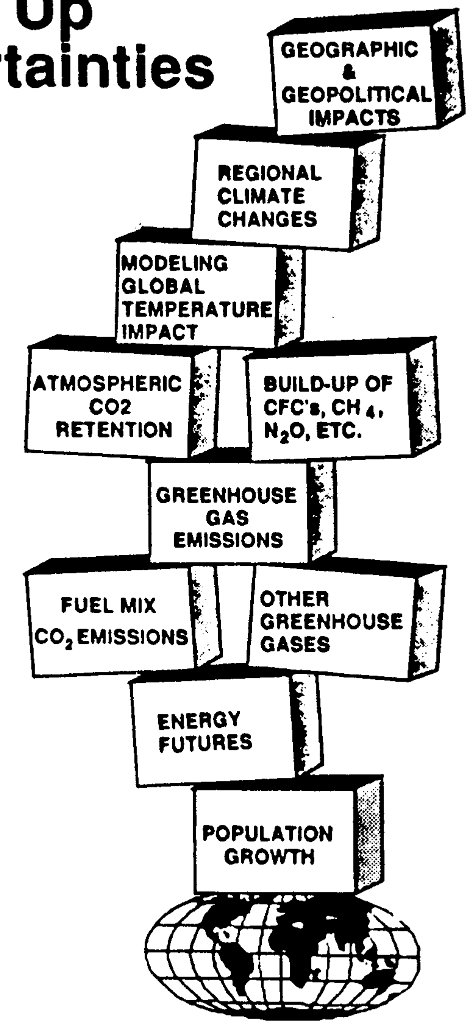
Knowledge Quality Assessment
an intro to The Guidance & NUSAP

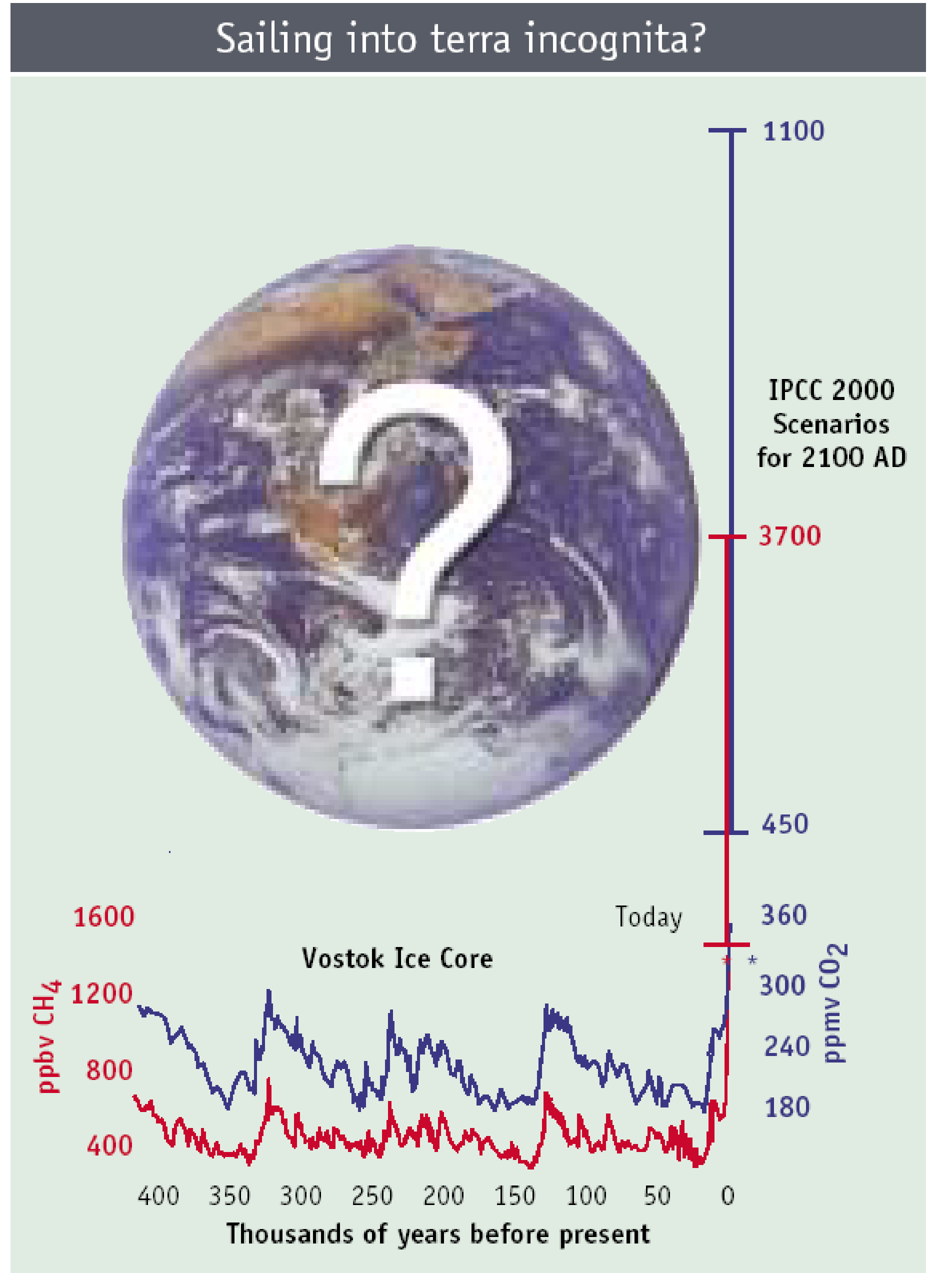
Prof. Dr Jeroen P. van der Sluijs

Senter for vitenskapsteori

GLOBAL CLIMATE CHANGE

Piling Up Uncertainties





How does science-policy interface cope with uncertainties

Two strategies dominate:

- **Overselling certainty**
  - to promote political decisions (enforced consensus), or
- **Overemphasising uncertainty**
  - to prevent political action

