conventional h2 power plant	710	40	0	30	0.65	0.4	0
Single cycle ppl bio	2.6	2.8	Conventional biomass power plant	1600	84	0	30	0.65	0.29	0.942
Single cycle ppl met	2.2	2.8	Conventional methanol power plant	730	52	0	30	0.65	0.40	0.549
Single cycle ppl eth	2.5	2.8	Conventional ethanol power plant	730	52	0	30	0.65	0.40	0.549
Combined Cycle ppl oil	2.3	2.8	Combined cycle oil power plant	800	55	0	30	0.65	0.5	0.631
Combined Cycle ppl gas	2.4	2.8	Combined cycle gas power plant	730	52	0	30	0.65	0.5	0.482
Combined Cycle ppl h2	2.9	2.8	Combined cycle h2 power plant	730	52	0	30	0.65	0.5	0
Combined Cycle ppl meth	2.2	2.8	Combined cycle methanol power plant	800	55	0	30	0.65	0.5	0.549
Combined Cycle ppl eth	2.5	2.8	Combined cycle ethanol power plant	800	55	0	30	0.65	0.5	0.549
IGCC ppl coal	2.1	2.8	IGCC coal power plant	1400	85	69.2	30	0.75	0.42	0.814
IGCC ppl bio	2.6	2.8	IGCC biomass power plant	1800	90	0	25	0.65	0.46	0.942
Engine ppl coal	2.1	2.8	Engine coal power plant	900	27	17.52	30	0.65	0.3	0.814
Engine ppl oil	2.3	2.8	Engine oil power plant	600	46	0	30	0.65	0.4	0.631
Engine ppl gas	2.4	2.8	Engine gas power plant	700	25	8.76	15	0.65	0.36	0.482
Engine ppl h2	2.9	2.8	Engine h2 power plant	700	25	8.76	15	0.65	0.36	0
Engine ppl bio	2.6	2.8	Engine biomass power plant	1200	60	0	30	0.65	0.3	0.942
Engine ppl meth	2.2	2.8	Engine methanol power plant	600	46	0	30	0.65	0.4	0.631

Technology	In	Out	Description	INV	FOM	VOM	PL	PF	EFF	CO2
Engine ppl eth	2 5	2 8	Engine ethanol power plant	600	46	0	30	0.65	0.4	0.631
Coal t/d bus	2 1	3 1	Coal distribution by bus	1	30	8	10	0.5	0.93	0.057
Coal t/d train	2 1	3 1	Coal distribution by train	2	20	6	10	0.5	0.93	0.057
Met t/d pipes	2 2	3 2	Methanol distribution by pepes	11.89	0.28	4.79	40	0.7	0.94	0.033
Met t/d bus	2 2	3 2	Methanol distribution by bus	1	30	8	10	0.5	0.93	0.033
Met t/d train	2 2	3 2	Methanol distribution by train	2	20	6	10	0.5	0.93	0.033
Oilp t/d pipes	2 3	3 3	Oil distribution by pepes	11.89	0.28	4.79	40	0.7	0.95	0.033
Oilp t/d bus	2 3	3 3	Oil distribution by bus	1	30	8	10	0.5	0.95	0.033
Oilp t/d train	2 3	3 3	Oil distribution by train	2	20	6	10	0.5	0.95	0.033
Gas t/d	2 4	3 5	gas pipes/grid	200	24	3.5	40	0.7	0.9	0.048
Eth t/d pipes	2 5	3 6	Ethanol distribution by pipes	11.89	0.28	4.79	40	0.7	0.95	0.033
Eth t/d bus	2 5	3 6	Ethanol distribution by bus	1	30	8	10	0.5	0.95	0.033
Eth t/d train	2 5	3 6	Ethanol distribution by train	2	20	6	10	0.5	0.95	0.033
Bio t/d bus	2 6	3 7	Biomass distribution by bus	1	30	8	10	0.5	0.93	0.033
Bio t/d train	2 6	3 7	Biomass distribution by train	2	20	6	10	0.5	0.93	0.033
H2 t/d	2 9	3 8	H2 distribution	220	25	7	42	0.5	0.85	0
Heat t/d	2 A	4 3	Heat distribution	400	36	0	40	0.52	0.97	0
Terrestrial steam coal bus_track_ship	3 1	4 4	Terrestrial steam coal bus	750	0.2	0.057	10	0.5	0.04	0.814
Terrestrial steam coal train	3 1	4 4	Terrestrial steam coal train	600	0.1	0.057	10	0.5	0.06	0.814
Terrestrial steam oil car	3 3	4 4	Terrestrial steam oil car	750	0.3	0.057	10	0.5	0.035	0.631
Terrestrial steam oil bus_track_ship	3 3	4 4	Terrestrial steam oil bus	600	0.2	0.057	10	0.5	0.042	0.631
Terrestrial steam bio train	3 7	4 4	Terrestrial steam bio train	550	0.1	0.057	10	0.5	0.05	0.942
Terrestrial IC oil car	3 3	4 4	Terrestrial IC oil car	900	0.3	0.057	10	0.65	0.2	0.631

