Uncertainty assessment of the IMAGE/TIMER B1 CO_2 emissions scenario using the NUSAP method

Jeroen P. van der Sluijs[1], Jose Potting[1,2], James Risbey[1], Detlef van Vuuren[2], Bert de Vries[2], Arthur Beusen[2], Peter Heuberger[2], Serafin Corral Quintana[3], Silvio Funtowicz[3], Penny Kloprogge[1], David Nuijten[1], Arthur Petersen[5], Jerry Ravetz[4]

Universiteit Utrecht

[1] *Department of Science Technology and Society, Utrecht University, j.p.vandersluijs@chem.uu.nl;*
[2] *National Institute of Public Health and the Environment. Bilthoven;*
[3] *EC Joint Research Centre, Institute for Systems, Informatics and Safety, Ispra (VA), Italy;*
[4] *The Research Method Consultancy, London;*
[5] *Faculty of Philosophy, Free University of Amsterdam*

Introduction

This project implemented a novel approach to uncertainty assessment, known as the NUSAP method (Numeral Unit Spread Assessment Pedigree) to assess qualitative and quantitative uncertainties in the TIMER energy model, part of RIVMs IMAGE Model. We used the IMAGE B1 scenario produced for the IPCC as case study.

Objective

Develop a framework for uncertainty assessment and management including both quantitative and qualitative dimensions, and test and demonstrate its usefulness in integrated assessment models.

- What are key uncertainties in TIMER?
- What is the role of model structure uncertainties in TIMER?
- Uncertainty in which input variables and parameters dominate uncertainty in model outcome?
- What is the strength of the sensitive parameters (pedigree)?

Method

The set of methods by which we implemented NUSAP in this project include:

- A comprehensive checklist for model quality assistance;
- A meta-level analysis of the results of the six SRES energy models, to explore model structure uncertainties;
- The Morris algorithm for global sensitivity analysis;
- A NUSAP expert elicitation workshop to systematically assess the pedigree of sensitive parameters
- A diagnostic diagram to prioritise uncertainties by the combination of criticality (based on Morris) and parameter strength (based on pedigree).

Table 1 Pedigree matrix for parameter strength. Note that the columns are independent.

Score	Proxy	Empirical	Theoretical basis	Method	Validation
4	Exact measure	Large sample direct mmts	Well established theory	Best available practice	Compared with indep. mmts of same variable
3	Good fit or measure	Small sample direct mmts	Accepted theory partial in nature	Reliable method commonly accepted	Compared with indep. mmts of closely related variable
2	Well correlated	Modeled/ derived data	Partial theory limited consensus on reliability	Acceptable method limited consensus on reliability	Compared with mmts not independent
1	Weak correlation	Educated guesses / rule of thumb est	Preliminary theory	Preliminary methods unknown reliability	Weak / indirect validation
0	Not clearly related	Crude speculation	Crude speculation	No discernible rigour	No validation

NUSAP expert elicitation workshop, Loosdrecht, June 12 and 13 2001

Results

TIMER is a non-linear model containing a large number of input variables, all liable to uncertainty of different orders of magnitude. A proper sensitivity analysis asks in such situation for an approach that covers the entire range of possible values for a given input variable. The Morris method (1991) facilitates such global sensitivity analysis in a minimum number of model runs. The analysis differentiated clearly between sensitive and less sensitive model components. The most sensitive turned out to be:

- Population levels and economic activity as main drivers;
- Variables related to intra-sectoral structural change;
- Progress ratios for technological improvements;
- Variables related to resources of fossil fuels (size and cost supply curves);
- Variables related to autonomous and price-induced energy efficiency improvement;
- Variables related to initial costs and depletion of renewables;

We assessed pedigree and value loading of these model components in a NUSAP expert elecitation (see phote) with 18 participants (in three parallel sessions).

For some parameters, we found a reasonable consistency in pedigree scores across the group results while for others, there was a considerable disagreement. The average scores were mostly in the middle ranges of the strenght matrix. An example result is presented in fig. 1.

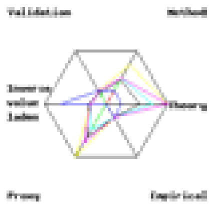


Fig. 1a Example of radar diagram of the gas depletion multiplier assessed by 6 experts

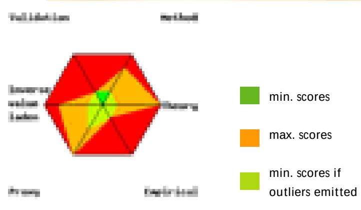
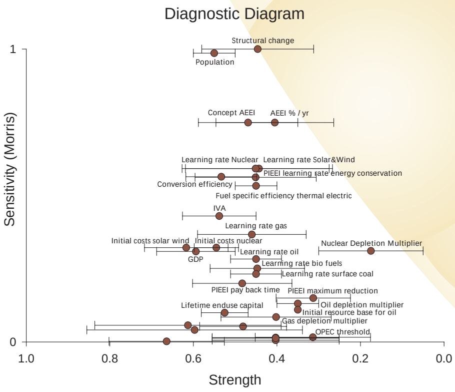


Fig 1b same, but represented as kite diagram.

Average scores for all five pedigree criteria over all parameters range from low (validation 1.1) to medium (empirical: 1.8, method 1.8, theory 2.0 and proxy 2.4, all on a scale from 0 to 4). The slightly higher score for theoretical understanding compared to empirical basis combined with the consistently low scores for validation nicely reflect the inherent theory ladeness of scenario studies, in this case based on not fully crystallised theory.



Findings for sensitivity and pedigree were combined in a diagnostic diagram (fig. 2). The Y-axis plots contribution to change in projected CO_2 emissions found with the sensitivity anaysis. The X-axis displays normalized average pedigree scores for each variable. The error bars about these values (one standard deviation) reflect expert disagreement on pedigree scores. The scale goes from 1 at the origin to zero on the right, placing the more `dangerous' variables in the top right quadrant of the plot.

Conclusions

- Our model quality assurance checklist proves a quick scan to flag major areas of concern and associated pitfalls in the complex mass of uncertainties.
- The meta-level intercomparison of TIMER with the other SRES models gave us some insight into the potential roles of model structure uncertainties.
- Global sensitivity analysis supplemented with expert elicitation constitutes an efficient selection mechanism to further focus the diagnosis of key uncertainties.
- Our pedigree elicitation procedure yields a differentiated insight into parameter strength.
- The diagnostic diagram puts spread and strength together to provide guidance in prioritisation of key uncertainties.

Overall, the project demonstrated that the NUSAP method can be applied to complex models in a meaningful way. The method provides a useful means to focus research efforts on the potentially most problematic parameters while it at the same time pinpoints specific weaknesses in these parameters.