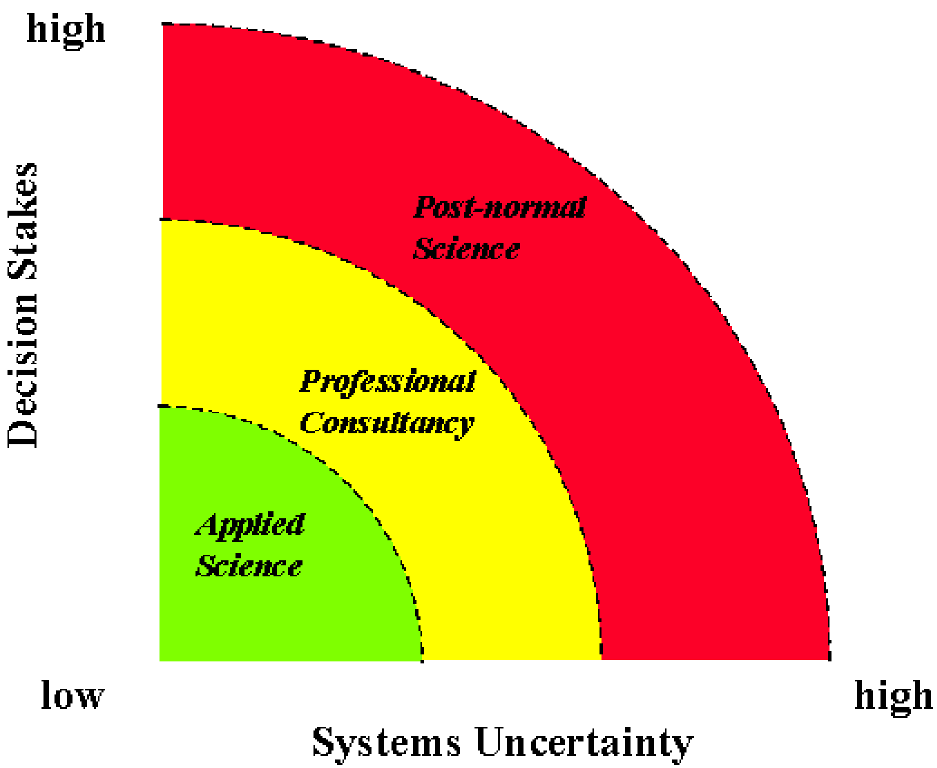
- Both promote decision strategies that are **not fit for meeting the challenges** posed by the uncertainties and complexities faced.

- Need for a third voice next to alarmists and skeptics: coping with uncertainty, scientific dissent & plurality in science for policy.

Complex - *uncertain* - risks

Typical characteristics:

- Decisions urgent
- Stakes high
- Values in dispute
- Irreducible & unquantifiable uncertainty



- Assessment: models, scenarios, assumptions, extrapolations
- (hidden) value loadings in problem frames, indicators chosen, assumptions made

- **Knowledge Quality Assessment!**

(Funtowicz & Ravetz, 1993)
http://www.uu.nl/wetfilos/wetfil10/sprekers/Funtowicz_Ravetz_Futures_1993.pdf

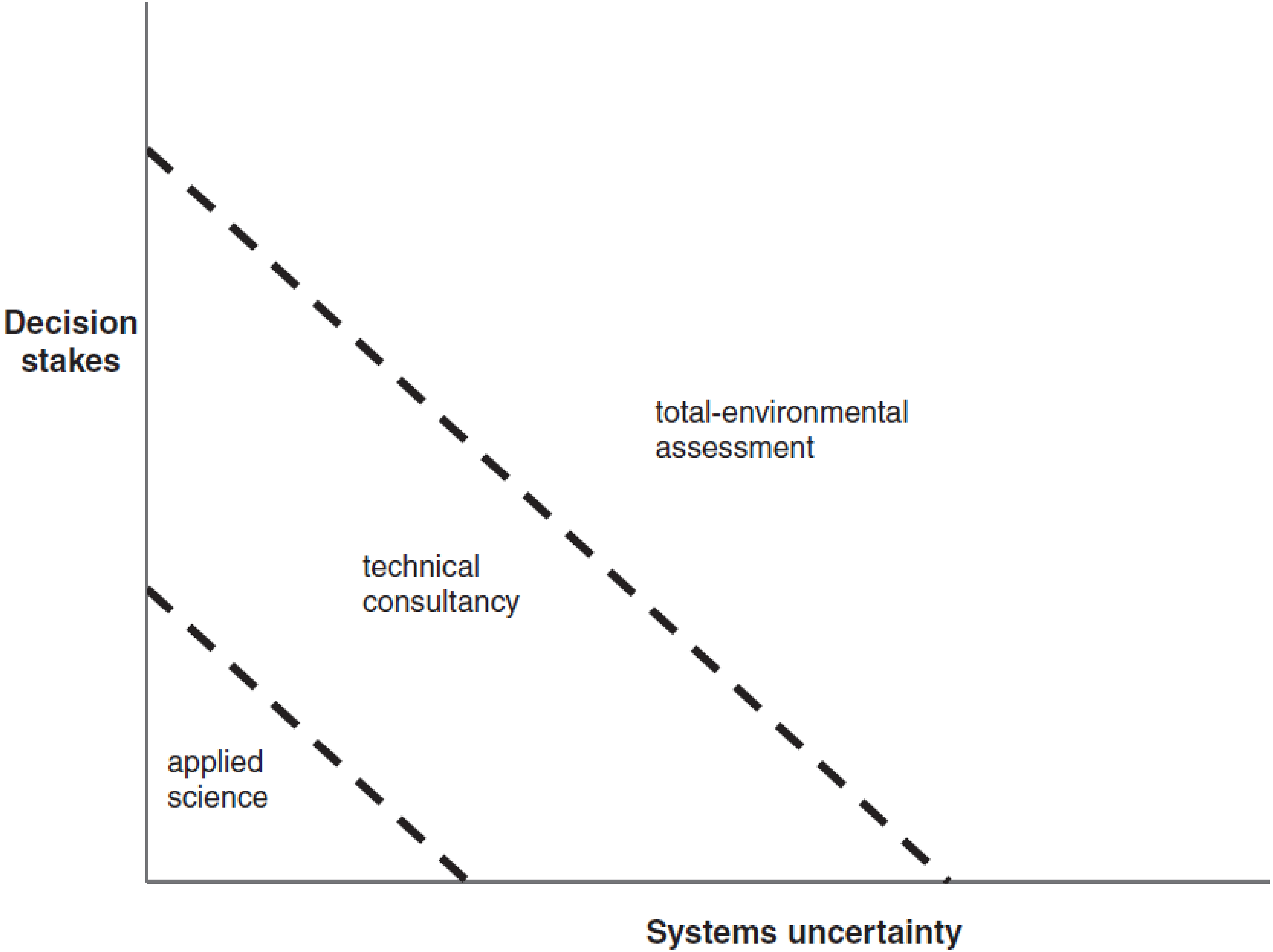


Figure 28.2 The original diagram of three types of risk assessment.
Source: redrawn after Funtowicz & Ravetz (1985).
Note: "Total-environmental assessment" would later be relabelled as "post-normal science".

