Comparing Behavioral Dynamics Across Models: the Case of Copper

J.H. Kwakkel (Corresponding author)
Delft University of Technology
Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft
J.H.Kwakkel@tudelft.nl
Tel: +31 (0)15 278 8457
Fax: +31 (0)15 278 6233

W. L. Auping
Delft University of Technology
Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

E. Pruyt
Delft University of Technology
Faculty of Technology, Policy and Management
Jaffalaan 5
2628 BX Delft

ABSTRACT
In many public policy issues diverging understandings of the system can be encountered. These diverging understandings can reside in the mental models of the different actors involved, or even be codified into structurally different models of the system. For an analyst it can be of great use to get insight into how and under what conditions the behavior of the models is different. In this paper, we address this problem. We present a general approach for comparing two or more structurally different models in the presence of additional uncertainties. This approach can be used to get insight into how different the results of two or more models are, and the conditions under which the models produce different results. The approach uses Exploratory System Dynamics Modeling in combination with the behavior pattern feature metric. We demonstrate the approach using a case study. This case study focusses on the future dynamics of the copper system. Here, there are experts favoring a top down way of modeling the system and there are experts favoring a bottom up way of modeling the system. We use both and find that for some outcomes of interest adopting either perspective makes no difference, while for other outcomes of interest only under specific additional assumptions about other uncertainties there is a difference in behavior.

1 Introduction

In a wide range of public policy issues, models are being used to support decision-making. One of the problems that can be encountered in model-based policy analysis is the presence of multiple structurally different models of the same system. The consequences of these models for policy could be could different. When confronted by structurally different models, it is important to provide insight into how the results of the model are different and under what conditions the models produce different results. Based on this, a structural explanation can be offered. The problem of alternative and potentially conflicting models for example can be encountered when using System Dynamics for addressing wicked problems (Rittel and Webber, 1973). In case of wicked problem, the available information is confusing, many stakeholders and decision-makers that each have their own mental model of the situation are involved, and many solutions are being proposed. (Churchman, 1967; Rittel and Webber, 1973). As a result, these situations are very resistant to being formulated as well defined problems (Rittel and Webber, 1973). Another example of a type of situation in which the problem addressed in this paper can be encountered is in decision-making under deep uncertainty (Kwakkel et al., 2010; Lempert et al., 2003). Under deep uncertainty, it is possible to enumerate alternative dynamic hypotheses without being able to indicate which of the alternative hypotheses is more likely or more probable (Kwakkel et al., 2010). Deep uncertainty can be encountered when the different actors involved in a decision-making problem do not agree on the relationships within the system, the input values for the model, the outcomes of interest and their relative importance (Lempert et al., 2003), or when decisions can be adapted in the future (Hallegatte et al., 2012). When dealing with wicked problems, or in case of decision-making under deep uncertainty, there is no a priori basis for guiding model selection. As recourse, it has been suggested that one should explore the consequences of the diverging understandings of the system. It is within such a context that the problem of comparing alternative models comes up.

When comparing alternative system dynamics models, the focus typically will be on comparing the dynamic behavior characteristics. The problem is that typical metrics for comparing model results do not focus on the dynamic behavior. As discussed in Yücel and Barlas (2011), typically one focuses on either comparing results for a given point in time, or on the deviation between two time series through something akin to the sum of squared error. The first is the most straightforward to understand and implement. The output of two models is compared for a shared outcome of interest and the difference is simply the difference between the two models for the value of the outcome of interest at a particular point in time. In such a comparison, the dynamics over time leading up to the value at a particular point in time are not taken into consideration. In a System Dynamics context such a metric is evidently questionable. The second way of comparing two models pays more attention to the dynamics over time. In this comparison the squared difference between the two models for a given outcome of interest over time is summed up over the time series. Although this metric considers the dynamics over time, this second metric can produce undesirable results where behaviorally identical outcomes are considered to be more different that outcomes which are not behaviorally identical (Yücel and Barlas, 2011).

In this paper we focus on the problem of comparing the simulation results arising from two or more alternative models. Following Lane (Lane, 2000b), we understand a System Dynamics model as being a concatenation of causal mechanism that taken together offer a plausible representation of a given system, and can explain a given model of behavior. These individual causal mechanisms are essentially hypothetical in character, as reflected in the use of the term 'dynamic hypothesis'. (Lane, 2000a; Sterman, 2000). So, a model for us in this paper is a dynamic hypothesis. To give an example to clarify this usage of the word model, consider a System Dynamics model implemented in say Vensim. There is a single Vensim file that contains the structure. Changing the value of a single parameter in this file would, under our usage of the word model, not be considered to produce an alternative model. This usage is in line with the usage found in for example sensitivity analysis, where one is inclined to speak of a single model where only one or more of the parameter values are being changed in order to study their influence on behavior. There are a few caveats with this usage of the word model. Most important, it is always possible to turn any given structure into any other structure by including one or more logical variables that switch parts of the model 'on' or 'off'. Such 'switch' variables over a convenient implementation to explore the influence of alternative model formulations. This suggests that the colloquial usage of the term model to refer to (e.g., Vensim) model files fails to make a clear distinction here. Under our understanding of the word model, if changing a few parameters in a Vensim file changes the causal structure, this single Vensim file would contain alternative models within a single file.

The starting point of this paper is that for whatever reason one is confronted with two or more models and one wants to compare the dynamics of these models. There are a variety of reasons why this could be relevant. In the context of model development and testing, knowing how models are different is a starting point for explaining this difference. As evidenced by the integrated assessment community and the climate community (Kriegler et al., 2014), there is a clear interest in understanding when and why different models of the same system produce different dynamics over time. For policy analysis, this insight is also critical. Knowing when and why models produce different results paves the way for designing policies that are insensitive to this (see e.g. Dalal et al., 2013; Hamarat et al., 2013; Kriegler et al., 2014; Lempert and Collins, 2007; Lempert et al., 2006). That is, policies that in either case produce satisficing results.

Methodologically, there are two issues that need to be addressed. First, we have to generate an ensemble of models that is consistent with the available information and encapsulates existing alternative and potentially conflicting mental models. To this end, we use Exploratory Modeling (Bankes, 1993; Kwakkel and Pruyt, 2013b) for systematically exploring the consequences of different dynamic hypotheses regarding a specific problem. Exploratory Modeling complements System Dynamics, for it focuses on how models are developed and used, but does not prescribe one particular modeling paradigm (Kwakkel and Pruyt, 2013b).

The second methodological issue is how to compare the behavioral dynamics of two models. The dynamic behavior over time can be understood as being a concatenation of atomic behavior patterns (Ford, 1999). The atomic behavior pattern is based on the sign (positive, negative, and zero) of the slope and curvature, resulting in nine possible atomic behavior patterns. In order to compare the behavioral dynamics of two models, we transform their behavior into a concatenation of atomic behavior patterns and compare these. The difference between two dynamics is then the average deviation across the entire concatenation. In essence, we are adapting the behavior pattern features discussed in (Yücel and Barlas, 2011) and further developed in (Yücel, 2012) to comparing the results from two models, rather than for validation and calibration.

