Examples of applications of the NUSAP approach

Dr. Jeroen van der Sluijs

j.p.vandersluijs@uu.nl

Copernicus Institute for Sustainable Development and Innovation
Utrecht University

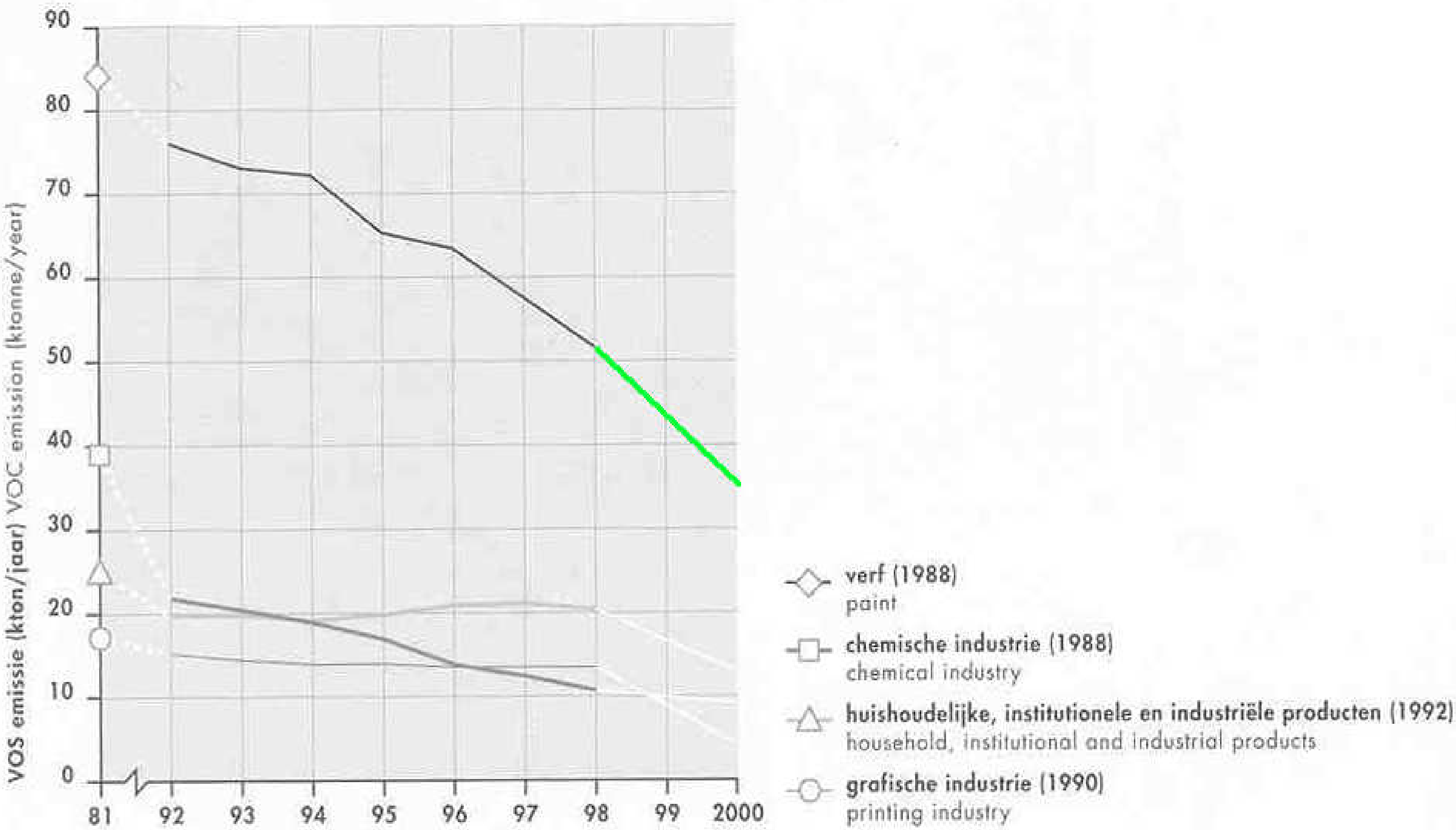
&

Centre d'Economie et d'Ethique pour l'Environnement
et le Développement,
Université de Versailles Saint-Quentin-en-Yvelines, France

Case 1

VOC emissions from paint in the Netherlands

figuur 3 **emissieverloop verf, chemische industrie, HIIP en grafisch industrie**
emission trends for paint, chemical industry, HIIP and printing industry



How is VOC from paint monitored?

VOC emission calculated from:

- VVVF national sales statistics NL-paint in NL per sector
- CBS paint import statistics
- Estimates of paint-related thinner use
- Assumption of VOC% imported paint
- Attribution imported paint over sectors

General steps in analysis

- Identification and classification
- Disaggregation
- Identification of expertise
- Assessment of assumptions
- Qualitative assessment
- Quantitative assessment
- Uncaptured assumptions
- Calculation of Monte Carlo and Pedigree results
- Communication

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

Expert Elicitation

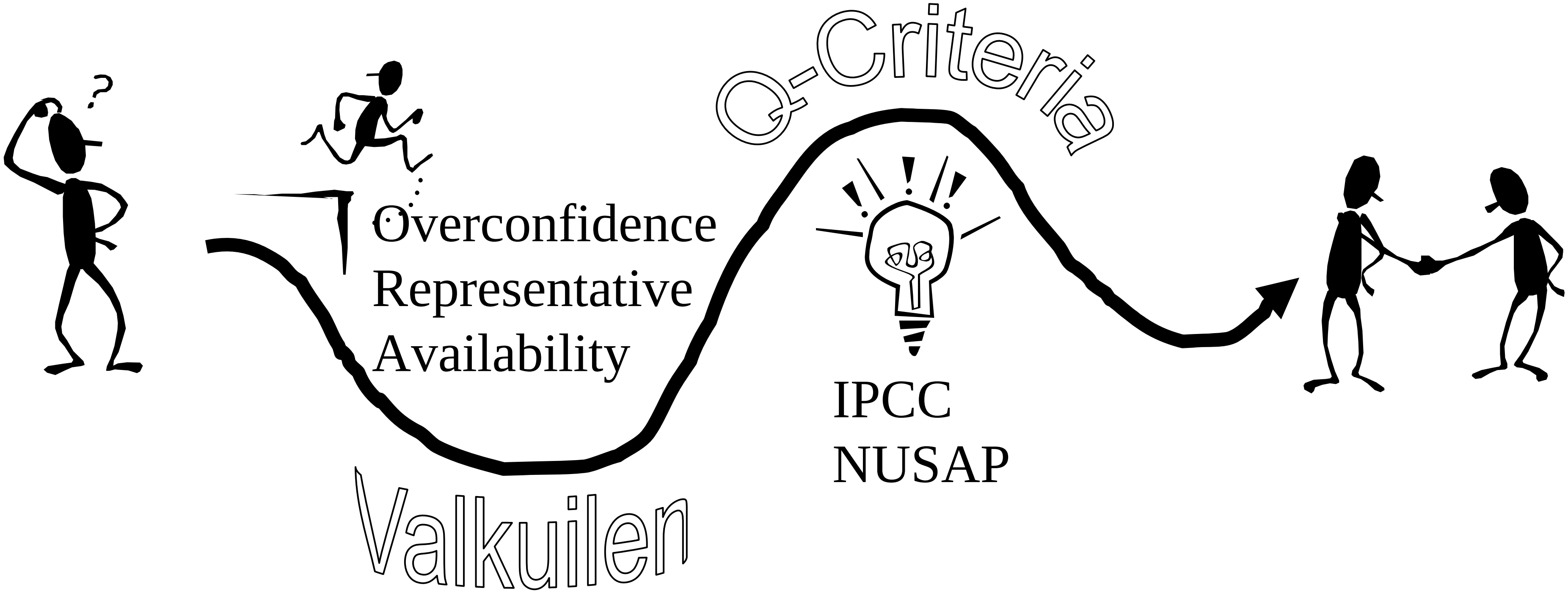
Goal

- To systematically make explicit and utilizable unwritten knowledge in the heads of experts, including insight in the limitations, strengths and weaknesses of that knowledge

Pitfalls in expert elicitation

- Overconfidence
- Representativeness
- Anchoring
- Bounded rationality
- Availability / lamp posting
- Implicit assumptions
- Motivational bias
  - Possibility of strategic answers
  - Interests with regard to outcome of analysis

Expert Elicitation proces



Sources of error

- Definitional inconsistency
- Interpretation of definitions
- Boundaries between raw materials, products, assortment
- Miscategorization
- Misreporting via unit confusion
- Deliberate misreporting
- Miscoding
- Non-response

Sources of error (continued)

- Not counting small firms (reporting threshold CBS)
- Not counting non-VVVF members
- Firm dynamics
- Paint dynamics
- Computer code errors