Technology	In	Out	Description	INV	FOM	VOM	PL	PF	EFF	CO2
Terrestrial IC oil bus_track_ship	3 3	4 4	Terrestrial IC oil bus	600	0.2	0.057	10	0.65	0.24	0.631
Terrestrial IC oil train	3 3	4 4	Terrestrial IC oil train	300	0.1	0.057	10	0.65	0.25	0.631
Terrestrial IC gas car	3 5	4 4	Terrestrial IC gas car	900	0.3	0.057	10	0.65	0.23	0.482
Terrestrial IC gas bus_track_ship	3 5	4 4	Terrestrial IC gas bus	600	0.2	0.057	10	0.65	0.3	0.482
Terrestrial IC h2 car	3 8	4 4	Terrestrial IC h2 car	900	0.3	0.057	10	0.65	0.2	0
Terrestrial IC h2 bus_track_ship	3 8	4 4	Terrestrial IC h2 bus	600	0.2	0.057	10	0.65	0.267	0
Terrestrial IC h2 train	3 8	4 4	Terrestrial IC h2 train	300	0.1	0.057	10	0.65	0.3	0
Terrestrial IC meth car	3 2	4 4	Terrestrial IC meth car	900	0.3	0.057	10	0.65	0.25	0.549
Terrestrial IC meth bus_track_ship	3 2	4 4	Terrestrial IC meth bus	600	0.2	0.057	10	0.65	0.3	0.549
Terrestrial IC eth car	3 6	4 4	Terrestrial IC eth car	900	0.3	0.057	10	0.65	0.3	0.549
Terrestrial IC eth bus_track_ship	3 6	4 4	Terrestrial IC eth bus	600	0.2	0.057	10	0.65	0.35	0.549
Terrestrial FC gas car	3 5	4 4	Terrestrial FC gas car	5400	0.3	0.057	10	0.65	0.3	0.482
Terrestrial FC gas bus_track_ship	3 5	4 4	Terrestrial FC gas bus	3600	0.2	0.057	10	0.65	0.4	0.482
Terrestrial FC gas train	3 5	4 4	Terrestrial FC gas train	1800	0.1	0.057	10	0.65	0.5	0.482
Terrestrial FC h2 car	3 8	4 4	Terrestrial FC h2 car	5400	0.3	0.057	10	0.65	0.4	0
Terrestrial FC h2 bus_track_ship	3 8	4 4	Terrestrial FC h2 bus	3600	0.2	0.057	10	0.65	0.5	0
Terrestrial FC h2 train	3 8	4 4	Terrestrial FC h2 train	1800	0.1	0.057	10	0.65	0.6	0
Terrestrial FC meth car	3 2	4 4	Terrestrial FC meth car	5400	0.3	0.057	10	0.65	0.25	0.549
Terrestrial FC meth bus_track_ship	3 2	4 4	Terrestrial FC meth bus	3600	0.2	0.057	10	0.65	0.3	0.549
Terrestrial FC meth train	3 2	4 4	Terrestrial FC meth train	1800	0.1	0.057	10	0.65	0.4	0.549