2 Method

2.1 Exploratory System Dynamics Modeling

Exploratory system dynamics modeling and analysis (ESDMA) is a way of developing and using system dynamics models in situations rife with deep uncertainties (Kwakkel and Pruyt, 2013b). Note that this usage of exploratory system dynamics should not be confused with the usage of Homer (Homer, 1996, 2013). Homer (Homer, 1996, 2013) uses the term 'exploratory system dynamics' to denote an impressionistic, typically qualitative, way of developing and using system dynamic modeling. In contrast, developing and using models in the context of ESDMA is explicitly quantitative and rigorous.

ESDMA combines System Dynamics modeling with Exploratory Modeling. Exploratory Modeling is not a modeling paradigm in the way that System Dynamics, Discrete Event Simulation, and Agent-Based modeling are modeling paradigms. Exploratory modeling does not focus on the way in which a system is being represented in a simulation model. Rather, it focuses on how irreducible uncertainties can be handled within model-based policy analysis. The starting point of exploratory modeling is that in the presence of deep uncertainties, models cannot be used for accurate prediction. As also recognized in the system dynamics literature, for many systems of interest, the construction of a model that may be validly used as a surrogate is simply not possible (see e.g. Lane, 2012; Sterman, 2000). This may be due to a variety of factors, including the impossibility of accurate measurements or observations, immaturity of theory, nonlinearity of system behavior, dynamic complexity, ambiguity, misperception of feedback, judgmental errors and biases, the problem of under determination, etc. (Cambell et al., 1985; Oreskes et al., 1994; Sterman, 2000). Exploratory modeling starts from this fact of not knowing enough to make predictions, while acknowledging that there is still a wealth of information and knowledge available that could be used to support decision making (Bankes, 1993).

Exploratory modeling can be useful when relevant information exists that can be exploited by building models, but where this information is insufficient to specify a single model that accurately describes system behavior. This is known as the non-uniqueness of models, or the problem of under determination (Oreskes et al., 1994). In this circumstance, multiple different models can be constructed that are consistent with the available information. This ensemble of different models typically can capture more of the available information than any of the individual models (Bankes, 2002). The implications of this ensemble of models for potential decisions may be quite diverse. A single model drawn from this potentially infinite set of plausible models is not a "prediction"; rather, it provides a computational experiment that reveals how the world would behave if the various hypotheses encapsulated in this single model about the various unresolvable uncertainties were correct. That is, a model is understood as being a concatenation of hypotheses. These hypotheses include hypotheses about parameter values, mathematical relations between variables, non-linear relations captures in table functions, etc. By conducting many such computational experiments, one can explore the implications of the combinations of these hypotheses. Model development for exploratory modeling aims at the explicit representation of the set of plausible models, through the explication of alternative hypotheses pertaining to parameter values, mathematical relations between variables, non-linear relations captures in table functions, etc. This in turn enables exploiting the information contained in such a set through a large number of computational experiments, the analysis of the results of these experiments, and the use of the set for robust policy design (Bankes, 1993; Hamarat et al., 2013). Thus, in the exploratory modeling literature, like in the system dynamics literature, researchers are arguing for making better use of the available information (Meadows, 1980; Meadows and Robinson, 1985).

For System Dynamics, the implications of adopting an Exploratory Modeling approach are the following. First, the endogenous point of view which is essential to System Dynamics (Richardson, 2011) is maintained. Exploratory modeling does not take a stance on how to describe a system. However, from an exploratory modeling point of view, any given dynamic hypothesis that offers an endogenous explanation for a particular problem is not unique. They are merely instances of a larger set of models that could have been developed. Note that this implication is consistent with SD literature where it is always maintained that models are only plausible. When using exploratory modeling, the process where one moves from mental models and other information about a situation to a single computer simulation model is being problematized. That is, the presence of multiple actors with different mental models and additional information from other sources, it is necessary to explore the extent to which the different mental models agree or are different. If disagreements between mental models become apparent, and these differences cannot be resolved through for example joint sense making as done in group model building (Vennix, 1999), a modeler should encapsulate these differences in the computational models and explicitly explore the implications of these differences on model outcomes.

Adopting an Exploratory Modeling perspective implies that when one is conceptualizing a problem, explicit attention should be given to the presence of diverging understandings of the system of interest. For example, one should assess whether there are alternative reference modes that should be considered. When formulating a dynamic hypothesis there is no guarantee that there is only a single dynamic hypothesis. In the presence of deep uncertainty, it is highly plausible that an ensemble of hypotheses can be articulated. The differences between these different hypotheses might be quite small, say only with respect to the functional form of a non-linear relation. However, it is quite plausible that the differences are more profound, resulting in at least partially disjoint dynamic hypotheses. Regardless, the modeler has to develop the simulation models such that the ensemble of dynamic hypotheses can be explored systematically and thoroughly. Typically, this results in one or more models, each with their associated set of uncertain parameters. The set of uncertainties associated with a given model is called the uncertainty space. For model testing, in addition to the typical questions addressed, specific attention should be given to explore the consequences of uncertainty (Hoffman, 2013): to what extent cover the developed models and their associated uncertainties the space of plausible models. For policy analysis, Exploratory Modeling adds the challenge of ensuring that whatever policy is being put forward produces satisficing results across the ensemble of plausible models.

2.2 Dynamic Pattern Features

Dynamic pattern features as a measure of similarity between two time-series has been proposed by Yücel (2012). He envisioned using it for calibration, validation, and during policy analysis in assessing the performance of alternative policy options. Dynamic pattern features have also been used for dynamic scenario discovery (Kwakkel et al., 2013). In this application, ESDMA is used to generate a wide variety of plausible dynamics of future copper price development. Dynamic pattern features are subsequently used to cluster the resulting time series based on their behavioral similarity. In this paper, we use dynamic pattern features to compare the dynamics resulting from alternative dynamic hypotheses.

The starting point of dynamic pattern features is the idea that a time-series can be decomposed into a sequence of atomic behavior modes (Ford, 1999). The atomic behavior mode is based on the sign (positive, negative, and zero) of the slope and curvature, resulting in nine possible atomic behavior modes. In order to compare the behavioral dynamics of two models, we transform their behavior dynamic into a concatenation of atomic behavior patterns and compare these. More specifically, we transform the time series by determining the sign of both the slope and curvature for each time step. Next, we truncate the resulting feature vector by grouping the atomic behavior patterns. So, if for several sequential time steps the atomic behavior pattern is identical, they are grouped together. Given the feature vectors of two time series, we can now calculate a similarity measure. This similarly is the average deviation across the entire feature vector. For a more elaborate discussion, see Yücel (2012).

3 Case

3.1 Background

In the debate about mineral and metal scarcity most focus is one 'risky' metals, like lithium (Angerer et al., 2009) and the rare earth metals (European Commission, 2011). Only limited attention is given to potential copper scarcity, in spite of today's historically high copper prices (LME, 2011), and the fact that copper is a bulk metal with enormous annual demand (ICSG, 2010a) which, contrary to other bulk metals such as iron and aluminium, could possibly become scarce (Gordon et al., 1987). There seem to be two causes for recent high prices: the growing demand for minerals and metals in rapidly developing economies like China and India (European Commission, 2011) and the growing demand for minerals and metals as a result of energy transitions (Kleijn and van der Voet, 2010). The lack of attention is surprising, given the fact that the future development of copper demand is deeply uncertain, as is the development of the ore grade in relation to mining operations (Gordon et al., 2007; Tilton, 2003; Tilton and Lagos, 2007).