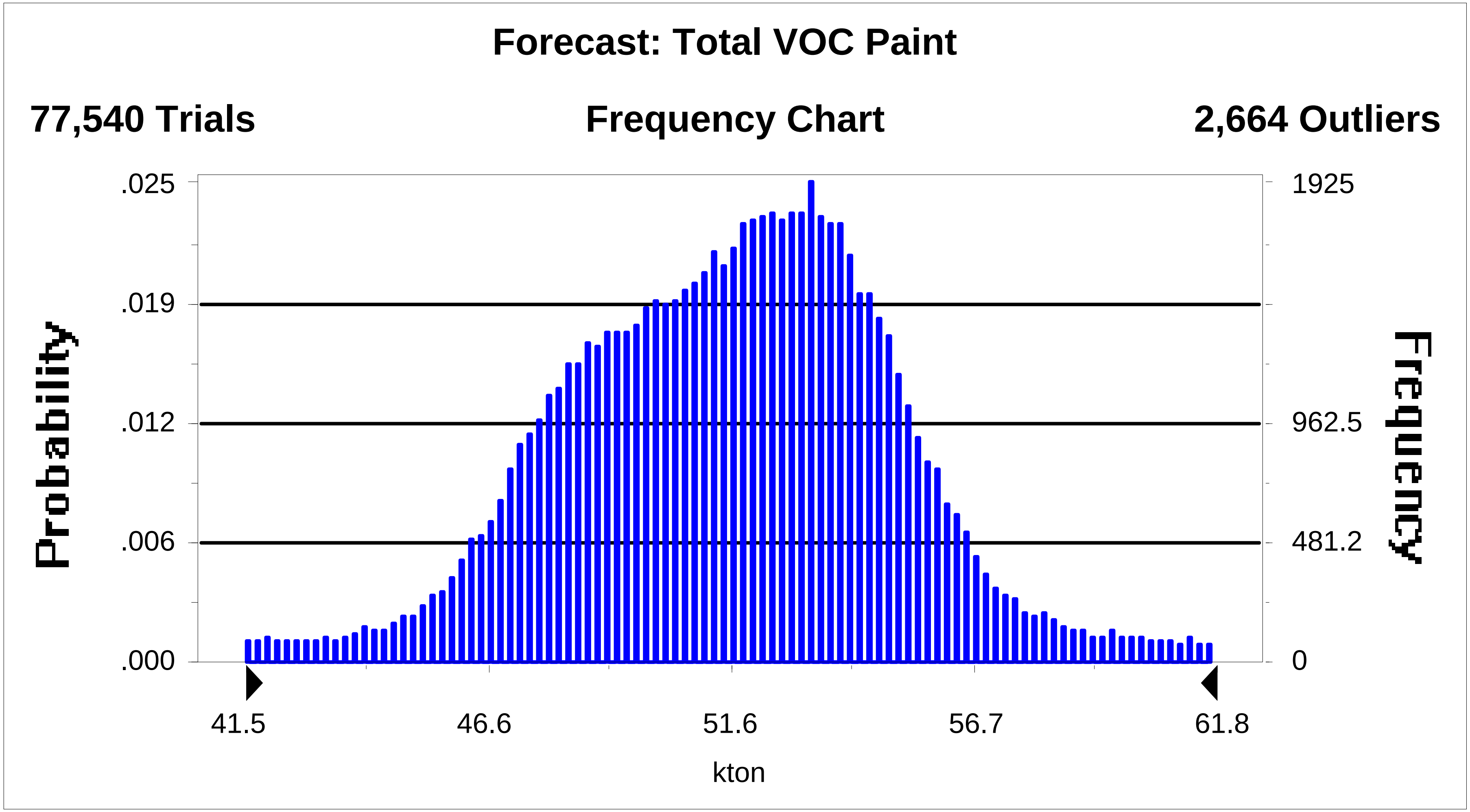
Conceivable sources of motivational bias

- Cross-check with CBS data by tax authorities
- Anonymity
- Membership due VVVF depends on sales figures
- Paint sales statistics is market-sensitive information; possibility of strategic reporting to mislead competitors
- KWS2000 - reputation of paint industry

Added during the elicitation

- Import below reporting threshold
- Gap between NS VVVF-members and total NS
- Overlap CBS import figure / VVVF NS figure

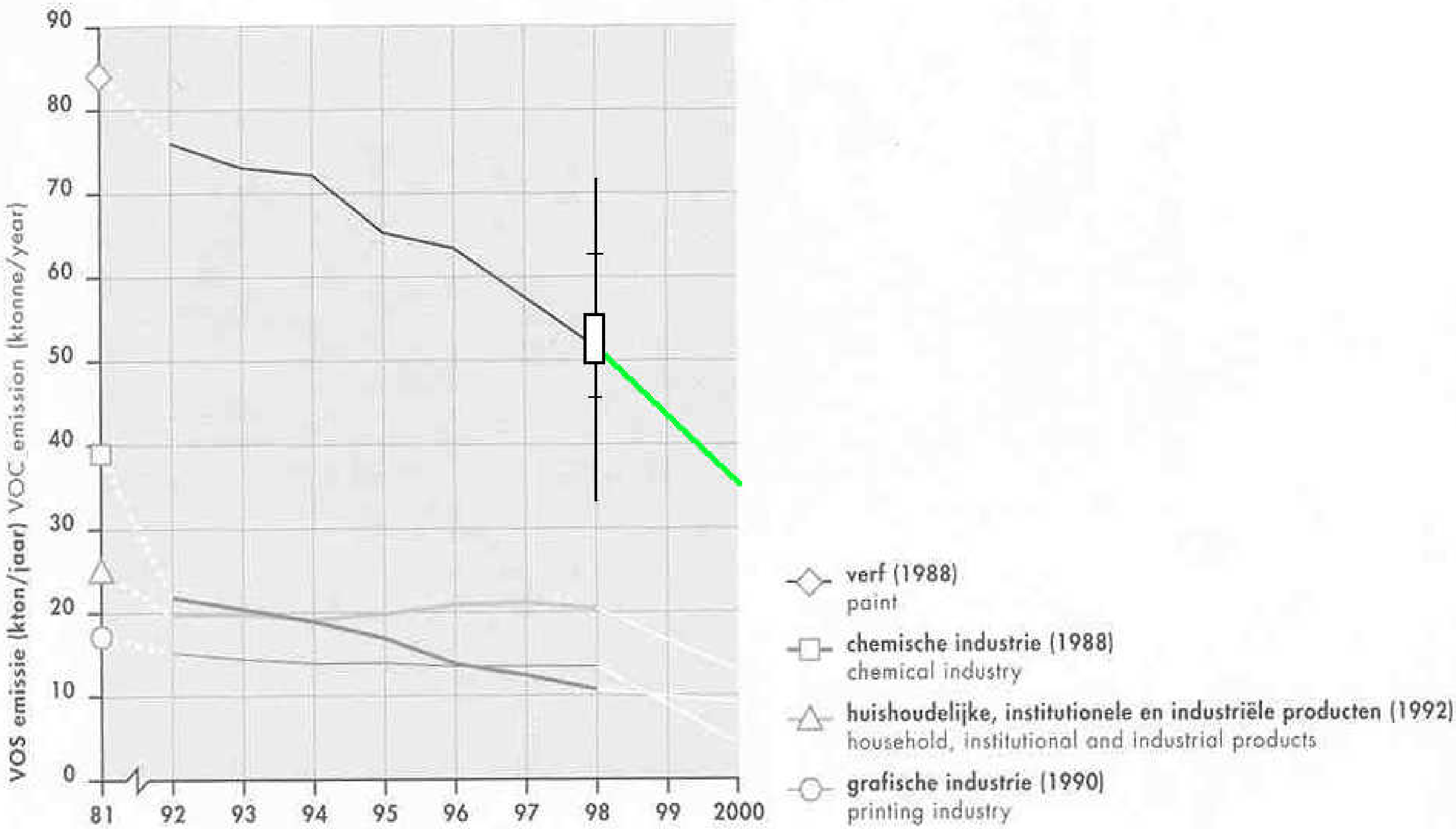
- Attribution of Thinner % DIY & Decoration to paint



Forecast: Total VOC Paint
77,540 Trials
Frequency Chart
2,664 Outliers
Probability
Frequency
.025
.019
.012
.006
.000
1925
962.5
481.2
0
41.5
46.6
51.6
56.7
61.8
kton

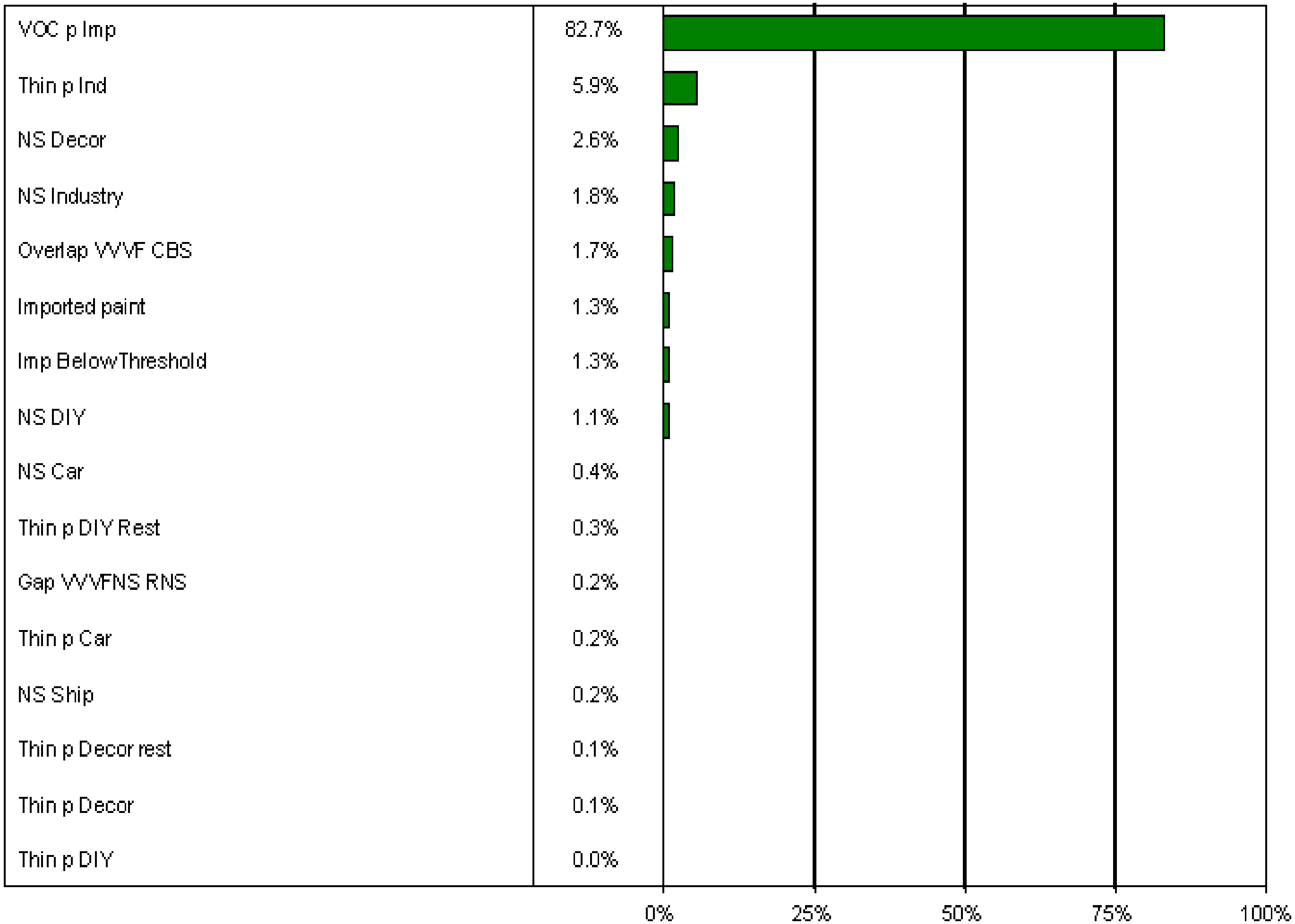
figuur 3 emissieverloop verf, chemische industrie, HIIP en grafisch industrie