Elements of Post Normal Science

- Appropriate management of uncertainty quality and value-ladenness
- Plurality of commitments and perspectives
- Internal extension of peer community *(involvement of other disciplines)*
- External extension of peer community *(involvement of stakeholders in environmental assessment & quality control)*

Conventional and new scientific inquiries: comparison of key features and basic beliefs

Conventional science (normal)	Conventional world view	New science (post-normal)	New, emerging world view
• Truth exists, knowledge = Truth	• Truth	• Truth ('t') is plural, context-dependent	• Truths depend on the observer
• Order, structure, certainty	• Simple	• Order, disorder, uncertainty	• Complex, uncertain (chaos?)
• Domination, control over parts	• Hierarchy	• Hierarchy as constraint, scales	• Heterarchy, holarchy
• Clockwork nature, mechanical universe	• Mechanical	• Self-organization, information, lumps	• Holographic, self-organizing
• Deterministic predictable nature	• Determinate	• Unpredictability inherent	• Indeterminate, unpredictable
• Nature evolves to a static climax	• (Most often) Linearly causal	• Non-linear, synergy discontinuity, dynamic	• Mutually causal, synergistic
• Reductionist, collection of parts	• Assembly	• Complex systems, nested holons	• Morphogenic, emergent complexity
• Values hidden (in hypothesis)	• Objective	• Values explicit, essential to inquiry	• Perspectival, contextual
• Expert-driven, exclusive to peers	• Expert	• Extended Peers, inclusive of others	• Participatory, interdisciplinary

Sources: Funtowicz and Ravetz, 1991, 1993b; Kuhn, 1962; Schwartz and Ogilvy, 1979. **Lister 1998**

Illustrative example

Protecting a strategic fresh-water resource under the Water Supply Act Denmark

Case:

- Important aquifer west of Copenhagen
- groundwater abstraction 12 million m^3/year
- Copenhagen County had to prepare an action plan for protection of groundwater against pollution
- Scientist were asked to assess aquifer's vulnerability to pollution in a 175 km^2 area

A practical problem:

Protecting a strategic fresh-water resource

5 scientists addressed same question:

“which parts of this area are most vulnerable to nitrate pollution and need to be protected?”

(Refsgaard, Van der Sluijs et al, 2006)

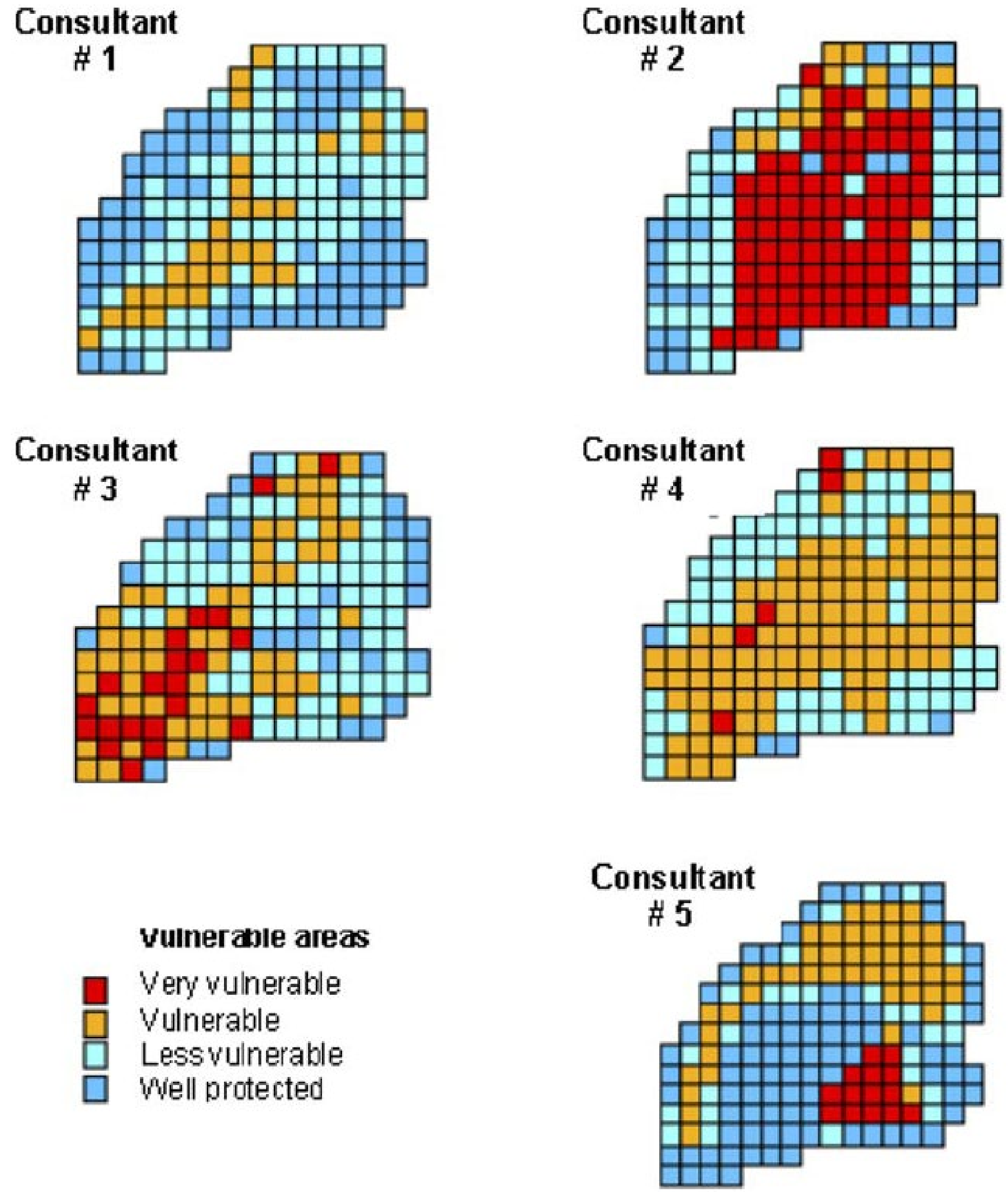


Fig. 1. Model predictions on aquifer vulnerability towards nitrate pollution for a 175 km^2 area west of Copenhagen [11].