A long tradition of modeling resource depletion and scarcity exists in System Dynamics (SD) modeling. The limits to growth study (Meadows et al., 1972) is probably the most well-known example. Many SD studies combine geological, technological, and economic aspects of mineral depletion (Davidsen et al., 1987; Kwakkel and Pruyt, 2013b; Pruyt, 2010; Sterman and Richardson, 1985; Sterman et al., 1988; Van Vuuren et al., 1999). Other SD studies focus on specific metals, like the platinum group metals (Alonso et al., 2008) or magnesium (Urbance et al., 2002), and are mostly linked to specific metal uses, such as electronics (Alonso et al., 2008) or the automotive industry

(Urbance et al., 2002). Copper markets and their interaction with aluminum markets have been the focus of three master theses in SD (Auping, 2011; Ballmer, 1961; Schlager, 1961).

In spite of the fact that the structure of the copper system is deeply uncertain, is it also well documented: Different perspectives on copper demand –from top-down to bottom-up and from global to regional– are described in the literature (Gordon et al., 2007; Meadows et al., 1982; Tilton and Lagos, 2007). The top-down perspective assumes copper demand is determined by the size of the population and the wealth per capita. In the bottom-up approach, copper demand is determined by different uses and their autonomous development.

Over thirty years ago, Cole already argued that “[w]hether a 'top-down' or 'bottom-up' approach is chosen […] may affect the results[, for s]imple recursive calculation of global or regional aggregates broken down by sector often gives surprisingly different results from systematically building up the global or regional aggregates from the sector or subsector levels” (Meadows et al., 1982). If modeling different perspectives indeed leads to different behavioral patterns, possibly expanding the set of plausible =long-term scenarios of the copper system, then different perspectives may have to be modeled, explored and used. The hypothesis that different models of the copper system generate different behavioral pattern for the same settings and sets of parameter values –and hence, that a multi-model approach is needed– will be tested in this paper by comparing runs generated with three different models of the copper system over the intersection of their input spaces, i.e. with identical settings and values for shared variables and parameters.

3.2 The ensemble of models[1]

In addition to the uncertainty regarding a top-down versus a bottom-up way of modeling the copper system, there are other important uncertainties related to the copper system that should be considered. These include the development of ore grades, energy prices, prices of substitutes, economic growth, infrastructure and capacities, and the resource base. Table 1 specifies how these uncertainties are dealt with. Some of these uncertainties are in turn composed of other deeply uncertain elements, e.g. demand development from a top-down perspective is calculated from global population scenarios (UNPD, 2011), economic development, and the relation between copper demand and GDP per capita (Wouters and Bol, 2009).

Table 1: Major uncertainties in the copper system

Uncertainty	Type of uncertainty	Description
Capacity development	Model uncertainty	The capacity for (deep sea) mines, smelters and refineries
Demand development	Model uncertainty	The intrinsic demand for copper, i.e. the demand without effects due to price and substitution
Economic growth	(Dynamic)	The growth of the GDP globally

[1] The description of the models is kept brief, for our interest is in comparing the results from different models, rather than the specifics of the case. Note that in the final version this section will be expanded. However, we do provide the Vensim models with the paper. Note that the models require Vensim double precision.

	parametric uncertainty	
Ore grade development	Model uncertainty	The ore grade declines with mining of copper, both the speed of this decline and the distribution of ore grades in the lithosphere are uncertain
Life times and construction times	(Dynamic) parametric uncertainty	The lifetimes of facilities and the construction time of facilities
Substitution behavior	(Dynamic) parametric uncertainty	The speed and strength with which substitution and re-substitution take place
Resources/resource base	Model uncertainty	What amount of copper is ultimately recoverable from the earth's crust

The copper demand is modelled as either a function of usages (bottom up) or as a function of population and GDP (top down). The supply chain of copper run runs from the conventional or unconventional resource base to refined copper, to copper in use and discarded copper. Discarded copper can be recycled. Part of the total copper consumption is added to the copper in use, but a relatively large part, 30% in this case, is lost during production and counts as primary scrap. Copper is use is on average scrapped after 50 years. Some of this is collected and recycled some of it is lost. The copper lost during production is completely recycled. The recycling of copper after its end of life depends on the efficiency rate of the recycling. The recycling efficiency rate is calculated by dividing the copper grade in EOL goods by the sum of the conventional copper ore grade and the copper grade in EOL goods. As the recycling adds to the availability of refined copper, it decreases the need for copper extraction. This model structure corresponds to supply chain diagrams in other copper studies (Auping et al., 2012; Glöser et al., 2013; ICSG, 2010b). In case of bottom up way of modeling copper demand, part of the structure shown in Figure 1 is subscripted.

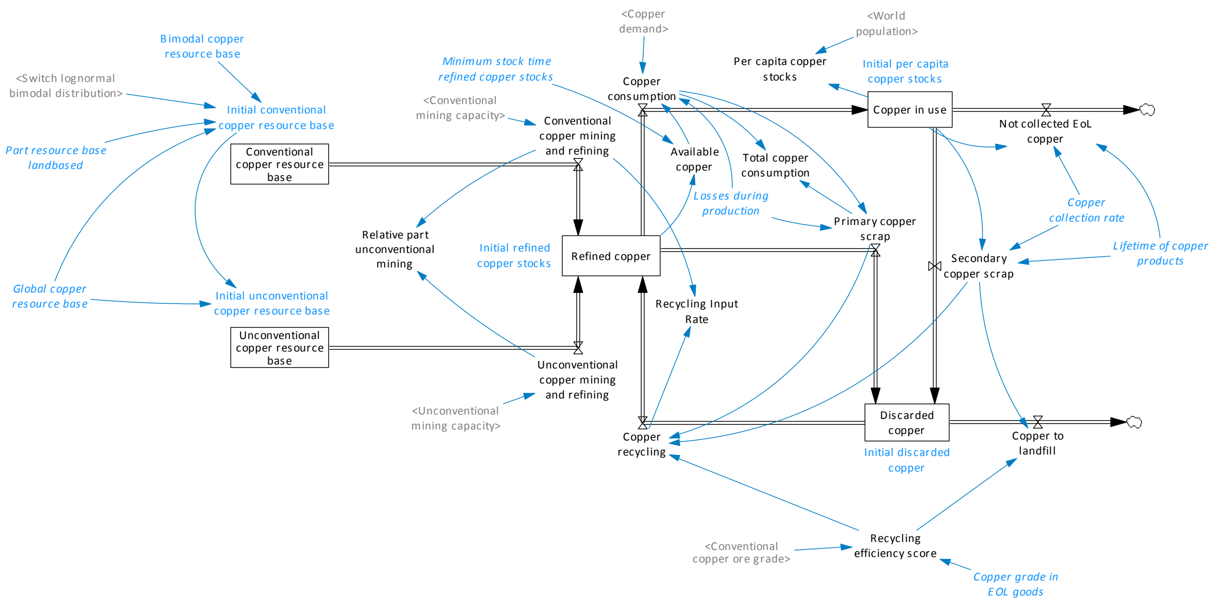

Figure 1. View of the supply chain sub-model

Similar to the supply chain, the extraction submodel is also divided into conventional and unconventional resources. The structure used is essentially the same. New extraction capacity is being developed in response to a shortfall of supply. There is a delay of 10 years before new capacity becomes online. The average lifetime of extraction capacity is 20 years. If demand falls short of supply, capacity can be mothballed. Mothballed capacity can be brought back online if and when necessary, or be decommissioned after prolonged mothballing. The part of new conventional capacity relative to the total new development is determined by the relative attractiveness of conventional resource compared to unconventional resources. This attractiveness is determined by the respective ore grades, taking into account that the energy demand for unconventional resources is ten times as high.