emission trends for paint, chemical industry, HIIP and printing industry



Sensitivity Chart

Target Forecast Total VOC Paint

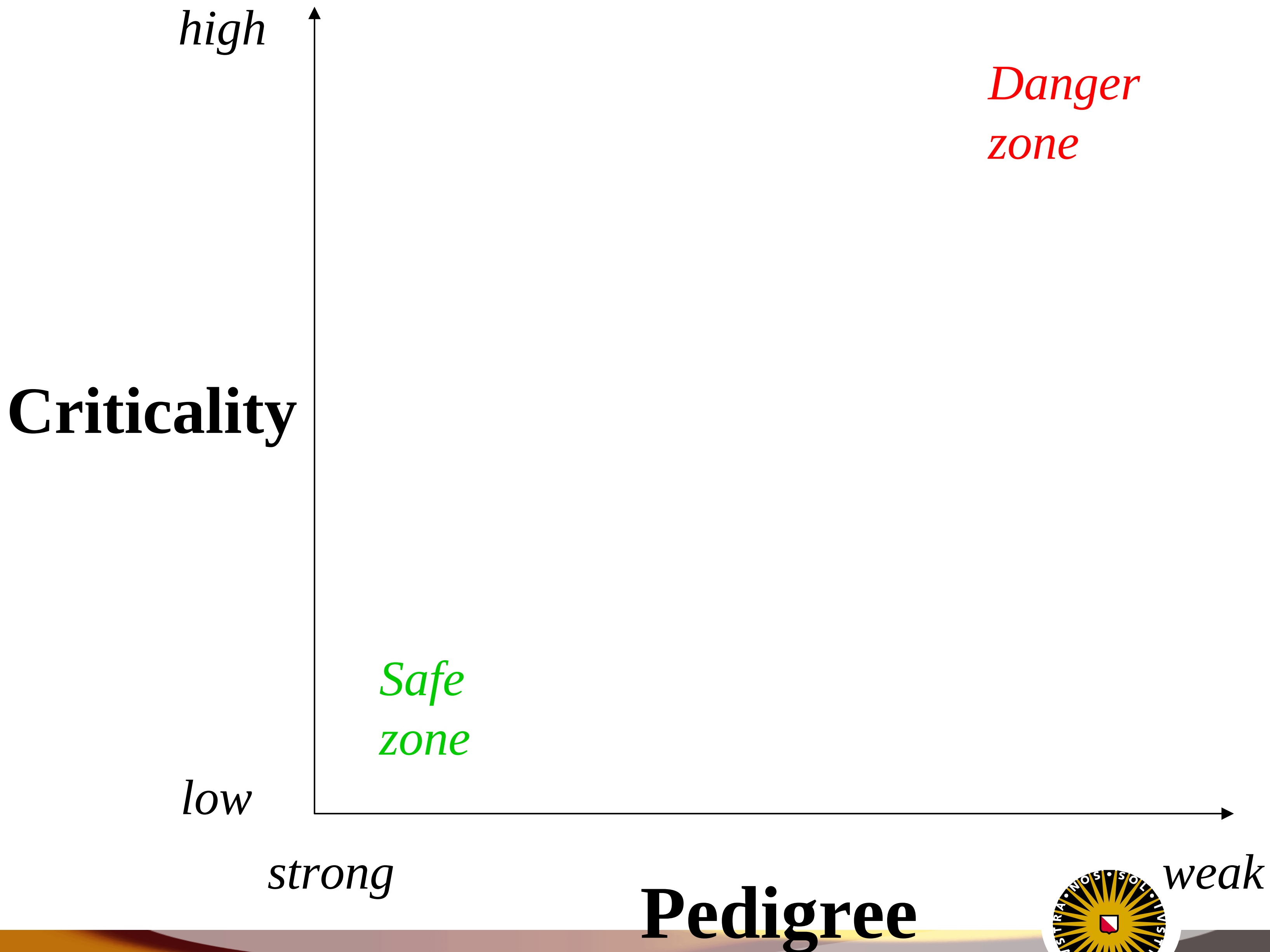


Pedigree scores

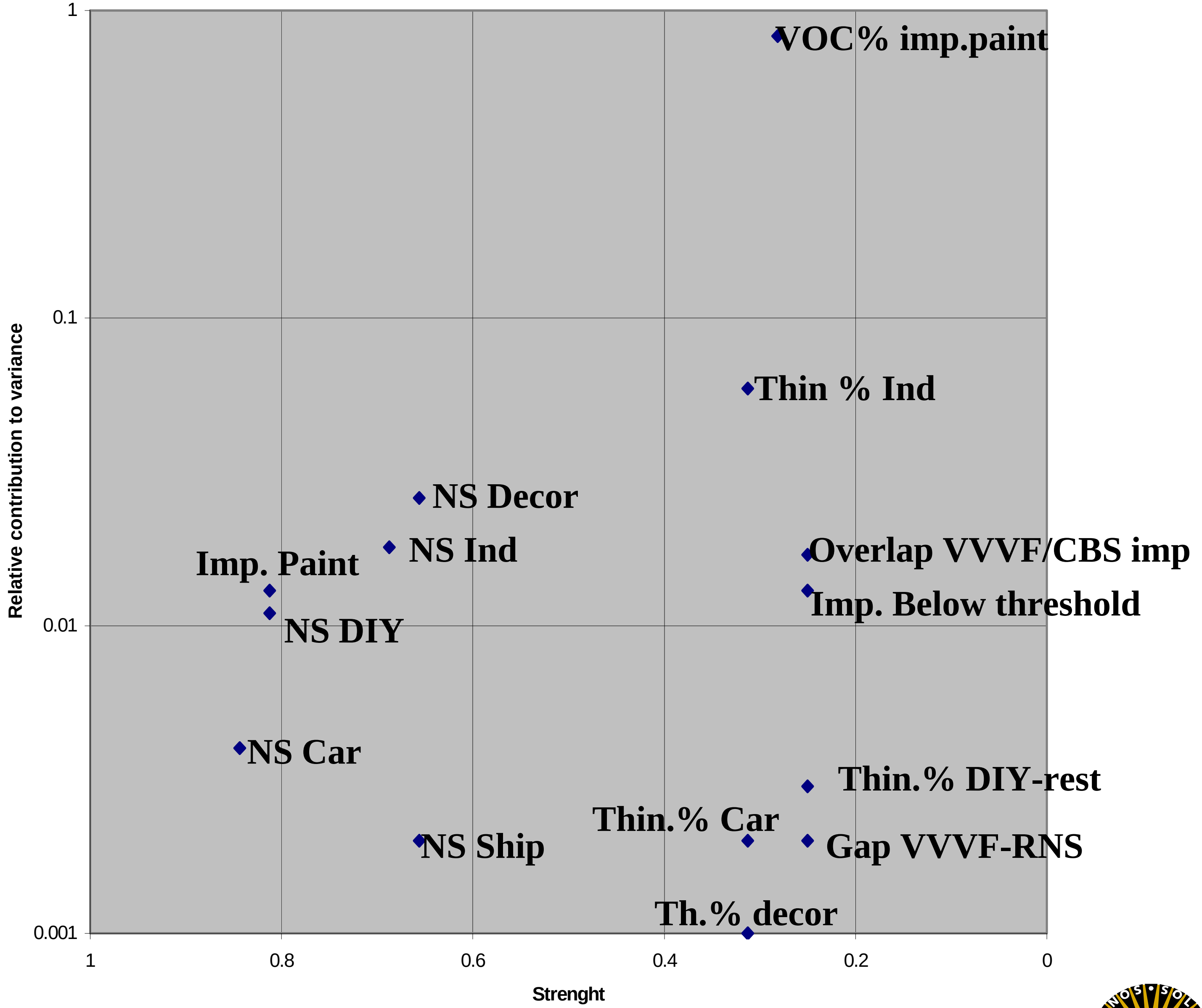
	Proxy	Empirical	Method	Validation	Strength
NS-SHI	3	3.5	4	0	0.66
NS-B&S	3	3.5	4	0	0.66
NS-DIY	2.5	3.5	4	3	0.81
NS-CAR	3	3.5	4	3	0.84
NS-IND	3	3.5	4	0.5	0.69
Th%-SHI	2	1	2	0	0.31
Th%-B&S	2	1	2	0	0.31
Th%-DIY	1	1	2	0	0.25
Th%-CAR	2	1	2	0	0.31
Th%-IND	2	1	2	0	0.31
VOS % import	1	2	1.5	0	0.28
Attribution import	1	1	2	0	0.25