3 framings of uncertainty

'deficit view'

- Uncertainty is provisional
- Reduce uncertainty, make ever more complex models
- *Tools:* quantification, Monte Carlo, Bayesian belief networks
  - ***Speaking truth to power***

'evidence evaluation view'

- Comparative evaluations of research results
- *Tools:* Scientific consensus building; multi disciplinary expert panels
- focus on robust findings
  - ***Speaking [consensus] to power***

'complex systems view / *post-normal* view'

- Uncertainty is intrinsic to complex systems
- Openly deal with deeper dimensions of uncertainty
- *Tools:* Knowledge Quality Assessment
  - ***Working deliberatively within imperfections***

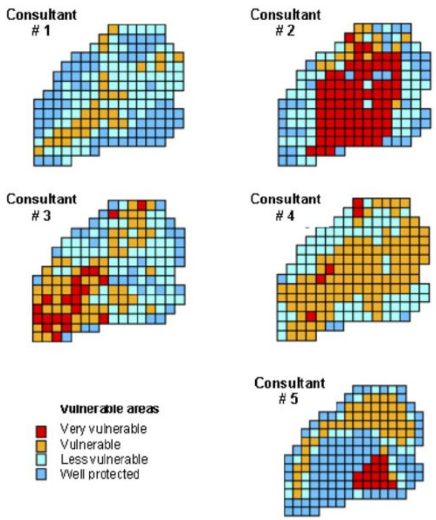


Fig. 1. Model predictions on aquifer vulnerability towards nitrate pollution for a 175 km^2 area west of Copenhagen [11].

How to act upon such uncertainty?

- **Bayesian** approach: 5 priors. Average and update likelihood of each grid-cell being red with data (but oooops, there is no data and we need decisions now)
- IPCC approach: Lock the 5 consultants up in a room and don't release them before they have **consensus**
- **Nihilist** approach: Dump the science and decide on an other basis
- **Precautionary** robustness approach: protect all grid-cells
- **Academic bureaucrat** approach: Weigh by citation index (or H-index) of consultant.
- Select the consultant that you **trust** most
- Real life approach: Select the consultant that best fits your **policy agenda**
- Post normal: explore the relevance of our ignorance: **working deliberatively within imperfections**

Clark & Majone 1985

Critical Appraisal of Scientific Inquiries with Policy Implications

1. Criticism by whom?

Critical roles

- Scientist
- Peer group
- Program Manager or Sponsor
- Policy maker
- Public interests groups

Clark & Majone 1985

Criticism of what?

Critical modes:

- Input
  - data; methods, people, competence, (im)matureness of field
- Output
  - problem solved? hypothesis tested?
- Process
  - good scientific practice, procedures for review, documenting etc.

Table 1. Critical criteria. **(Clark & Majone, 1985)**

Critical Role	Critical Mode: Input	Output	Process
Scientist	Resource and time constraints; available theory; institutional support; assumptions; quality of available data; state of the art.	Validation; sensitivity analyses; technical sophistication; degree of acceptance of conclusions; impact on policy debate; imitation; professional recognition.	Choice of methodology (e.g., estimation procedures); communication; implementation; promotion; degree of formalization of analytic activities within the organization.
Peer Group	Quality of data; model and/or theory used; adequacy of tools; problem formulation. Input variables well chosen? Measure of success specified in advance?	Purpose of the study. Are conclusions supported by evidence? Does model offend common sense? Robustness of conclusions; adequate coverage of issues.	Standards of scientific and professional practice; documentation; review of validation techniques; style; interdisciplinarity.
Program Manager or Sponsor	Cost; institutional support within user organization; quality of analytic team; type of financing (e.g., grant vs. contract).	Rate of use; type of use (general education, program evaluation, decisionmaking, etc.); contribution to methodology and state of the art; prestige. Can results be generalized, applied elsewhere?	Dissemination; collaboration with users. Has study been reviewed?
Policymaker	Quality of analysts; cost of study; technical tools used (hardware and software). Does problem formulation make sense?	Is output familiar and intelligible? Did study generate new ideas? Are policy indications conclusive? Are they consonant with accepted ethical standards?	Ease of use; documentation. Are analysts helping with implementation? Did they interact with agency personnel? With interest groups?
Public Interest Groups	Competence and intellectual integrity of analysts. Are value systems compatible? Problem formulation acceptable? Normative implications of technical choices (e.g., choices of data).	Nature of conclusions; equity. Is analysis used as rationalization or to postpone decision? All viewpoints taken into consideration? Value issues.	Participation; communication of data and other information; adherence to strict rules of procedure.

Clark & Majone 1985

Meta quality criteria:

- Adequacy
  - reliability, reproducibility, uncertainty analysis etc.
- Value
  - Internal: how well is the study carried out?
  - External: fitness for purpose, fitness for function
  - Personal: subjectivity, preferences, choices, assumptions, bias
- Effectiveness
  - Does it help to solve practical problems
- Legitimacy
  - numinous: natural authority, independence, credibility, competence
  - civil: agreed procedures

CoCliServ's deliberative knowledge quality assessment tool

Critical mode Critical role	Input (and context)	Process	Output	Use
(Climate) scientists				
Local government				
Water boards				
Delta program				
Bird's nest				
Provinces				
Residents				
…				
…				

https://cocliserv.cearc.fr/sites/cocliserv.cearc.fr/files/resultats/CoCliServ_D5.1.pdf