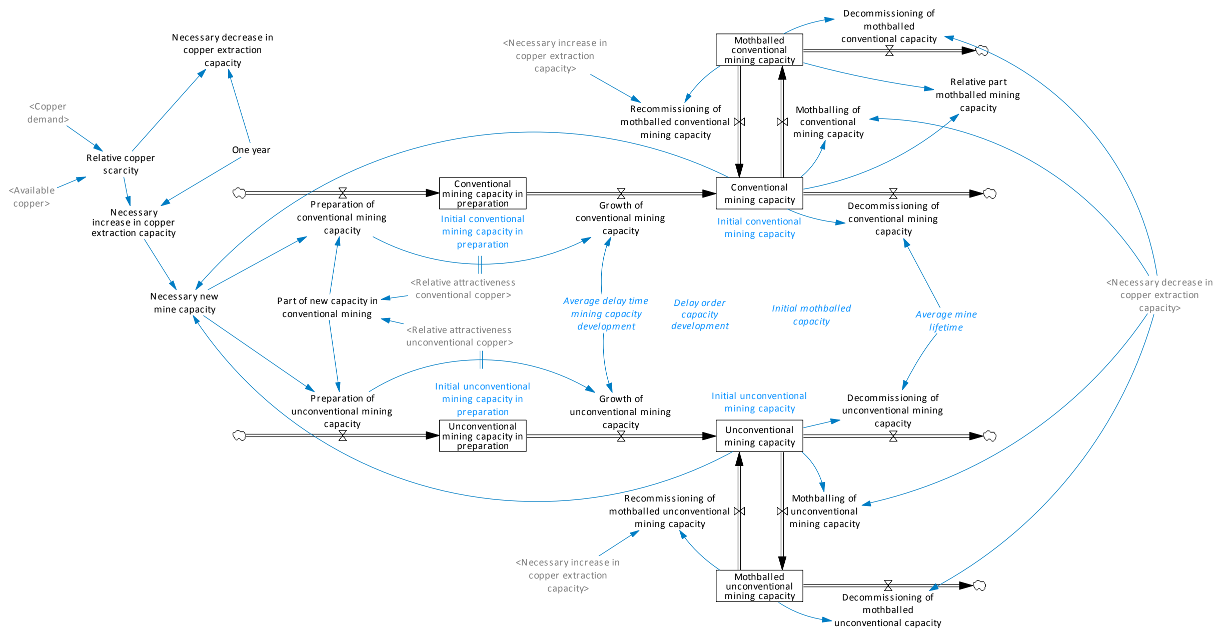

Figure 2. View of the extraction capacity sub-model

4 Results

4.1 Design of computational experiments

To summarize the foregoing, we are using ESDMA to generate an ensemble of simulation runs. This ensemble is composed of two alternative models, each with its associated uncertainty space. Using the dynamic pattern features metric, we can now compare the dynamics for the alternative dynamic hypotheses. Given this comparison, the final step is to explain differences in behavior. Explaining the difference in behavior can be done through differences in parameter values or differences in structure. We can maximize the degree of comparability between two models by making sure that any parameter that exists in both models has the same values. This requires some care in designing the computational experiments that are being used to explore the uncertainty space associated with a given model. In order to maximize the comparability, we first identify which uncertainties the two models are sharing. That is we identify the intersection of the uncertainty spaces and generate computational experiments for this intersection. Next, for both models, we complement these experiments by sampling the model specific uncertainties. This guarantees that uncertainties that exist in both models will have the same values. We use Latin Hypercube sampling for sampling the uncertainties and generated 1000 experiments for each model.

Below the results are shown for three outcomes of interest. We have grouped the outcomes by model. We show, the envelope of outcomes, a few characteristic dynamics within this envelope, and the distribution of outcomes at the end of the runtime using a boxplot. Figure 3 shows the results for the part of the potential copper demand that is being substituted. Over the course of time there is substantial substitution taken place, resulting in 25%-75% substitution in 2050. Although the models differ with respect to the exact values, the behavioral dynamic appears to be identical. That is. Adopting a bottom up or top down perspective on modeling the copper system appears not to affect the behavior of substitution.

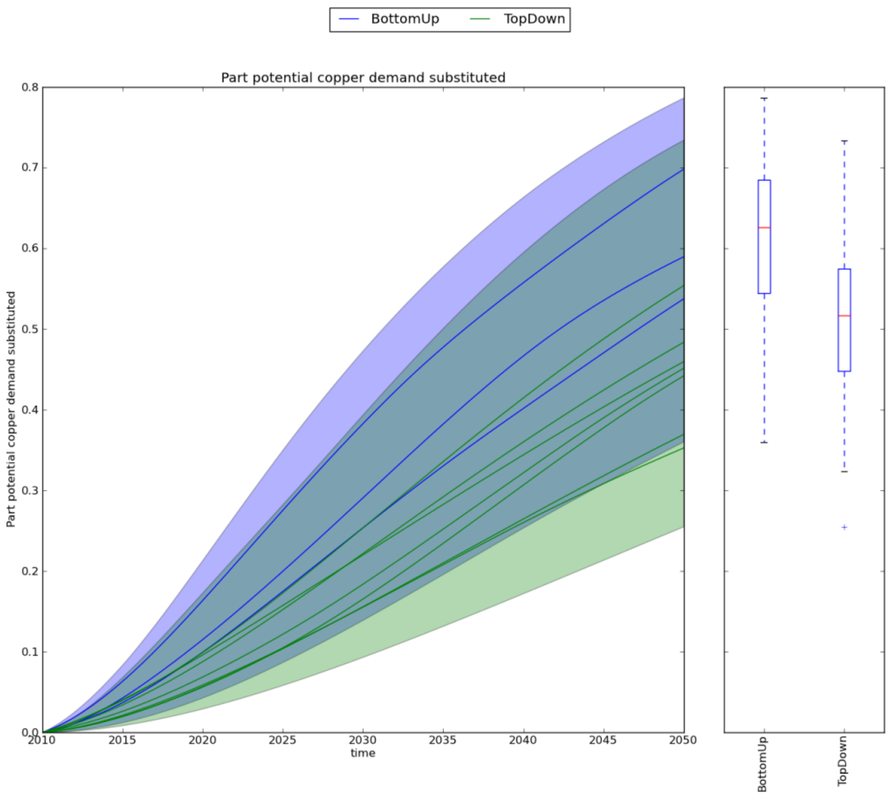

Figure 3. The dynamics for the part of potential copper demand being substituted, grouped by model structure.

To assess whether our visual impression based on Figure 3 is correct, we compare the individual experiments. Recall that the experimental design guaranteed that uncertainties that occur in both models have the same values. This means we can compare individual experiments. We use the dynamic pattern feature metric (Yücel, 2012) to calculate the behavioral distance between the top down model and the bottom up model. Next, we sort these results from low to high. The results of this analysis for the of part potential copper demand substituted is shown in Figure 4. As can be seen, the score is 0 for all experiments, meaning that there is no behavioral difference for any of the experiments. This confirms our visual impression. Adopting a top down or bottom up perspective does not affect the behavioral dynamic of substitution.