Technology	In	Out	Description	INV	FOM	VOM	PL	PF	EFF	CO2
Terrestrial FC eth car	3 6	4 4	Terrestrial FC eth car	5400	0.3	0.057	10	0.65	0.25	0.549
Terrestrial FC eth bus_track_ship	3 6	4 4	Terrestrial FC eth bus	3600	0.2	0.057	10	0.65	0.3	0.549
Terrestrial FC eth train	3 6	4 4	Terrestrial FC eth train	1800	0.1	0.057	10	0.65	0.4	0.549
Terrestrial ele car	3 4	4 4	Terrestrial ele car	1800	0.3	0.057	10	0.65	0.4	0
Terrestrial ele bus_track_ship	3 4	4 4	Terrestrial ele bus	1200	0.2	0.057	10	0.65	0.5	0
Terrestrial ele train	3 4	4 4	Terrestrial ele train	600	0.1	0.057	10	0.65	0.52	0
Airborne IC oil	3 3	4 4	Airborne IC oil	900	1	0.2	10	0.4	0.074	0.631
Airborne IC gas	3 5	4 4	Airborne IC gas	900	1	0.2	10	0.4	0.1	0.482
Airborne IC h2	3 8	4 4	Airborne IC h2	900	1	0.2	10	0.4	0.15	0
Airborne IC meth	3 2	4 4	Airborne IC meth	900	1	0.2	10	0.4	0.12	0.549
Airborne IC eth	3 6	4 4	Airborne IC eth	900	1	0.2	10	0.4	0.12	0.549
Airborne turbine oil	3 3	4 4	Airborne turbine oil	900	1	0.2	10	0.4	0.3	0.631
Airborne turbine gas	3 5	4 4	Airborne turbine gas	900	1	0.2	10	0.4	0.35	0.482
Airborne turbine h2	3 8	4 4	Airborne turbine h2	900	1	0.2	10	0.4	0.4	0
Airborne turbine meth	3 2	4 4	Airborne turbine meth	900	1	0.2	10	0.4	0.31	0.549
Airborne turbine eth	3 6	4 4	Airborne turbine eth	900	1	0.2	10	0.4	0.31	0.549
Boiler coal	3 1	4 3	Coal heating plant	275	29	0	30	0.4	0.85	0.814
Boiler oil	3 3	4 3	Oil heating plant	155	21	0	30	0.4	0.9	0.631
Boiler gas	3 5	4 3	Gas heating plant	95	15	0	30	0.4	0.9	0.482
Boiler biomass	3 7	4 3	Biomass heating plant	275	29	0	30	0.4	0.83	0.942
Boiler h2	3 8	4 3	H2 heating rc	300	15	0	30	0.4	0.95	0
Boiler methanol	3 2	4 3	Methanol heating rc	155	21	0	30	0.4	0.7	0.549
Boiler ethanol	3 6	4 3	Ethanol heating rc	155	21	0	30	0.4	0.7	0.549
Boiler solar	08	43	Solar heating	4000	100	0	20	0.3	0.7	0
H2 device	3 8	4 2	Device powered by h2	2000	1	0	30	0.6	1	0

* Notations

- In: technology input energy form
- Out: technology output energy form
- INV: technology investment cost, \$/kwyr
- FOM: fixed operation and maintenance cost, \$/kwyr
- VOM: variable operation and maintenance cost, \$/kwyr
- PL: technology plant life

- PF: technology plant factor
- EFF: technology efficiency
- CO2: technology CO2 emission factor, tC/kwyr
- Time: groups technologies by the time when they become available in the model
- LBD: if equals 0, then there is no learning effect. In case of 1 denotes a 20% mean learning rate in the base line.