Trafic-light analogy <1.4 red; 1.4-2.6 amber; >2.6 green

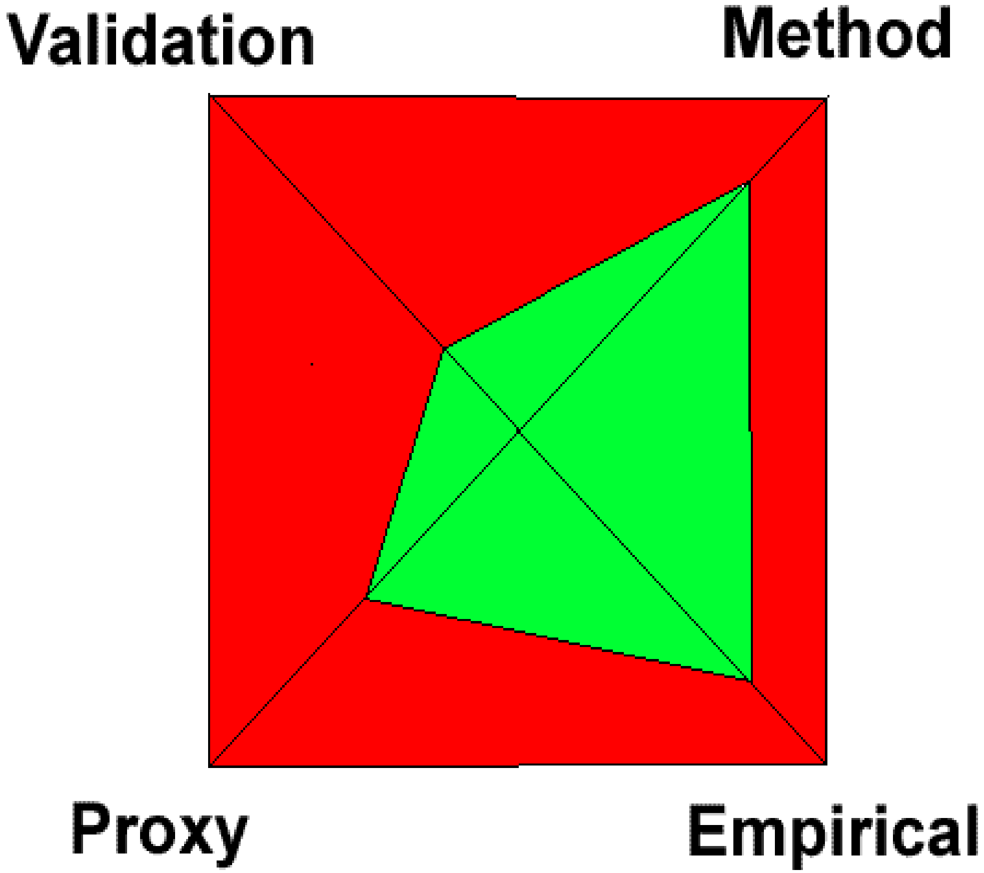
NUSAP Diagnostic Diagram



NUSAP Diagnostic Diagram



Pedigree 1998 VOC emission from Paint



Conclusions

- NUSAP provides a strong diagnostic tool which yields a richer insight in sources and nature of uncertainty than Monte Carlo analysis alone
- Helpful in priority setting for uncertainty management and quality improvement
- Time intensive
  - need for two-step approach: quick-scan / thorough analysis

Case 2

The IMAGE/TIMER B1 scenario

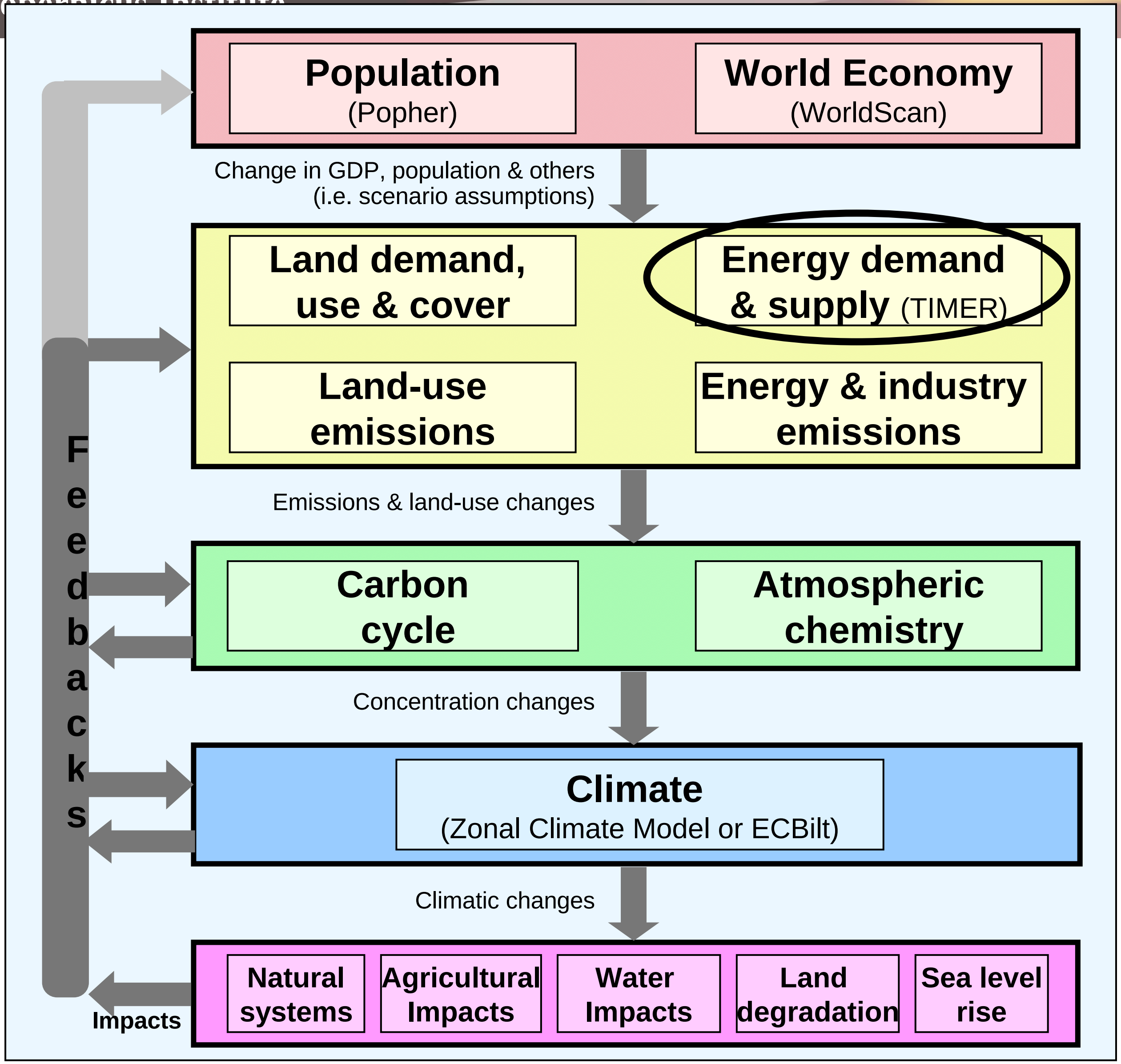
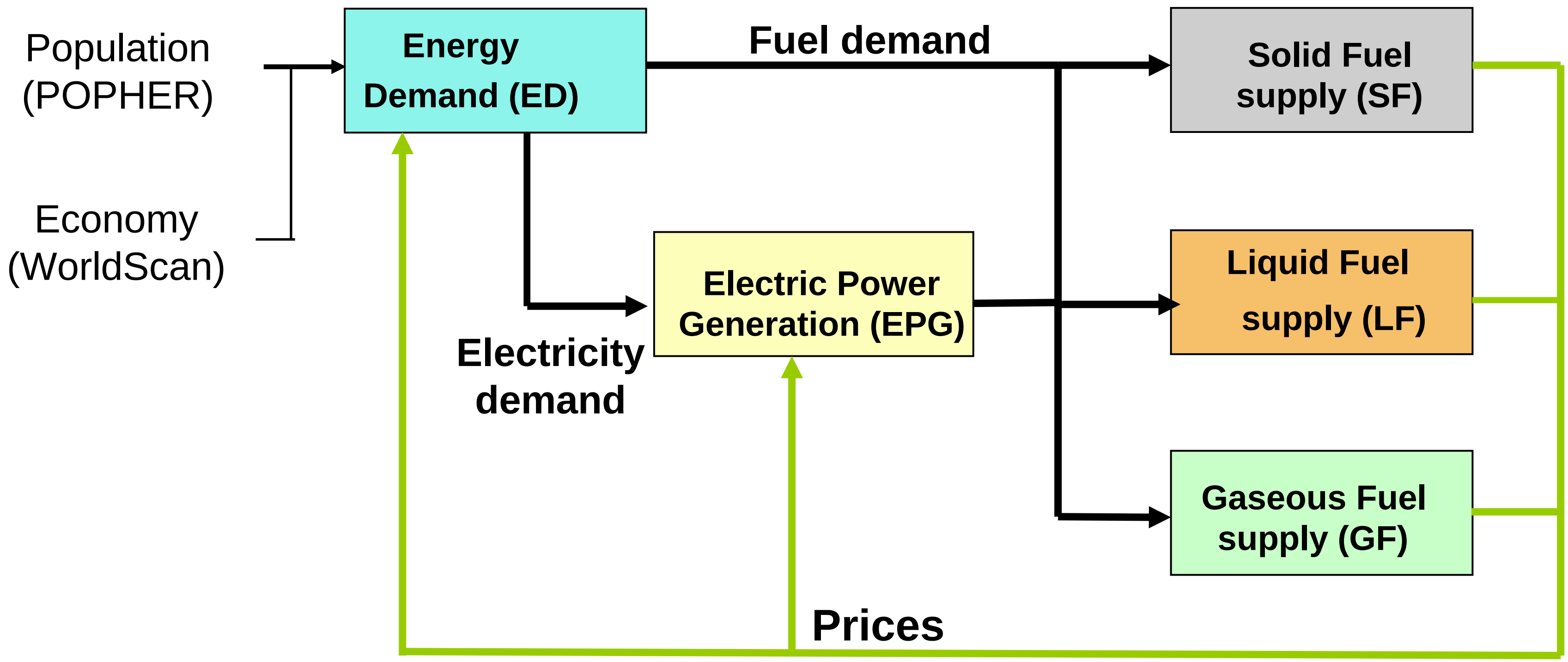


IMAGE 2:

Framework of models and Linkages

TIMER Model : five submodels



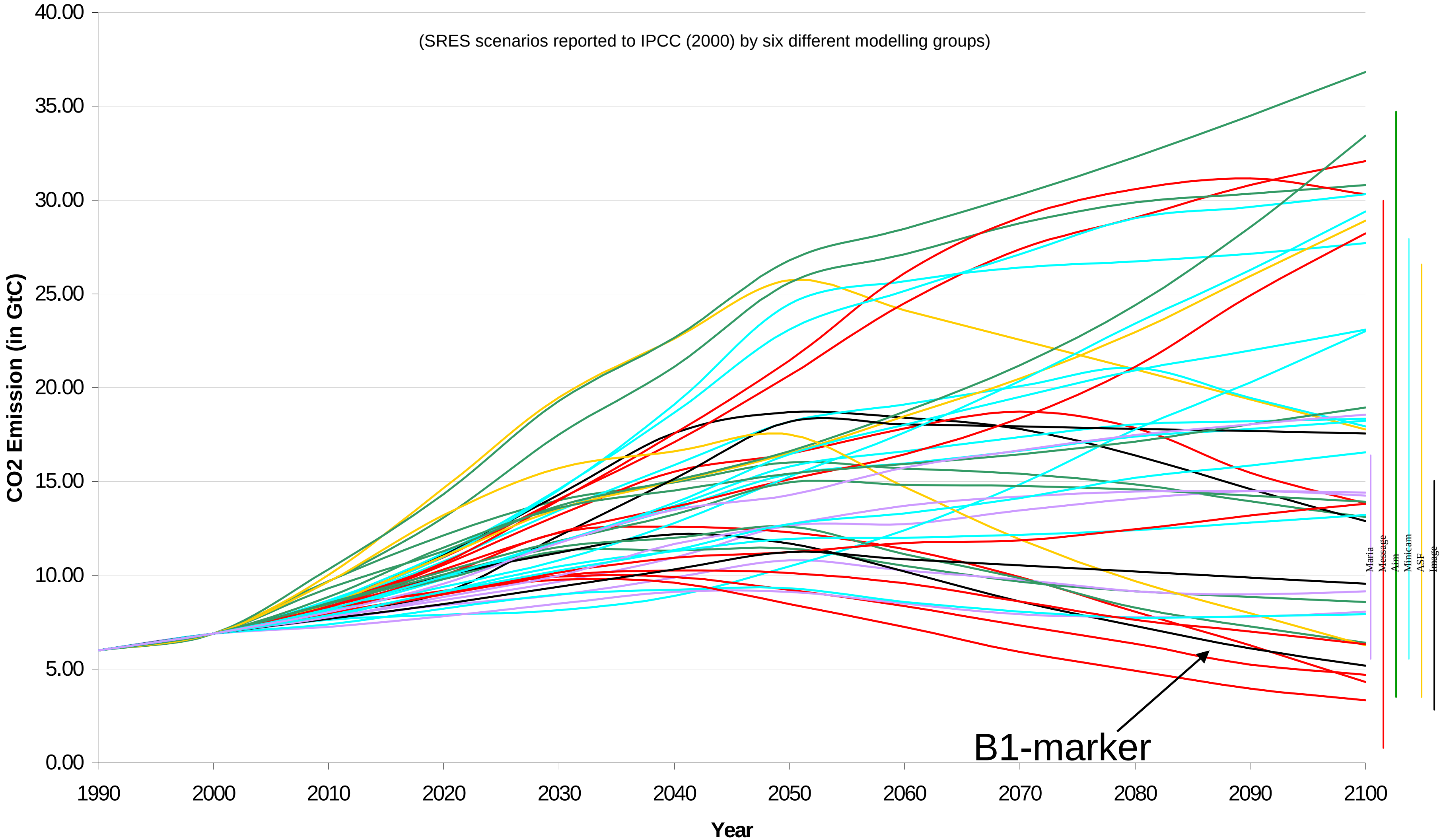
Inputs: Population, GDP capita^{-1}, activity in energy sectors, assumptions regarding technological development, depletion and others.

Outputs: End-use energy consumption, primary energy consumption.

Main objectives TIMER

- To analyse the long-term dynamics of the energy system, and in particular changes in energy demand and the transition to non-fossil fuels within an integrated modeling framework;
- To construct and simulate greenhouse gas emission scenarios that are used in other submodels of IMAGE 2.2 or that are used in meta-models of IMAGE;

Global CO2 emission from fossil fuels



(Van Vuuren et al. 2000)

Key questions

- 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

- Checklist for model quality assistance
- Meta-level analysis SRES scenarios to explore model structure uncertainties
- Global sensitivity analysis (Morris)
- NUSAP expert elicitation workshop to assess pedigree of sensitive model components
- Diagnostic diagram to prioritise uncertainties by combination of criticality (Morris) and strength (pedigree)

Applying NUSAP to a complex model was quite a challenge

TIMER model:

- 300 variables
- 19 world regions
- 5 economic sectors
- 5 types of energy carriers
- 2 forms of energy
- some are time series

$\Rightarrow$ about 160,000 numbers

Checklist

- Assist in quality control in complex models
- Not *models* are good or bad but ‘better’ and ‘worse’ forms of modelling *practice*
- Quality relates to fitness for function
- Help guard against poor practice
- Flag pittfalls

Checklist structure

- Screening questions
- Model & problem domain
- Internal strength
- Interface with users
- Use in policy process
- Overall assessment

Morris (1991)

- facilitates global sensitivity analysis in minimum number of model runs
- covers entire range of possible values for each variable
- parameters varied one step at a time in such a way that if sensitivity of one parameter is contingent on the values that other parameters may take, Morris captures such dependencies

Most sensitive model components:

- Population levels and economic activity
- Intra-sectoral structural change
- Progress ratios for technological improvements
- Size and cost supply curves of fossil fuels resources
- Autonomous and price-induced energy efficiency improvement
- Initial costs and depletion of renewables

Parameter Pedigree

- Proxy
- Empirical basis
- Theoretical understanding
- Methodological rigour
- Validation

Code	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

Elicitation workshop

- Focussed on 40 key uncertain parameters grouped in 18 clusters
- 18 experts (in 3 parallel groups of 6) discussed parameters, one by one, using information & scoring cards
- ***Individual*** expert judgements, informed by group discussion

Structural change / Growth elasticity	Sub module: dem

Definition: These parameters describe the structural change curve. When an economy grows it is assumed to go through successive stages of development. In TIMER, based on historic analysis, that this is also reflected in terms of the demand for energy services in different energy end-use sectors. For instance, in early stages of development the industry sector is dominated by light industry; in a next stage heavy industry dominates and finally industry with high-value added. Consequently the energy intensity of a economy is assumed to go through a maximum with increasing GDP per capita (at PPP). In TIMER, the structural change formulation can be characterised by two important parameters:

Position maximum: Position of the maximum in the GDP per capita (at PPP) vs energy intensity curve

Saturation level: This parameter represents a theoretical minumum in energy intensity, associated with a saturation in energy demand per capita as a function of GDP per capita (at PPP). Note that this saturation point is assumed to be strongly scenario dependent. In a A-storyline the saturation is not met before 2100, in a B storyline it is.

Background Information:

	B1 range:	Range over which sensitivity was tested:
Position maximum:	1189.22, 1.0E+05 1995US$	100.00, 1.0E+05
Saturation level:	0, 3.5E-03 GJ/1995 US$	0, B1 value +50%

Rank in Morris Sensitivity Analysis (maximums are listed from this group of parameters)

Grouped by	Rank	μ(μ)	σ(μ(μ))	μ(σ)
Type:	1	873%	587%	2008%
Module	1	423%	278%	1051%

Dimension	17 Regions	5 Sectors	heat/electricity	5 energy carriers	Other
Variable	x	x	x		