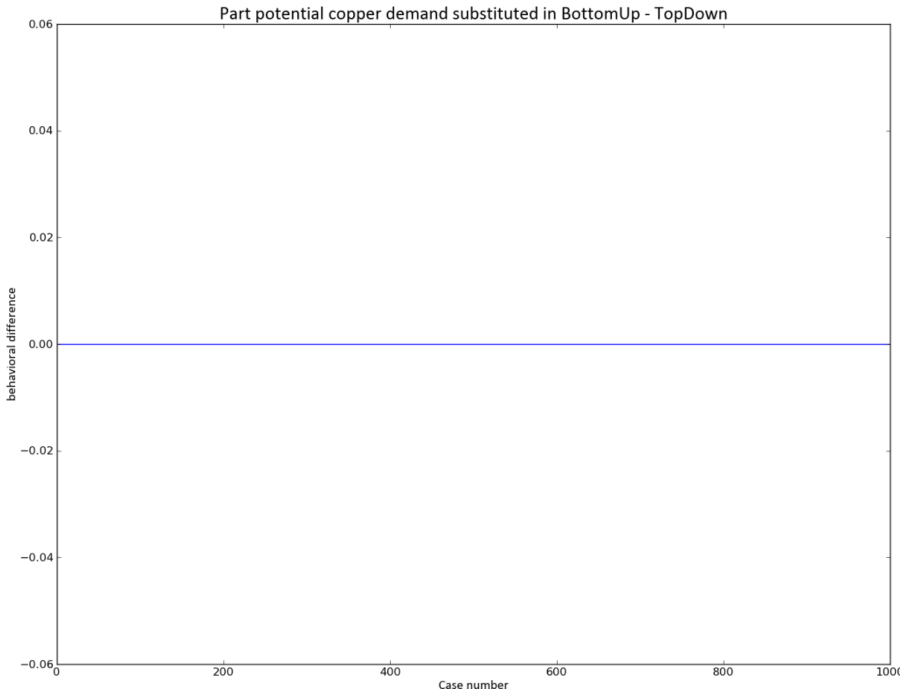

Figure 4. The behavioral difference between the top down model and the bottom up model for the part of the

Next, we look at the unconventional copper ore grade. The envelopes with characteristic dynamics and box plots of the terminal values are shown in Figure 5. These results again appear to be quite similar, although the bottom up model appears to stay stable for a longer duration of the run than the top down model. To assess the degree of behavioral difference, we again use the dynamic pattern feature metric and sort the results. This results in Figure 6. As can be seen a little over 600 experiments are behaviorally identical. To be precise, 397 experiments are behaviorally different.

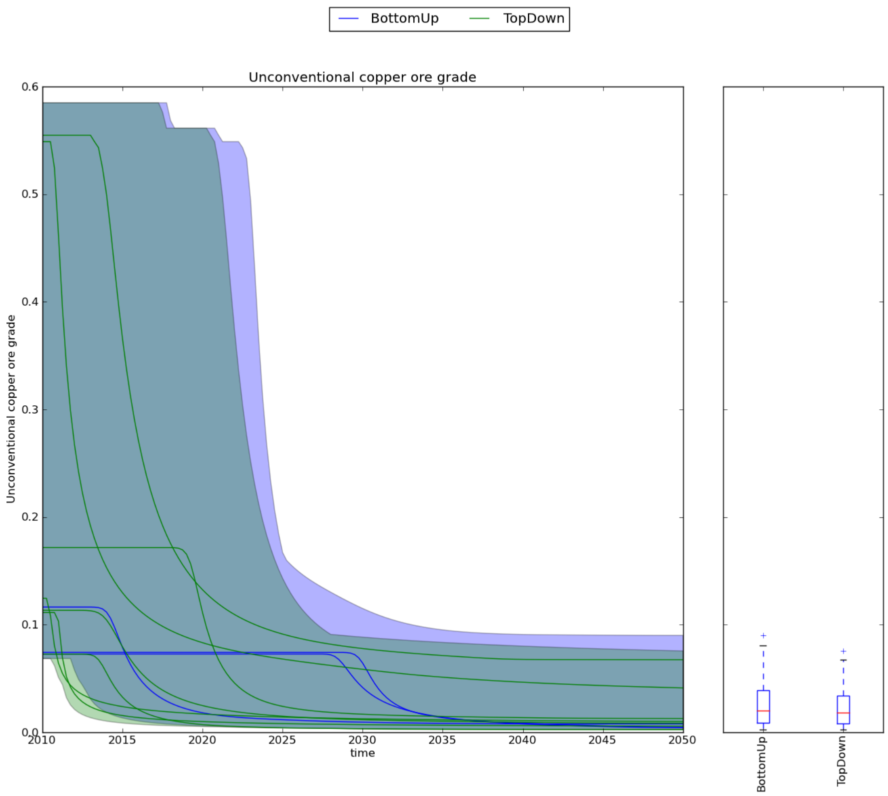

Figure 5. The dynamics for per capita copper stocks, grouped by model structure. potential copper demand substituted. The differences are ordered from small (left) to large (right).

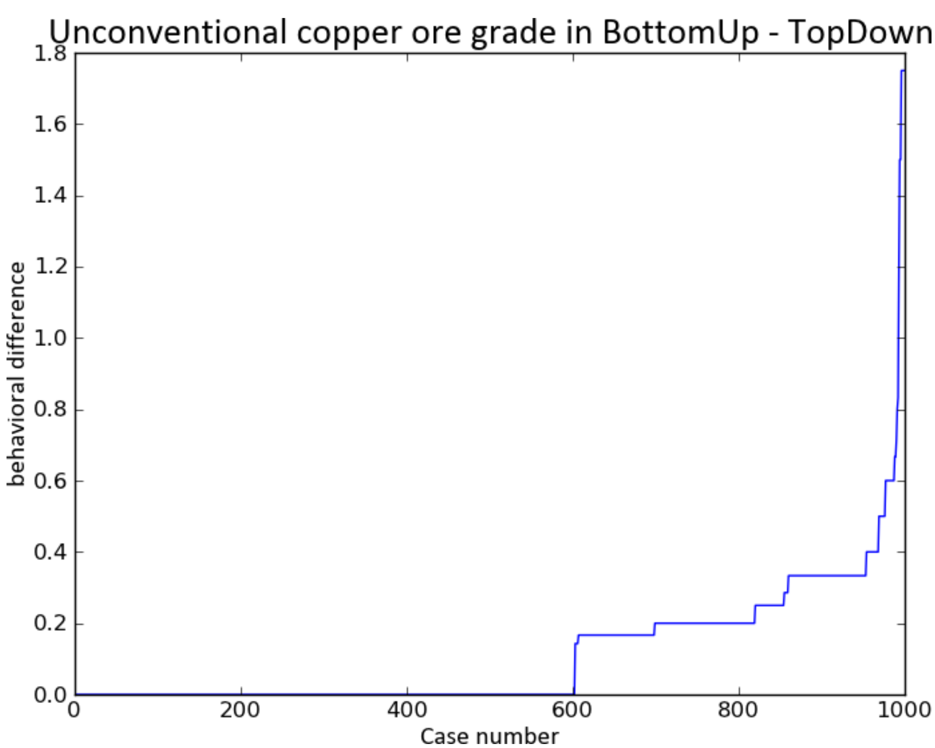

Figure 6. The behavioral difference between the top down model and the bottom up model for the unconventional copper ore grade. The differences are ordered from small (left) to large (right).