References

- [1] Antonini, G., Bierlaire, M. & Weber M. (2006). Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological*, 40 (8): 667-687.
- [2] Axelrod, R. (1997). The dissemination of culture a model with local convergence and global polarization. *Journal of Conflict Resolution*, 41 (2): 203-226.
- [3] Ayres, R. & Kneese, A. (1969). Production, consumption, and externalities. *The American Economic Review*, 59 (3): 282-297.
- [4] Balci, O. (1994). Validation, verification and testing techniques throughout the life cycle of a simulation study. In: Tew, J., Manivannan, S., Sadowski, D., Seila, A. (eds.), *Winter Simulation Conference*: 215-220, Orlando, December 11-14, 1994, IEEE, Piscataway, NJ.
- [5] Barreteau, O. (2003). Our companion modelling approach. *Journal of Artificial Societies and Social Simulation*, 6 (1). Available via <http://jasss.soc.surrey.ac.uk/6/2/1.html>. Cited October 29, 2015.
- [6] Beyrer, C., Baral, S., van Griensven, F., Goodreau, S., Chariyalertsak, S., Wirtz, A. & Brookmeyer, R. (2012). Global epidemiology of HIV infection in men who have sex with men. *Lancet*, 380 (9839): 367-377.
- [7] Bonabeau, E. (2002). Adaptive agents, intelligence, and emergent human organization: capturing complexity through agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99: 7280–7287.
- [8] Bowles, S., Choi, J. & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, 223 (2): 135-147.
- [9] Bunn, D. & Oliveira, F. (2001). Agent-based simulation - An application to the new electricity trading arrangements of England and Wales. *IEEE Transactions on Evolutionary Computation*, 5 (5): 493-503.
- [10] Choi, J. & Bowles, S. (2007). The coevolution of parochial altruism and war. *Science*, 318 (5850): 636-640.
- [11] Cioppa, T., Lucas, T. & Sanchez, S. (2004). Military applications of agent-based simulations. In: Ingalls R., Rossetti M., Smith J., Peters B. (eds), *Winter Simulation Conference*, 1. 171-180, Washington, December 5-8, 2004, IEEE.
- [12] Delre, S., Jager, W. & Janssen, M. (2007). Diffusion dynamics in small-world networks with heterogeneous consumers. *Computational and Mathematical Organization Theory*, 13 (2): 185-202.
- [13] De Stercke, S. (2014). Dynamics of Energy Systems: a Useful Perspective. *International*

- Institute for Applied Systems Analysis, Laxenburg, Austria, Interim Report IR-14-013.
- [14] Dia, H. (2002). An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C: Emerging Technologies*, 10 (5): 331-349.
- [15] Fagiolo, G., Moneta, A. & Windrum, P. (2007). A critical guide to empirical validation of agent-based models in economics: methodologies, procedures, and open problems. *Computational Economics*, 30 (3): 195-226.
- [16] Grubler, A. (2012). Energy transitions research: Insights and cautionary tales. *Energy Policy*, 50: 8-16.
- [17] Grubler, A. (1991). Diffusion: Long-term patterns and discontinuities. *Technological Forecasting and Social Change*, 39: 159-180.
- [18] Halsnaes, K., Shukla, P., Ahuja, D., Akumu, G., Beale, R., Edmonds, J., Gollier, C., Grubler, A. et al. (2007). Framing issues. In: Metz, B., Davidson, O., Bosch, P., Dave, R., Meyer, L. (eds), *Climate Change 2007: Mitigation. Contribution of Working Group III to the 4th Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 117-167. Cambridge: Cambridge University Press.
- [19] Isley, S., Lempert, R., Poper, S. & Vardavas, R. (2013). An Evolutionary Model of Industry Transformation and the Political Sustainability of Emission Control Policies. RAND Corporation technical report.
- [20] Kaipio, J. & Somersalo, E. (2005). *Statistical and Computational Inverse Problems*. Springer, New York.
- [21] Kleijnen, J. (1995). Verification and validation of simulation models. *European Journal of Operational Research*, 82: 145-162.
- [22] Ma, T. & Grubler, A. (2008). The evolution of technological complexity: An agent-based simulation model of the global energy system. In: Golub, A., Markandya, A. (eds), *Modeling Environment - Improving Technological Innovations under Uncertainty*, pp. 205-244., London: Routledge.
- [23] Ma, T., Grubler, A., Nakicenovic, N. & Arthur, W. (2008). Technologies as agents of change: a simulation model of the evolving complexity of the global energy system. International Institute for Applied Systems Analysis, Laxenburg, Austria, Interim Report IR-08-021.
- [24] Messner, S. & Strubegger, M. (1994). The energy model MESSAGE III. In: Hake, J. et al. (eds.), *Advances in Systems Analysis: Modelling Energy-Related Emissions on a National and Global Scale.*, Konferenzen des Forschungszentrums Juelich, Germany.
- [25] Moss, S. (2008). Alternative Approaches to the Empirical Validation of Agent-Based Models. *Journal of Artificial Societies and Social Simulation*, 11 (1). Available via <http://jasss.soc.surrey.ac.uk/11/1/5.html>. Cited October 29, 2015.
- [26] Norman, T., Preece, A., Chalmers, S., Jennings, N., Luck, M., Dang, V., Nguyen, T., Deora, V., Shao, J., Gray, W. & Fiddian, N. (2004). Agent-based formation of virtual organizations. *Knowledge Based Systems*, 17 (2): 103-111.
- [27] Palmer, R., Arthur, W., Holland, J., LeBaron, B. & Tayler, P. (1994). Artificial