Likely Uncertainty Range: *Maximum*: ± ______ % *Saturation*: ± ______ %

Characterization of variable

		0	1	2	3	4		Elaboration/justification
Value-ladenness	Negligible						High	

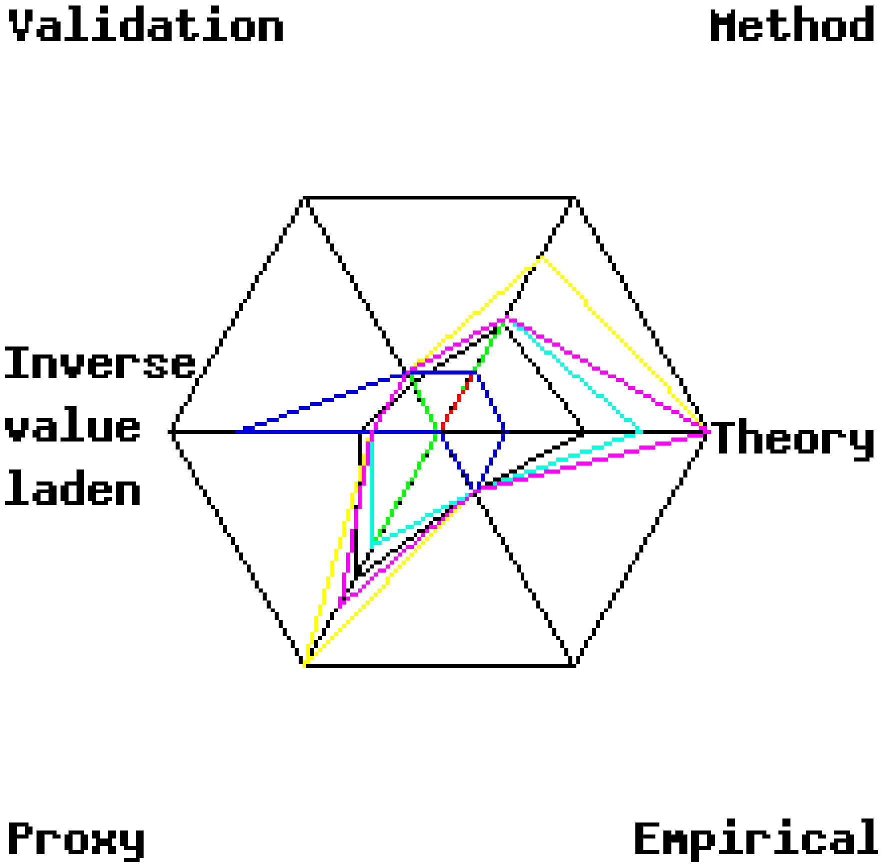
Pedigree

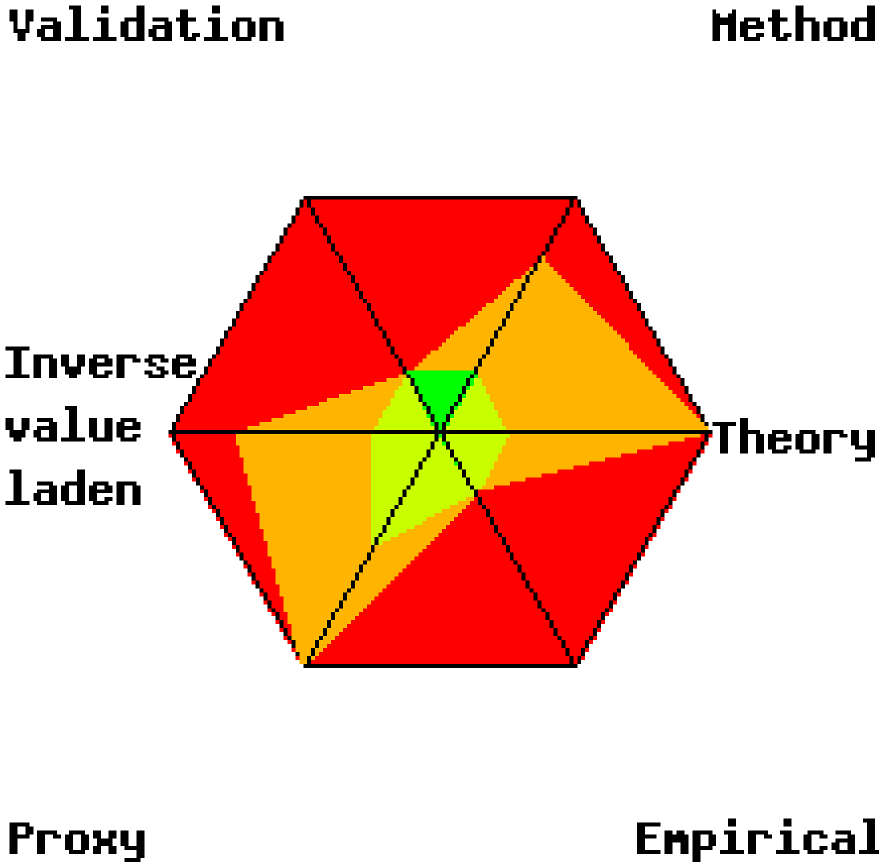
		0	1	2	3	4		Elaboration/justification
Proxy	Not Related						Exact Measure	
Empirical basis	Weak						Strong	
Theoretical understanding	Weak						Strong	
Methodological rigour	Low						High	
Validation	No						Complete	

Instructions

- Do the Pedigree assessment as an **individual** expert judgement, we do not want a group judgement
- Main function of group discussion is clarification of concepts
- Group works on one card at a time
- If you feel you cannot judge the pedigree scores for a given parameter, leave it blank

Example result *gas depletion multiplier*



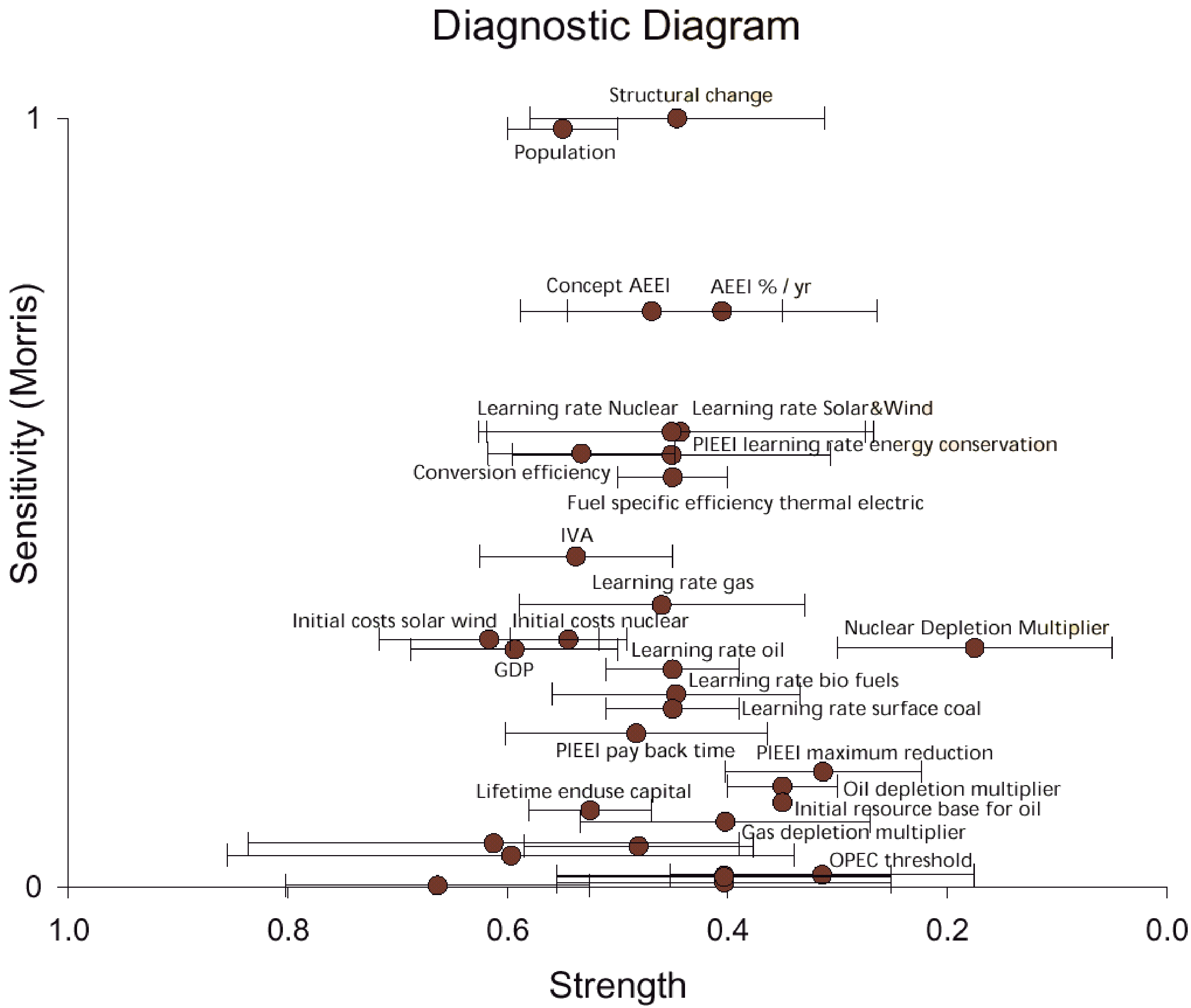


Radar diagram:
Each coloured line represents scores given by one expert

Same data represented as kite diagram:
Green = min. scores, Amber= max scores,
Light green = min. scores if outliers omitted
(Traffic light analogy)

Average scores (0-4)

• proxy	2½	± ½
• empirical	2	± ½
• theory	2	± ½
• method	2	± ½
• validation	1	± ½
• valueladeness	2½	± 1
• competence	2	± ½


Diagnostic Diagram
Structural change
Population
Concept AEEI
AEEI % / yr
Learning rate Nuclear
Learning rate Solar&Wind
PIEEI learning rate energy conservation
Conversion efficiency
Fuel specific efficiency thermal electric
IVA
Learning rate gas
Initial costs solar wind
Initial costs nuclear
Nuclear Depletion Multiplier
Learning rate oil
GDP
Learning rate bio fuels
Learning rate surface coal
PIEEI pay back time
PIEEI maximum reduction
Lifetime enduse capital
Oil depletion multiplier
Initial resource base for oil
Gas depletion multiplier
OPEC threshold
Sensitivity (Morris)
1
0
1.0
0.8
0.6
0.4
0.2
0.0
Strength

Conclusions (1)

- 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.

Conclusions (2)