Given the design of the computational experiment, we can now try and explain differences in behavior between the two models. We define the model intersection as the intersection of the uncertainty spaces of two models. In set theory, the intersection I of two sets X and Y is defined as the part of the sets that is an element of both sets, hence $I = X \cap Y$. The relative complement C_X of set X in set Y is the part of X that is not part of Y, hence $C_X = X \setminus Y$. The difference D in behaviour between two models with parameter sets A and B can thus be explained both from the complement of the inputs C_A and C_B and the structural differences between the models. The challenge now is to identify subspaces in the uncertainty space that produce different behavior in the two models. To this end, we use the Patient Rule Induction Method (Friedman and Fisher, 1999). PRIM can be used for data analytic questions, where one tries to find combinations of values for input variables that result in similar characteristic values for an outcome of interest. In this particular context, we seek one or more subspaces of the joint uncertainty spaces within which the behavioral difference is larger than 0. PRIM describes these subspaces in the form of hyper-rectangular boxes of the joint uncertainty spaces.

Table 2 shows the results of the PRIM analysis for the behavioral difference between the top down model and the bottom up model for the unconventional ore grade. The table shows the coverage and density metrics for two boxes (Bryant and Lempert, 2010), and the definition of the two boxes. Note that for the definition of the boxes, only the uncertainties that are restricted are shown. Coverage specifies the fraction of experiments that are behaviorally different that are within the identified box. As can be seen, we are able to find a single box that can explain 42% of the behaviorally different cases. Density indicates out of all the experiments that fall within the box, how many are behaviorally different. As can be seen, of all the experiments within box 1, 70% are behaviorally different. Turning to the definition of the box, we see that the first box is primarily defined by the delay order for the capacity development. This means that the behavior between the top down and bottom up model for the unconventional ore grade is most apparent when using a first order delay for the capacity development.

Attempts to find another subspace with a high concentration of behaviorally different results where unsuccessful.

Table 2: Prim results for unconventional copper ore grade

boxes	coverage	density
1	0.42	0.7
rest	0.58	0.3

uncertainty	boxes	
	1	**rest**
Delay order capacity development	1	1, 3, 10, 100
Copper collection rate other uses	0.5-0.78	0.5-0.8

We continue our analysis by looking at a third outcome of interest, namely the recycling input rate. The envelopes with characteristic dynamics and box plots of the terminal values are shown in Figure 7. For this outcome of interest, both the dynamics and the numerical values appear to be quite different. Moreover, the bandwidth of the dynamics is substantially larger for the bottom up model than it is for the top down model. We next calculate the behavioral difference and sort these. The result of this is shown in Figure 8. As can be seen, there is no experiment behaviorally identical. All experiments are different. For some, this difference is quite small, but there is a substantial number for which this behavioral difference is more profound.

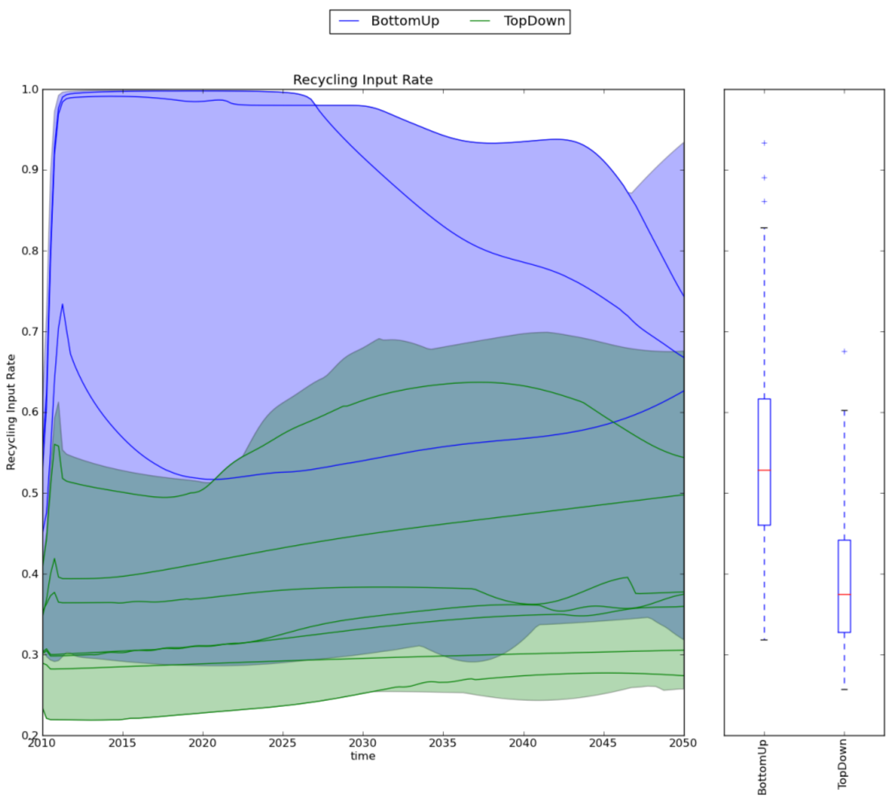

Figure 7. The dynamics for the recycling input rate, grouped by model structure.

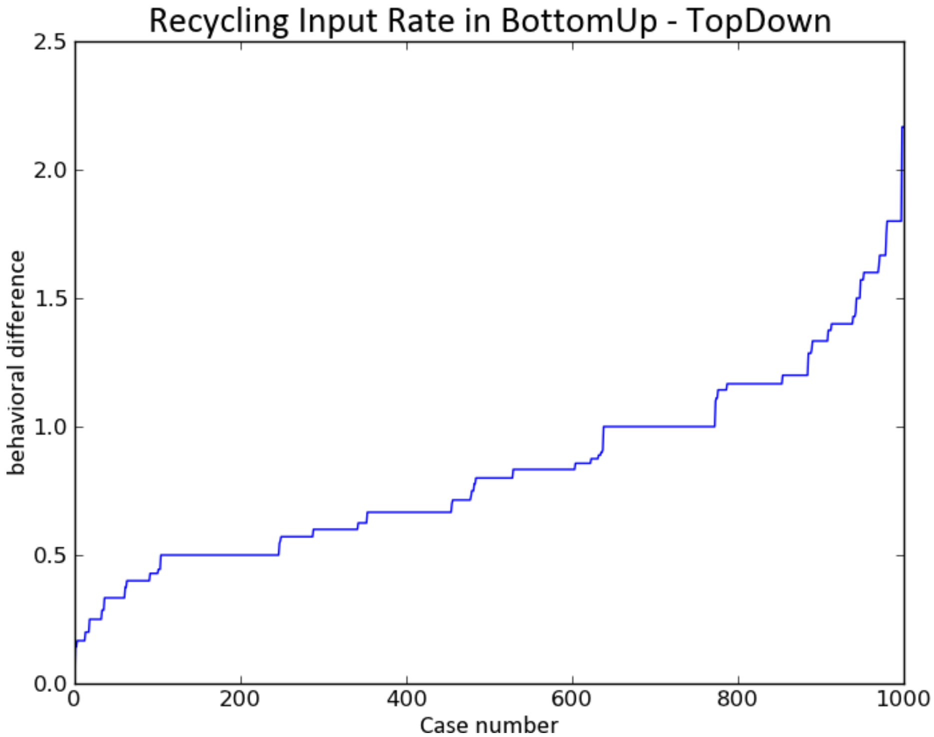

Figure 8. The behavioral difference between the top down model and the bottom up model for the recycling input rate. The differences are ordered from small (left) to large (right).