Short list of quality criteria

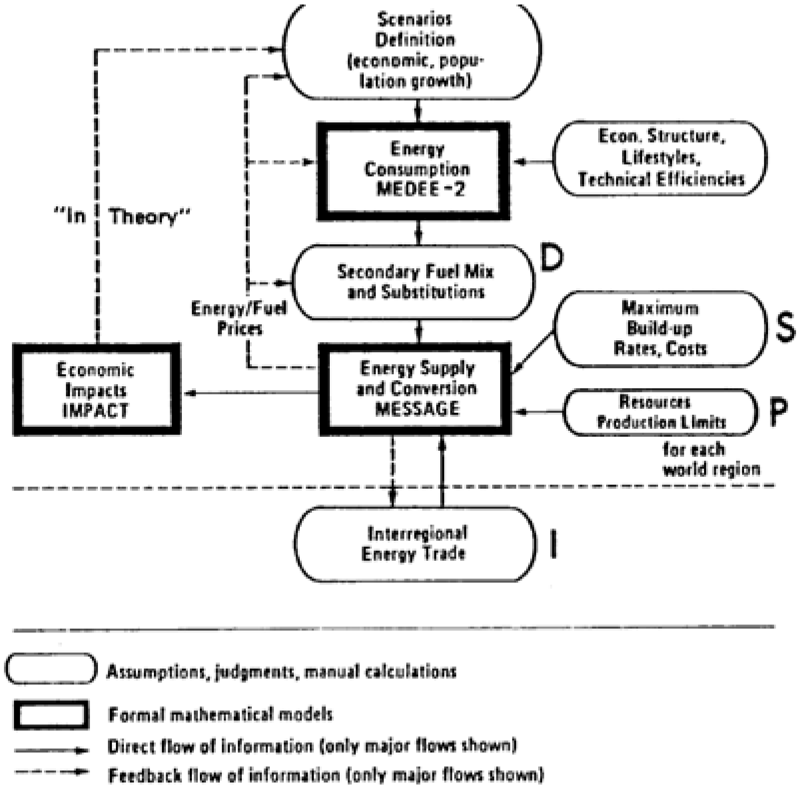
|Input |Process |Output |Use|

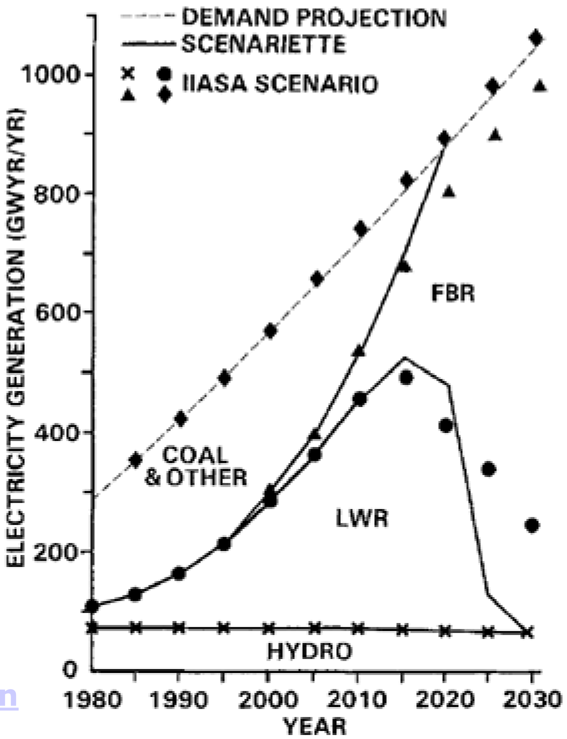
To what degree are the climate services:

- Salient and fit for function
- Uncertainty aware
- Based on credible data
- Inclusive and interactive
- Legitimate and deliberative
- Transparent and responsible
- Intelligible and usable
- Flexible and adaptable
- Iterative and accounting for progressing insights
- Encouraging of learning

https://cocliserv.cearc.fr/sites/cocliserv.cearc.fr/files/resultats/CoCliServ_D5.1.pdf

Energy modelling 1980s





http://www.nature.com/nature/journal/v312/n5996/pdf/312691a0.pdf
https://proxy.reeds.uvsq.fr/galleries/broceliande7/node/986/1584/grain

Uncertainty in knowledge based society: the problems

1984 Keepin & Wynne:

"Despite the appearance of analytical rigour, IIASA's widely acclaimed global energy projections are highly unstable and based on informal guesswork. This results from inadequate peer review and quality control, raising questions about political bias in scientific analysis."

Keepin and Wynne, 1984, *Nature*, **312**, p. 691-695.
http://www.nature.com/nature/journal/v312/n5996/pdf/312691a0.pdf

Scandal at the Netherlands Environmental Assessment Agency

RIVM / De Kwaadsteniet (1999)

"RIVM over-exact prognoses based on virtual reality of computer models"

Newspaper headlines:

- Environmental institute lies and deceits
- Fuss in parliament after criticism on environmental numbers
- The bankruptcy of the environmental numbers
- Society has a right on fair information, RIVM does not provide it