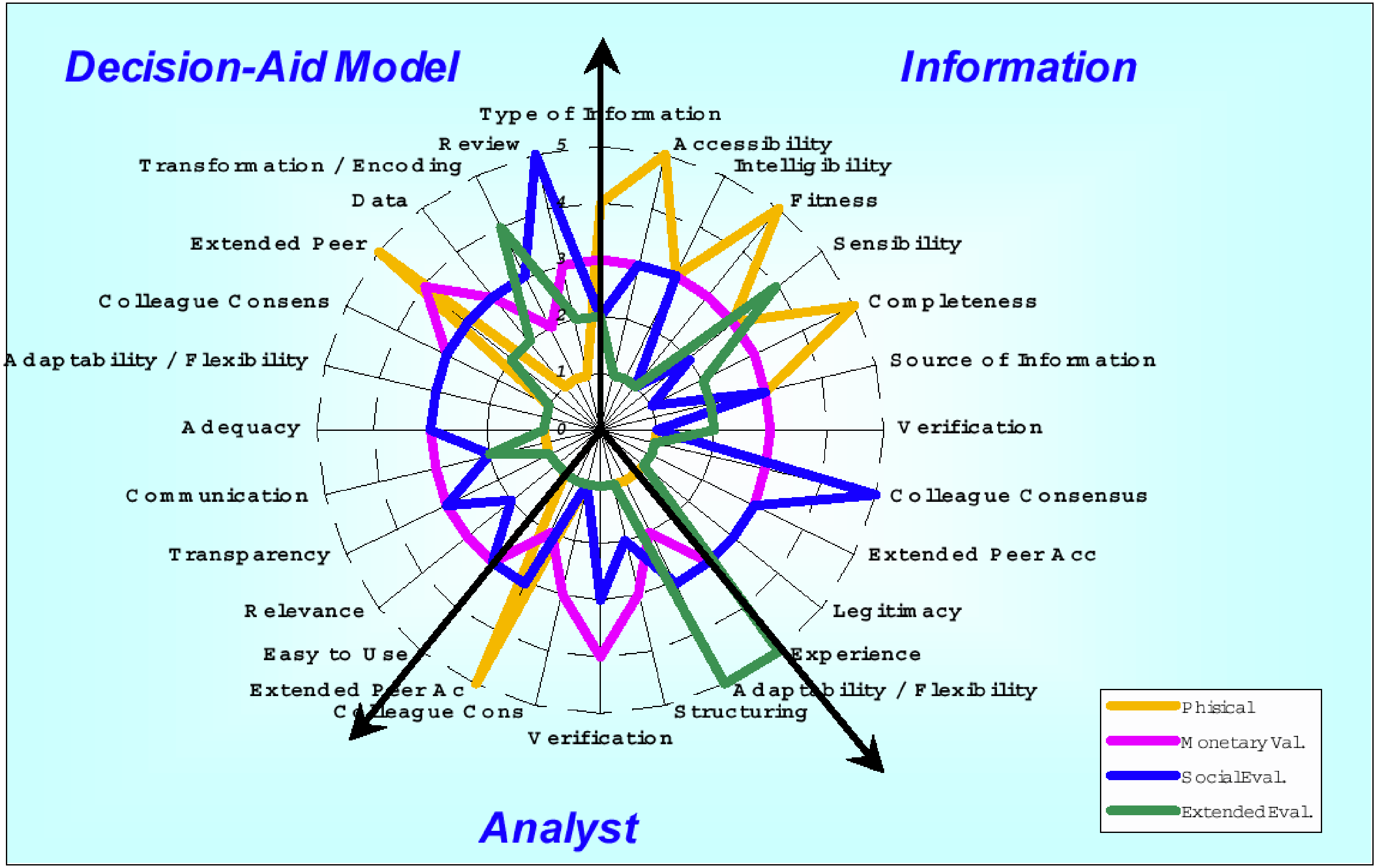
- The diagnostic diagram puts spread and strength together to provide guidance in prioritisation of key uncertainties.

Conclusions (3)

NUSAP method:

- can be applied to complex models in a meaningful way
- helps to focus research efforts on the potentially most problematic model components
- pinpoints specific weaknesses in these components

The Pedigree Exploring Tool (PET)

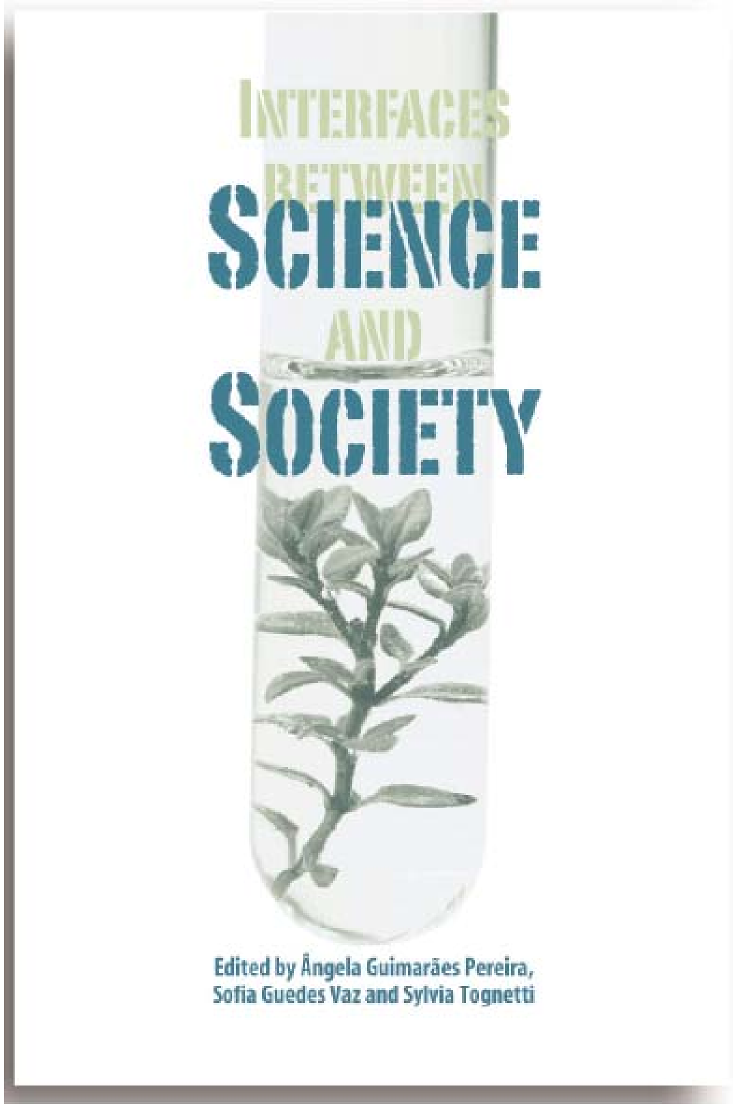


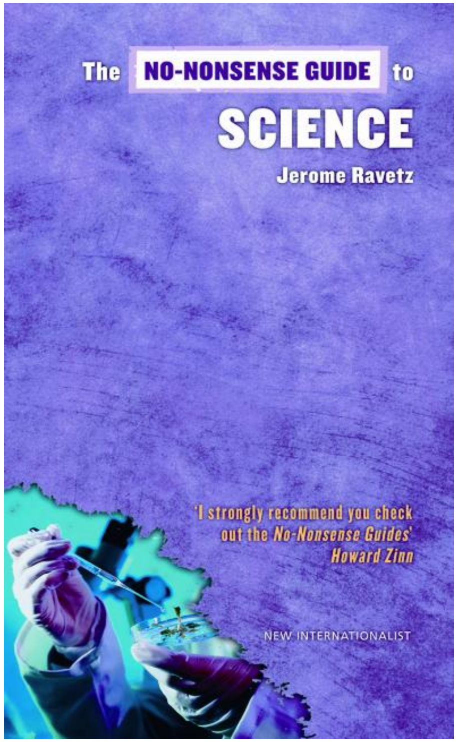
Serafín Corral Quintana, 2000, 2002

In summary, NUSAP

- Has a strong theoretical foundation in the theory of knowledge and the philosophy of science
- Addresses all three dimensions of uncertainty: technical (inexactness), methodological (unreliability) and epistemological (border with ignorance) in an coherent way
- Provides a systematic framework for synthesising qualitative and quantitative assessments of uncertainty
- Can act as a bridge between the quantitative mathematical disciplines and traditions and the qualitative discursive and participatory disciplines and traditions in the field of uncertainty management.
- Helps to focus research efforts on the potentially most problematic model components
- Pinpoints specific weaknesses in these components
- Provides those who produce, use and are affected by policy-relevant knowledge a tool for a critical self-awareness of their engagement with that knowledge. It thereby fosters extended peer review processes.

Books





Websites:

http://www.jvds.nl

http:// www.postnormaltimes.net

http:// www.nusap.net

http://alba.jrc.it/ibss

References

Funtowicz SO, Ravetz JR. Uncertainty and quality in science for policy. Dordrecht: Kluwer; 1990. 229 pp.

Funtowicz SO, Ravetz JR. Science for the post-normal age. Futures 1993;25(7):735–55.

Funtowicz S, Strand R. Models of science and policy. in: Traavik, T. and Lim, L.C. (eds.) Biosafety First: Holistic Approaches to Risk and Uncertainty in Genetic Engineering and Genetically Modified Organisms (forthcoming) Tapir Academic Press, Trondheim

Funtowicz SO. 2006. Why knowledge assessment? In: Interfaces between Science and Society (Guimarães Pereira Â, Guedes Vaz S, Tognetti, S, eds). Sheffield: Greenleaf Publishers, 138–145.

Guimarães Pereira Â, Guedes Vaz S, Tognetti, S, (eds), 2006. Interfaces between Science and Society. Sheffield: Greenleaf Publishers

Janssen PHM, Petersen AC, van der Sluijs JP, Risbey JS, Ravetz JR. 2005. A guidance for assessing and communicating uncertainties. Water Sci Technol 52:125–131.

Ravetz, J. 2006, The no-nonsense guide to science, New Internationalist.

Saltelli A, Chan K, Scott M. 2000, Sensitivity analysis. Probability and statistics series. John Wiley & Sons Publishers; 475 pp.

Saltelli A, Tarantola S, Campolongo F, Ratto M. 2004, Sensitivity analysis in practice: a guide to assessing scientific models. JohnWiley & Sons Publishers; 219 pp.

Smits, M. 2006, Taming monsters: The cultural domestication of new technology, *Technology in Society* 28 (4), pp. 489-504.

van der Sluijs, JP, 2002, A way out of the credibility crisis around model-use in Integrated Environmental Assessment, Futures, 34 133-146.

van der Sluijs JP. 2005. Uncertainty as a monster in the science policy interface: four coping strategies. Water Sci Technol 52:87–92.

van der Sluijs JP, Craye M, Funtowicz SO, Kloprogge P, Ravetz JR, Risbey JS. 2005. Combining quantitative and qualitative measures of uncertainty in model based environmental assessment: the NUSAP system. Risk Anal. 25:481–492.

van der Sluijs, J.P. 2006, Uncertainty, assumptions, and value commitments in the knowledge-base of complex environmental problems, in: Ângela Guimarães Pereira, Sofia Guedes Vaz and Sylvia Tognetti, Interfaces between Science and Society, Green Leaf Publishing, 2006, p.67-84.

van der Sluijs JP, 2007, Uncertainty and Precaution in Environmental Management: Insights from the UPEM conference, Environmental Modelling and Software, 22, (5), 590-598.