We again use PRIM to identify one or more subspaces within the model input space where the behavior is substantially different. Given that the behavior is at least slightly different for all experiments, we concentrate our analysis on the experiments that show the largest deviation in behavior. We choose to use a cutoff value of 1, so we try to find one or more subspaces that have a behavioral difference larger than 1. This is a choice of the analyst, and one can experiment with different cutoff values. The results of this analysis are shown in Table 3. We are able to find one subspace that contains 41% of all the cases of interest, with a density of 59%. An attempt to find a second subspace did not yield any conclusive results.

Table 3: Prim results for recycling input rate

box	**coverage**	**density**
1	0.41	0.59
rest	0.59	0.16
uncertainty	**boxes**	
	1	**rest**
Global copper resource base exponent	12,13	12-17
Switch lognormal bimodal distribution	1	1,2
Threshold value aluminum price	1.53-0.2	1.5-2.0
Copper grade in EOL architecture	0.0008-0.00118	0.0008-0.002

5 Discussion and Conclusion

The starting point for this paper was the problem of comparing the simulation results arising from two or more alternative models. This problem can arise in case of offering model based decision support for wicked problems, or problems characterized by deep uncertainty. In such situations, there are various actors involved each with their own understanding of the system. Sometimes these diverging understandings can be resolved through joint sense making, but there is no a priori guarantee that this will always succeed. In such situations the prudent course is to explore the implications of the diverging mental models on the problem at hand and design policies that are insensitive to these differences. A necessary step then becomes comparing the results from the alternative models, understand how there results are different, when there results are different, paving the way for offering a structural explanation of the difference in behavior.

In order to identify the extent to which the behavior arising from two or more models is actually behaviorally different, we adapted the dynamic feature pattern approach of Yücel (2012). This metric first transform a given time-series into a sequence of atomic behavior patterns, and subsequently uses this feature vector when comparing different time series. In order to systematically address deep uncertainties, we used ESDMA Kwakkel and Pruyt (2013a). We demonstrated the approach with a case study of the copper system. There are two dominant alternative perspectives on how to model the copper system: top down or bottom up. We used two models instantiation each of these perspectives and designed a series of computational experiments to systematically explore the dynamics of both models across other key uncertainties. Using the dynamic feature pattern, we looked at the difference in behavior between the two models for three outcomes of interest. We found that for the substitution dynamics, the rival perspectives do not matter behaviorally. For the dynamics of the ore grade of unconventional ore reserves, we found that in many experiments there is no behavioral difference. For almost 40% of the experiments, however, there was a difference. Using the patient rule induction method, we were able to trace back this behavioral difference to an assumption regarding the order of a delay. This offers a clue for offering a more in depth structural explanation of why the models produce different dynamics. The third outcome of interest was related to recycling. Here in all experiments there was a behavioral difference. This suggests that the structural differences between the top down perspective and the bottom up perspective always affect the dynamics of recycling. Again, using PRIM, we traced the most profound behavioral differences to assumptions regarding the amount of available copper ore in the lithosphere and whether this ore is distributed according to a log normal or a bimodal log normal distribution. The behavior between the two models is most pronounced in case of a relatively small amount of ore, which declines relatively quickly. This insight can be used as a starting point for offering a more in depth structural explanation.

In this paper, we have focused on comparing the results from two models and we offered a first step towards explaining this difference in terms of the underlying causal structure. The focus in this paper has been on the approach and the case was meant to illustrate the approach. Future work is needed. A primary direction for future work is to link the insights from PRIM more explicitly to an explanation of the behavioral differences between two models in terms of differences of the underlying feedback structure and delays. The analysis of PRIM offers a direction in the sense that it provides insights into which uncertain assumptions make the behavioral difference stand out the most. As such it provides valuable guidance to the analyst, but only at the level of individual uncertain assumptions. Another direction for further work is to link the presented approach to the design of policies. Knowing when, how, and why to alternative models produce different results is highly relevant for policy analysis. To avoid a policy deadlock due to contested knowledge claims, or inaction grounded in the desire to reduce uncertainty prior to taking a decision, policies should be designed to be insensitive to these uncertainties. Knowing when, how and why models produce different results can help the analyst in designing a policy that in either model produces desirable results, avoiding the policy deadlock. The presented case might be expanded in future work to include a demonstration of this idea.

Reference

Alonso, E., Field, F.R., Kirchain, R.E., (2008) A case study of the availability of platinum group metals for electronics manufacturers, 2008 IEEE International Symposium on Electronics and the Environment. Electronics and the Environment, International Symposium on, pp. 1-6.

Angerer, G., Marscheider-Weidemann, F., Wendl, M., Wietschel, M., (2009) Lithium für Zukunftstechnologien. Fraunhofer ISI, Karlsruhe, p. 58.

Auping, W.L., (2011) The uncertain future of copper: An Exploratory System Dynamics Model and Analysis of the global copper system in the next 40 years, Faculty of Technology, Policy, and Management. Delft University of Technology, Delft.

Auping, W.L., Pruyt, E., Kwakkel, J.H., (2012) Analysing the Uncertain Future of Copper with Three Exploratory System Dynamics Models The 30th International Conference of the System Dynamics Society, St. Gallen.

Ballmer, R.W., (1961) Copper Market Fluctuations: An Industrial Dynamics Study, School of Industrial Management. MIT, Boston.

Bankes, S.C. (1993) Exploratory Modeling for Policy Analysis. Operations Research 4, 435-449.

Bankes, S.C. (2002) Tools and Techniques for Developing Policies for Complex and Uncertain Systems. Proceedings of the National Academy of Sciences of the United States of America 99, 7263-7266.

Bryant, B.P., Lempert, R.J. (2010) Thinking Inside the Box: a participatory computer-assisted approach to scenario discovery. Technological Forecasting and Social Change 77, 34-49.

Cambell, D., Crutchfield, J., Farmer, D., Jen, E. (1985) Experimental Mathematics" the role of computation in nonlinear science. Communications of the ACM 28, 374-384.

Churchman, C.W. (1967) Wicked problems. Guest editorial. Management Science 14.

Dalal, S., Han, B., Lempert, R., Jaycocks, A., Hackbarth, A. (2013) Improving Scenario Discovery using Orthogonal Rotations. Environmental Modelling & Software 48, 49-64.

Davidsen, P.I., Sterman, J.D., Richardson, G.P., (1987) A Petroleum Life Cycle Model for the United States with Endogenous Technology, Exploration, Recovery, and Demand, The 1987 International Conference of the System Dynamics Society, China.

European Commission, (2011) Tackling the challenges in commodity markets and on raw materials, Brussels, p. 22.