Science, Technology, & Human Values
36(3) 362-388

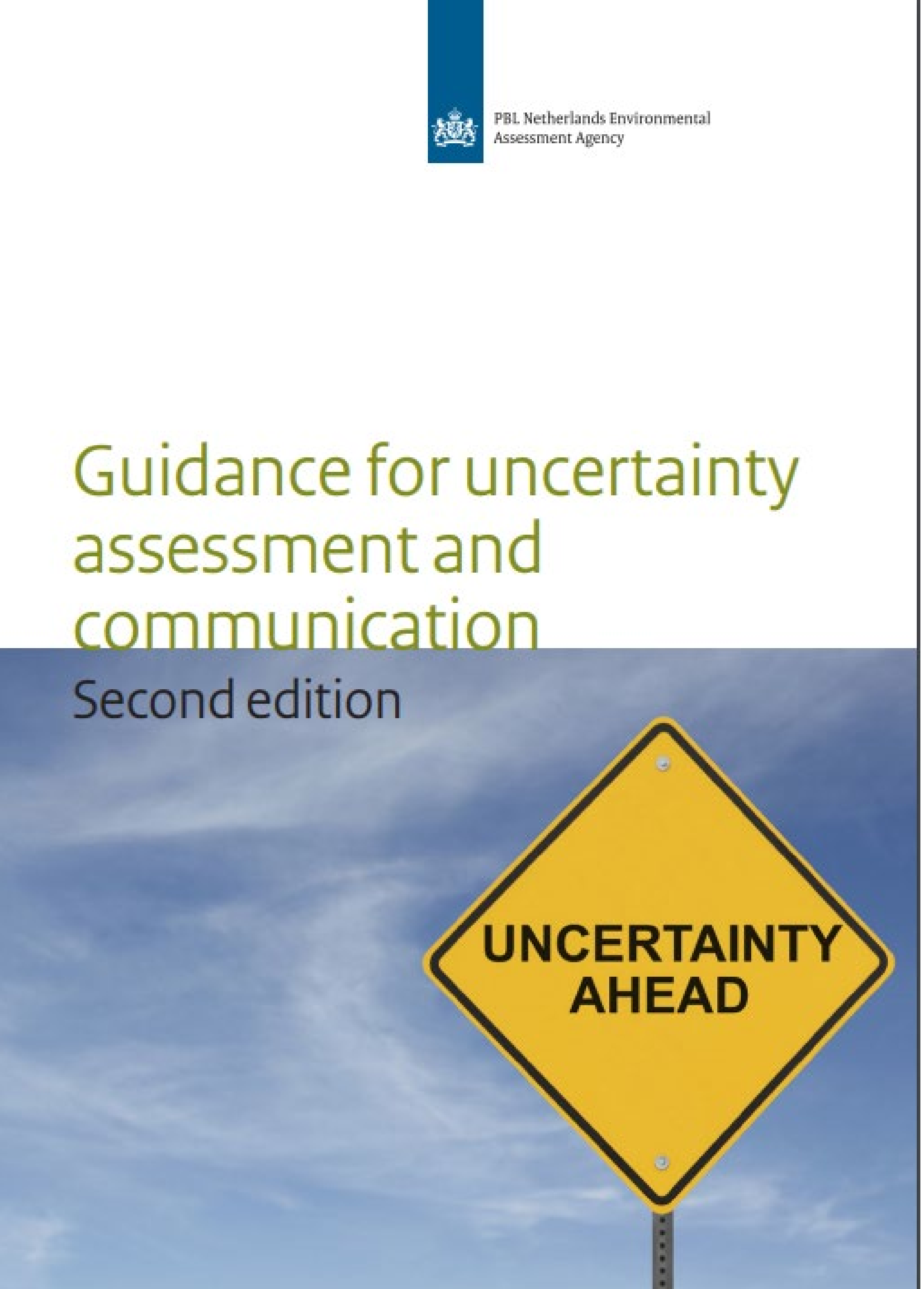
DOI: 10.1177/0162243910385797
http://sthv.sagepub.com

Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency

Arthur C. Petersen,[1] Albert Cath,[2] Maria Hage,[1] Eva Kunseler,[1] and Jeroen P. van der Sluijs[3,4]

Abstract

About a decade ago, the Netherlands Environmental Assessment Agency (PBL) unwittingly embarked on a transition from a technocratic model of science advising to the paradigm of "post-normal science" (PNS). In response to a scandal around uncertainty management in 1999, a Guidance for "Uncertainty Assessment and Communication" was developed with advice from the initiators of the PNS concept and was introduced in 2003. This was followed in 2007 by a "Stakeholder Participation" Guidance. In this article, the authors provide a combined insider/outsider perspective on the transition process. The authors assess the extent to which the PNS paradigm has delivered new approaches in the agency's practice and analyze two projects—on long-term options for Dutch sustainable development policy and for urban development policy—the latter in somewhat more detail. The authors identify several paradoxes PBL encounters when putting the PNS concept into practice. It is concluded that an openness to other


PBL Netherlands Environmental Assessment Agency

Guidance for uncertainty assessment and communication

Second edition

http://www.pbl.nl/en/publications/guidance-for-uncertainty-assessment-and-communication

NL Environmental Assessment Agency (RIVM/MNP) Guidance: Systematic reflection on uncertainty & quality in:

Foci	Key issues
Problem framing	Other problem views; interwovenness with other problems; system boundaries; role of results in policy process; relation to previous assessments
Involvement of stakeholders	Identifying stakeholders; their views and roles; controversies; mode of involvement
Selection of indicators	Adequate backing for selection; alternative indicators; support for selection in science, society, and politics
Appraisal of knowledge base	Quality required; bottlenecks in available knowledge and methods; impact of bottlenecks on quality of results
Mapping and assessing relevant uncertainties	Identification and prioritisation of key uncertainties; choice of methods to assess these; assessing robustness of conclusions
Reporting uncertainty information	Context of reporting; robustness and clarity of main messages; policy implications of uncertainty; balanced and consistent representation in progressive disclosure of uncertainty information; traceability and adequate backing

Problem framing and context

- Explore rival problem frames
- Relevant aspects / system boundary
- Typify problem structure
- Problem lifecycle / maturity
- Role of study in policy process
- Uncertainty in socio-political context

Type-III error:
Assessing the wrong problem by incorrectly accepting the false meta-hypothesis that there is no difference between the boundaries of a problem, as defined by the analyst, and the actual boundaries of the problem (Dunn, 1997).

Context validation (Dunn, 1999).
The validity of inferences that we have estimated the proximal range of rival hypotheses.

Context validation can be performed by a participatory bottom-up process to elicit from scientists and stakeholders rival hypotheses on causal relations underlying a problem and rival problem definitions.

Involvement of stakeholders

- Identify relevant stakeholders.
- Identification of areas of agreement and disagreement among stakeholders on value dimensions of the problem.
- Recommendations on when to involve different stakeholders in the assessment process.