Case 3
Chains of models

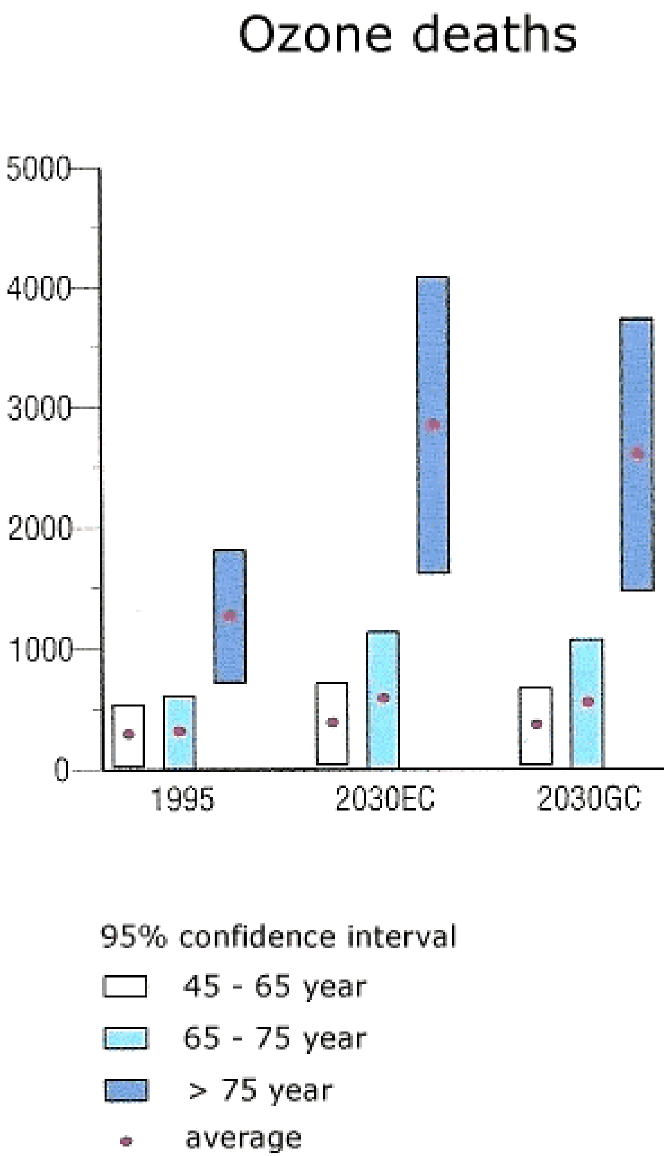
EO5 Environmental Indicators

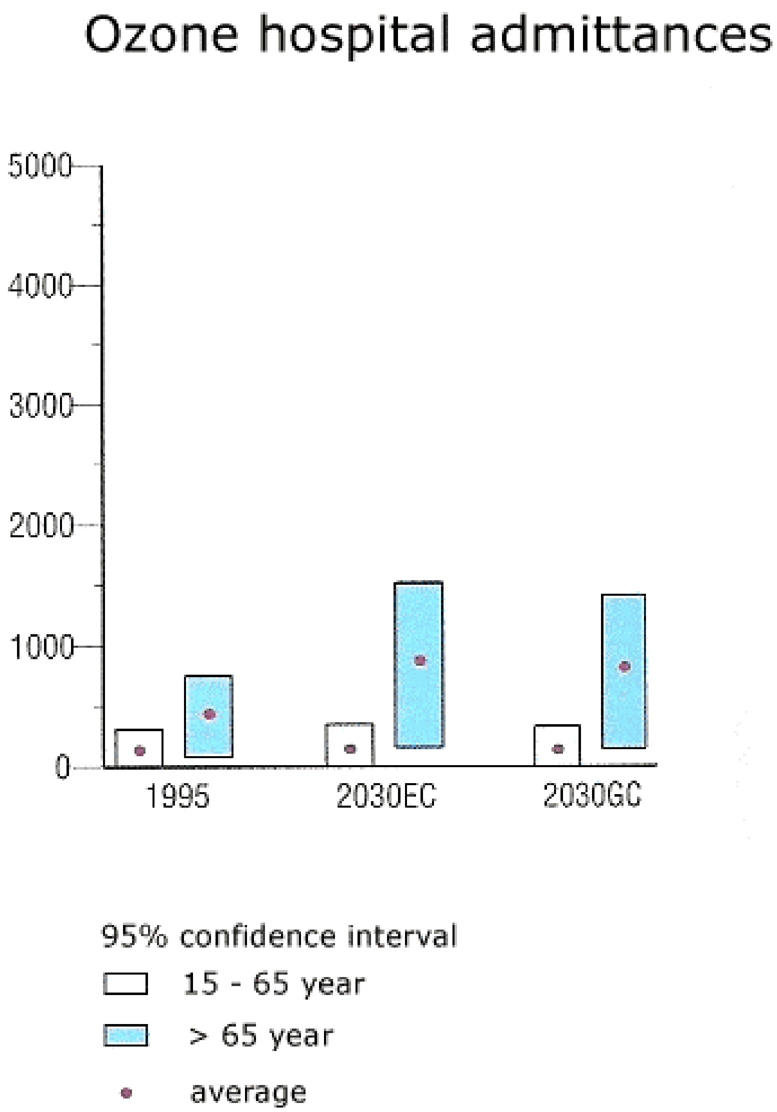
Value-ladenness

- Value orientations and biases of an analyst, an institute, a discipline or a culture can co-shape the way scientific questions are framed, data are selected, interpreted, and rejected, methodologies are devised, explanations are formulated and conclusions are formulated.
- Since theories are always underdetermined by observation, the analysts' biases will fill the epistemic gap which makes any assessment to a certain degree value-laden.
- In a context of (potential) controversy, stakeholder participation and transparency are essential in coping with value ladenness

RIVM Environmental Outlook

- Scenario study issued every 4 years
- hundreds of environmental indicators
- basis for NL Environmental Policy Plan
- Strongly based on chains of model calculations


Ozone deaths
5000
4000
3000
2000
1000
0
1995
2030EC
2030GC
95% confidence interval
45 - 65 year
65 - 75 year
> 75 year
average


Ozone hospital admittances
5000
4000
3000
2000
1000
0
1995
2030EC
2030GC
95% confidence interval
15 - 65 year
> 65 year
average

Calculation chain deaths and hospital admittances due to ozone

1. Societal/demographical developments
2. VOC and NOx emissions in the Netherlands and abroad
3. Ozone concentrations
4. Potential exposure to ozone
5. Number of deaths/hospital admittances due to exposure

Pedigree criteria for reviewing assumptions

- Plausibility
- Inter-subjectivity peers
- Inter-subjectivity stakeholders
- Choice space
- Influence of situational restrictions (time, money, etc.)
- Sensitivity to view and preferences of analyst
- Estimated influence on results

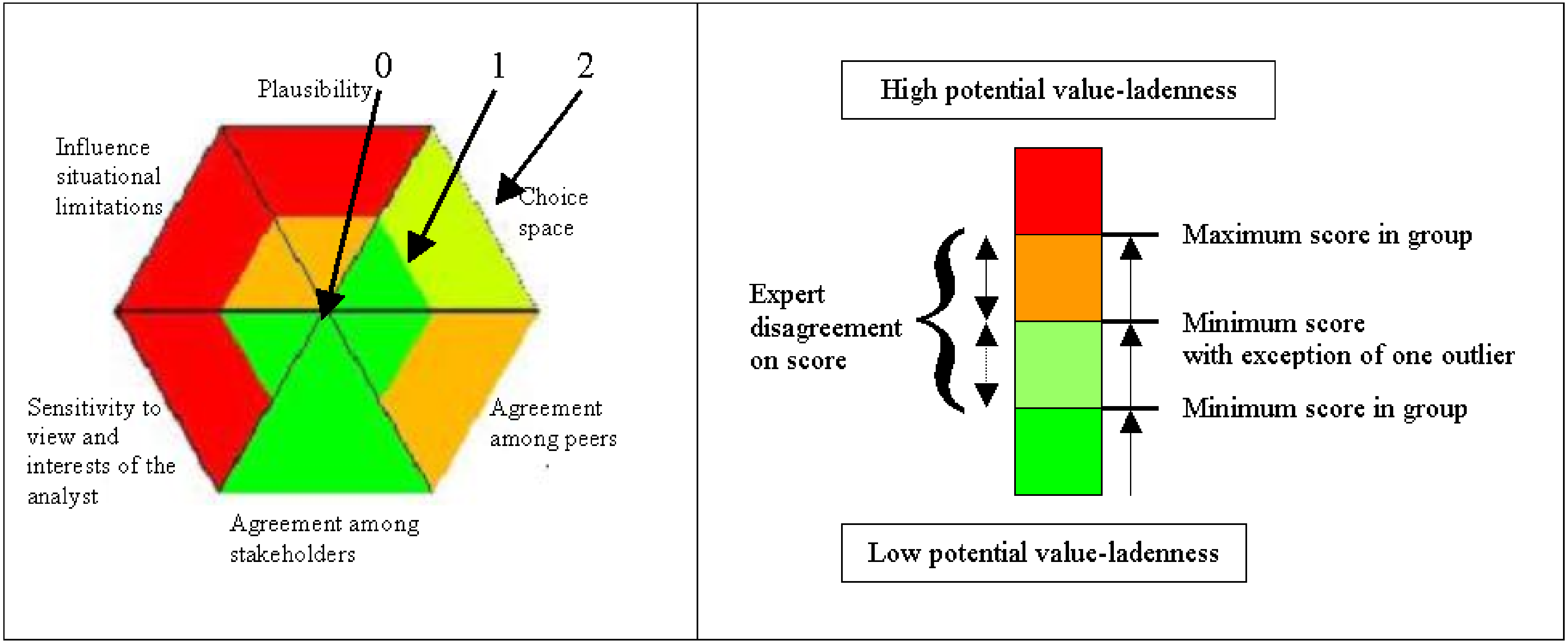
Workshop reviewing assumptions

- Completion of list of key assumptions
- Rank assumptions according to importance
- Elicit pedigree scores
- Evaluate method

Key assumptions deaths and hospital admittances due to ozone

- Uncertainty mainly determined by uncertainty in Relative Risk (RR)
- No differences in emissions abroad between the two scenarios
- Ozone concentration homogeneously distributed in 50 x 50 km grid cells
- Worst case meteo now = worst case future
- RR constant over time (while air pollution mixture may change!)
- Linear dose-effect relationship

Example "traffic light" graph



Assumption that there is a linear dose-effect relationship