Ford, D.N. (1999) A behavioral approach to feedback loop dominance analysis. System Dynamics Review 15, 3-36.

Friedman, J.H., Fisher, N.I. (1999) Bump hunting in high-dimensional data. Statistics and Computing 9, 123-143.

Glöser, S., Soulier, M., Tercero Espinoza, L.A. (2013) Dynamic Analysis of Global Copper Flows. Global Stocks, Postconsumer Material Flows, Recycling Indicators, and Uncertainty Evaluation. Environmental Science & Technology 47, 6564-6572.

Gordon, R.B., Bertram, M., Graedel, T.E. (2007) On the sustainability of metal supplies: A response to Tilton and Lagos. Resources Policy 32, 24-28.

Gordon, R.B., Koopmans, T.C., Nordhaus, W.D., Skinner, B.J. (1987) Towards a New Iron Age? Quantitative Modeling of Resource Exhaustion. Harvard University Press, Cambridge, Massachusetts.

Hallegatte, S., Shah, A., Lempert, R., Brown, C., Gill, S., (2012) Investment Decision Making Under Deep Uncertainty: Application to Climate Change. The World Bank.

Hamarat, C., Kwakkel, J.H., Pruyt, E. (2013) Adaptive Robust Design under Deep Uncertainty. Technological Forecasting and Social Change 80, 408-418.

Hoffman, M. (2013) Simulation-based exploratory data generation and analysis (data farming): a critical reflection on validity and methodology. The journal of defense modeling and simulation: applications, methodology, technology, 1-13.

Homer, J. (1996) Why we iterate: scientific modeling in theory and practice. System Dynamics Review 12, 1-19.

Homer, J. (2013) The aimless plateau, revisted: why the field of system dynamics needs to establish a more coherent identity. System Dynamics Review 29, 124-127.

ICSG, (2010a) Release of ICSG 2010 Statistical Yearbook, ICSG Press Release. ICSG.

ICSG (2010b) The World Copper Factbook, Lisbon.

Kleijn, R., van der Voet, E. (2010) Resource constraints in a hydrogen economy based on renewable energy sources: An exploration. Renewable and Sustainable Energy Reviews 14, 2784-2795.

Kriegler, E., Petermann, N., Krey, V., Schwanitz, V.J., Luderer, G., Ashina, S., Bosetti, V., Eom, J., Kitous, A., Méjean, A., Paroussos, L., Sano, F., Turton, H., Wilson, C., van Vuuren, D.P. (2014) Diagnostic indicators for integrated assessment models of climate policy. Technological Forecasting and Social Change.

Kwakkel, J.H., Auping, W.L., Pruyt, E. (2013) Dynamic scenario discovery under deep uncertainty: the future of copper. Technological Forecasting and Social Change 80, 789-800.

Kwakkel, J.H., Pruyt, E. (2013a) Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. Technological Forecasting and Social Change 80, 419-431.

Kwakkel, J.H., Pruyt, E. (2013b) Using System Dynamics for Grand Challenges: The ESDMA Approach. Systems Research and Behavioral Science.

Kwakkel, J.H., Walker, W.E., Marchau, V.A.W.J. (2010) Classifying and communicating uncertainties in model-based policy analysis. International Journal of Technology, Policy and Management 10, 299-315.

Lane, D.C. (2000a) Diagramming conventions in System Dynamics. The journal of the Operational Research Society 51, 241-245.

Lane, D.C. (2000b) Should System Dynamics be described as a 'hard' or 'deterministic' systems approach? Systems Research and Behavioral Science 17, 3-22.

Lane, D.C. (2012) What is a 'policy insight'? Systems Research and Behavioral Science 29, 590-595.

Lempert, R.J., Collins, M. (2007) Managing the Risk of Uncertain Threshold Response: Comparison of Robust, Optimum, and Precautionary Approaches. Risk Analysis 24, 1009-1026.

Lempert, R.J., Groves, D.G., Popper, S., Bankes, S. (2006) A general analytic method for generating robust strategies and narrative scenarios. Management Science 52, 541-528.

Lempert, R.J., Popper, S., Bankes, S., (2003) Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis. RAND, Santa Monica, CA, USA.

LME, (2011) Copper graphs. London Metal Exchange, London.

Meadows, D.H., (1980) The unavoidable a priori, in: Randers, J. (Ed.), Elements of the System Dynamics Method. Productivity Press, Cambridge, Massachusetts.

Meadows, D.H., Meadows, D.L., Randers, J., Behrens III, W.W. (1972) The limits to growth. Universe Books.

Meadows, D.H., Richardson, J., Bruckmann, G., (1982) Groping in the dark. J.W. Arrowsmith Limited, Bristol.

Meadows, D.H., Robinson, J. (1985) The Electronic Oracle. Computer Models and Social Decisions. Joh Wiley & Sons.

Oreskes, N., Shrader-Frechette, K., Belitz, K. (1994) Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. Science 263, 641-647.

Pruyt, E., (2010) Scarcity of Minerals and Metals: A Generic Exploratory System Dynamics Model, 18th International Conference of the System Dynamics Society, Seoul, Korea.

Richardson, G.P. (2011) Reflections on the foundations of system dynamics. System Dynamics Review 27, 219-243.

Rittel, H.W.J., Webber, M.W. (1973) Dilemmas in a General Theory of Planning. Policy Sciences 4, 155-169.

Schlager, K.J., (1961) A Systems Analysis of the Copper and Aluminum Industries: An Industrial Dynamics Study, School of Industrial Management. MIT, Boston.

Sterman, J.D. (2000) Business Dynamics: Systems Thinking and Modeling for a Complex World. McGraw-Hill.

Sterman, J.D., Richardson, G.P. (1985) An experiment to evaluate models for estimating fossil fuel resources. Journal of Forecasting 4, 197-226.

Sterman, J.D., Richardson, G.P., Davidsen, P.I. (1988) Modeling the estimation of petroleum resources in the United States. Technological Forecasting and Social Change 33, 219-249.

Tilton, J.E. (2003) On Borrowed Time? Assessing the Threat of Mineral Depletion. RFF Press, Washington, DC.

Tilton, J.E., Lagos, G. (2007) Assessing the long-run availability of copper. Resources Policy 32, 19-23.

UNPD, (2011) World Population Prospects: The 2010 Revision. Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York.

Urbance, R., Field, F.R., Kirchain, R.E., Roth, R., Clark, J. (2002) Market model simulation: The impact of increased automotive interest in magnesium. JOM Journal of the Minerals, Metals and Materials Society 54, 25-33.

Van Vuuren, D.P., Strengers, B.J., De Vries, H.J.M. (1999) Long-term perspectives on world metal use - a system dynamics model. Resources Policy 25, 239-255.

Vennix, J.A.M. (1999) Group model-building: tackling messy problems. System Dynamics Review 15, 37-401.

Wouters, H., Bol, D., (2009) Material Scarcity. An M2i study. Stichting Materials innovation institute, Delft, p. 72.

Yücel, G., (2012) A novel way to measure (dis)similarity between model behaviors based on dynamic pattern features, The 30th International Conference of the System Dynamics Society, St. Gallen, Switzerland.

Yücel, G., Barlas, Y. (2011) Automated parameter specification in dynamic feedback models based on behavior pattern features. System Dynamics Review 27, 195-215.