Roles of stakeholders

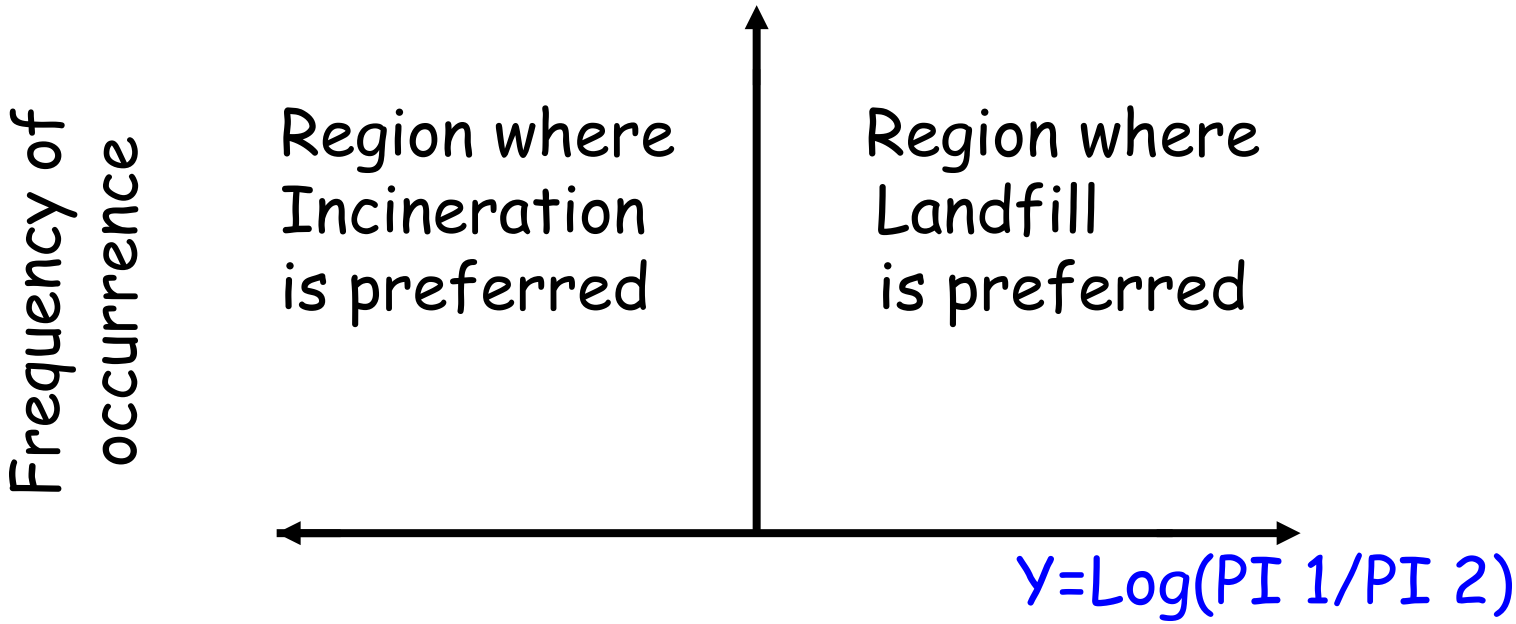
- (Co-) definer of the problems to be addressed
  - What knowledge is relevant?
- Source of knowledge
- Quality control of the science (for instance: review of assumptions)

Indicators

- How well do indicators used address key aspects of the problem?
- Use of proxies
- Alternative indicators?
- Limitations of indicators used?
- Scale and aggregation issues
- Controversies in science and society about these indicators?

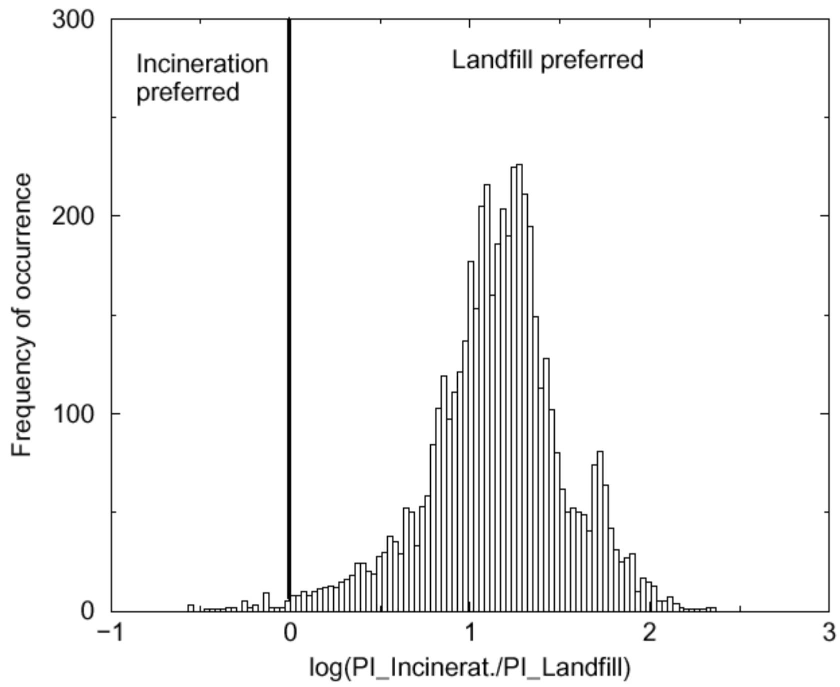
High uncertainty is not the same as low quality

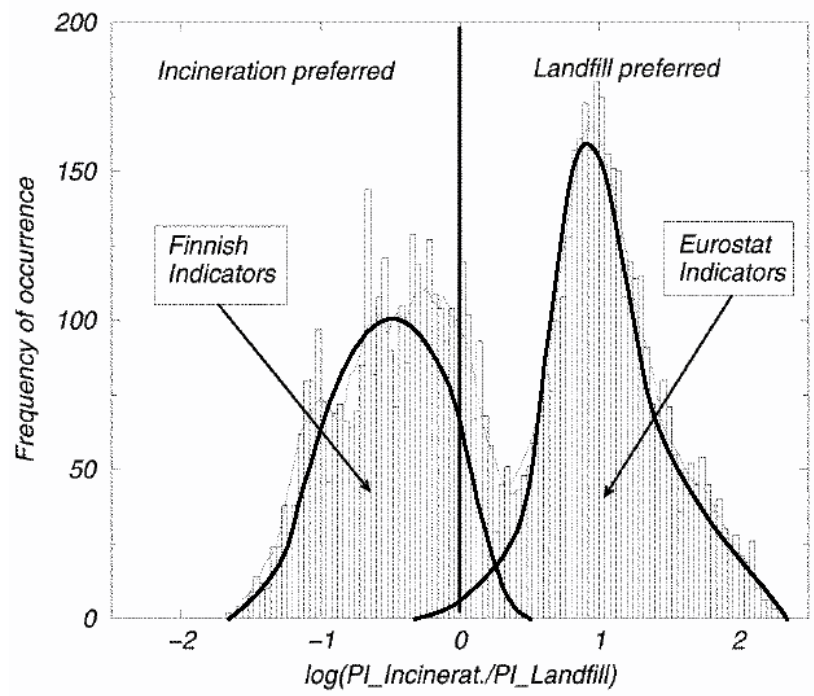
Example: imagine the inference is Y = the logarithm of the ratio between the two pressure-on-decision indices PI1 and PI2



High uncertainty is not the same as low quality,

but..... methodological uncertainty can de dominant





(slide borrowed from Andrea Saltelli)

Do we know enough to quantify?