- Economic Life: A Simple Model of a Stockmarket. *Physica D*, 75: 264-274.
- [28] Riahi, K., Grubler, A. & Nakicenovic, N. (2007). Scenarios of long-term socioeconomic and environmental development under climate stabilization. *Technological Forecasting and Social Change*, 74 (7): 887-935.
- [29] Riahi, K., Dentener, F., Gielen, D., Grubler, A., Jewell, J., Klimont, Z., Krey, V., McCollum, D., Pachauri, S., Rao, S., van Ruijven, B., van Vuuren, D. & Wilson, C. (2012). Energy Pathways for Sustainable Development. In: *Global Energy Assessment - Toward a Sustainable Future*, pp. 1203-1306. Cambridge, UK and New York, NY, US: Cambridge University Press and International Institute for Applied Systems Analysis, Laxenburg, Austria.
- [30] Rukhin, A. (2000). Approximate entropy for testing randomness. *Journal of Applied Probabilit.*, 37: 88-100.
- [31] Sargent, R. (2013). Verification and validation of simulation models. *Journal of Simulation*, 7: 12-24.
- [32] Shannon, C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27 (3): 379-423.
- [33] Shi, J., Ren, A. & Chen, C. (2009). Agent-based evacuation model of large public buildings under fire conditions. *Automation in Construction*, 18 (3): 338-347.
- [34] Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, Pa.
- [35] Tikhonov, A. & Arsenin, V. (1977). *Solution of Ill-posed Problems*. Wiley.
- [36] Welsh, D. (1988). *Codes and cryptography*. Oxford University Press.
- Anna Shchiptsova** is a Research Scholar at the International Institute for Applied Systems Analysis, Austria with a MSc in applied mathematics and informatics (Petrozavodsk State University, Russia) and a PhD in mathematics from the same university. Her main fields of interest include game theory, probability theory and statistics with their applications in mathematical modeling.
- Jiangjiang Zhao** is a PHD candidate at the school of Business at East China University of Science and Technology. His research is concerned with the various issues related to the agent-based model and optimization problems.
- Arnulf Grubler** is a Senior Research Scholar and Acting Program Director of the Transitions to New Technologies Program at the International Institute for Applied Systems Analysis, Austria. He also holds a part-time appointment as Professor in the Field of Energy and Technology at the School of Forestry and Environmental Studies at Yale University, New Haven, USA. His teaching and research focuses on the long-term history and future of technology and the environment with emphasis on energy, transport, and communication systems.
- Arkady Kryazhimskiy** † was a Principal Research Scholar at the Steklov Institute of Mathematics of the Russian Academy of Sciences, Russia; a Professor at the Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, Russia; and a Senior Research Scholar at the International Institute for Applied Systems Analysis, Austria. He was a Full Member of the

Russian Academy of Sciences. His research interests included game theory, optimization, and ill-posed problems.

Tieju Ma is a professor at School of Business, East China University of Science and Technology. His research interest is modeling technological dynamics with uncertainties and heterogeneous agents.