Risbey & Kandlikar (2007): What format is in accordance with the level of knowledge on the quantity?

- Full probability density function
  - Robust, well defended distribution
- Bounds
  - Well defended percentile bounds
- First order estimates
  - Order of magnitude assessment
- Expected sign or trend
  - Well defended trend expectation
- Ambiguous sign or trend
  - Equally plausible contrary trend expectations
- Effective ignorance
  - Lacking or weakly plausible expectations

Uncertainty is more than a number

Dimensions of uncertainty:

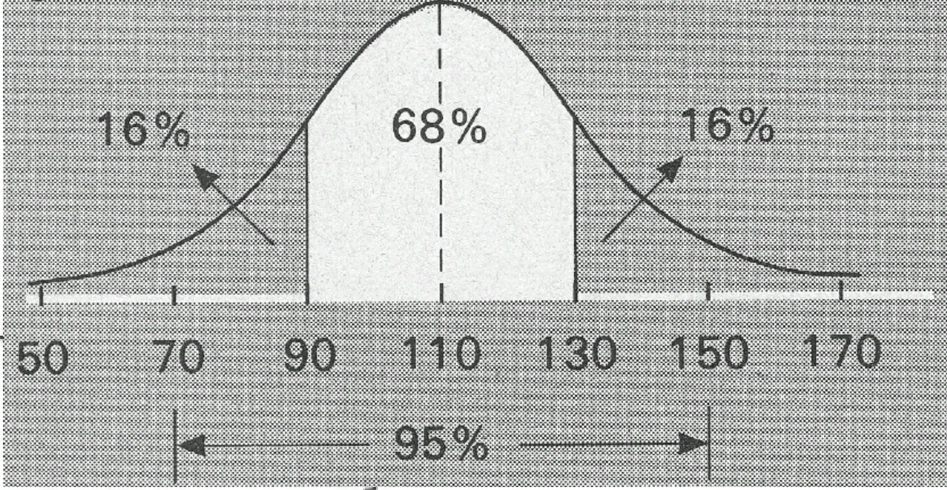
- Technical (inexactness)
- Methodological (unreliability)
- Epistemological (ignorance)
- Societal (limited social robustness)

Reliability intervals normal distributions

± σ = 68 %

± 2σ = 95 %

± 3σ = 99.7 %



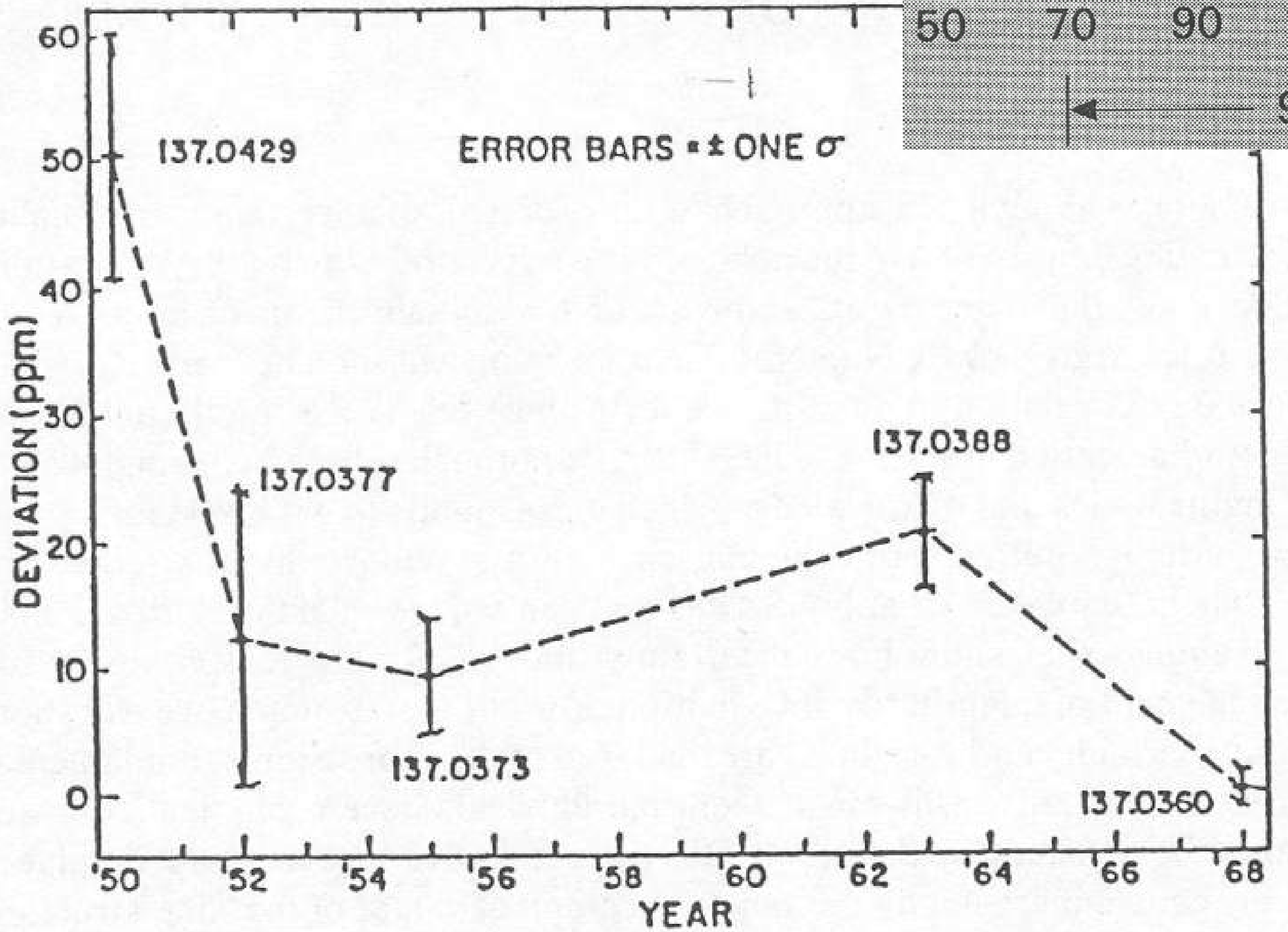
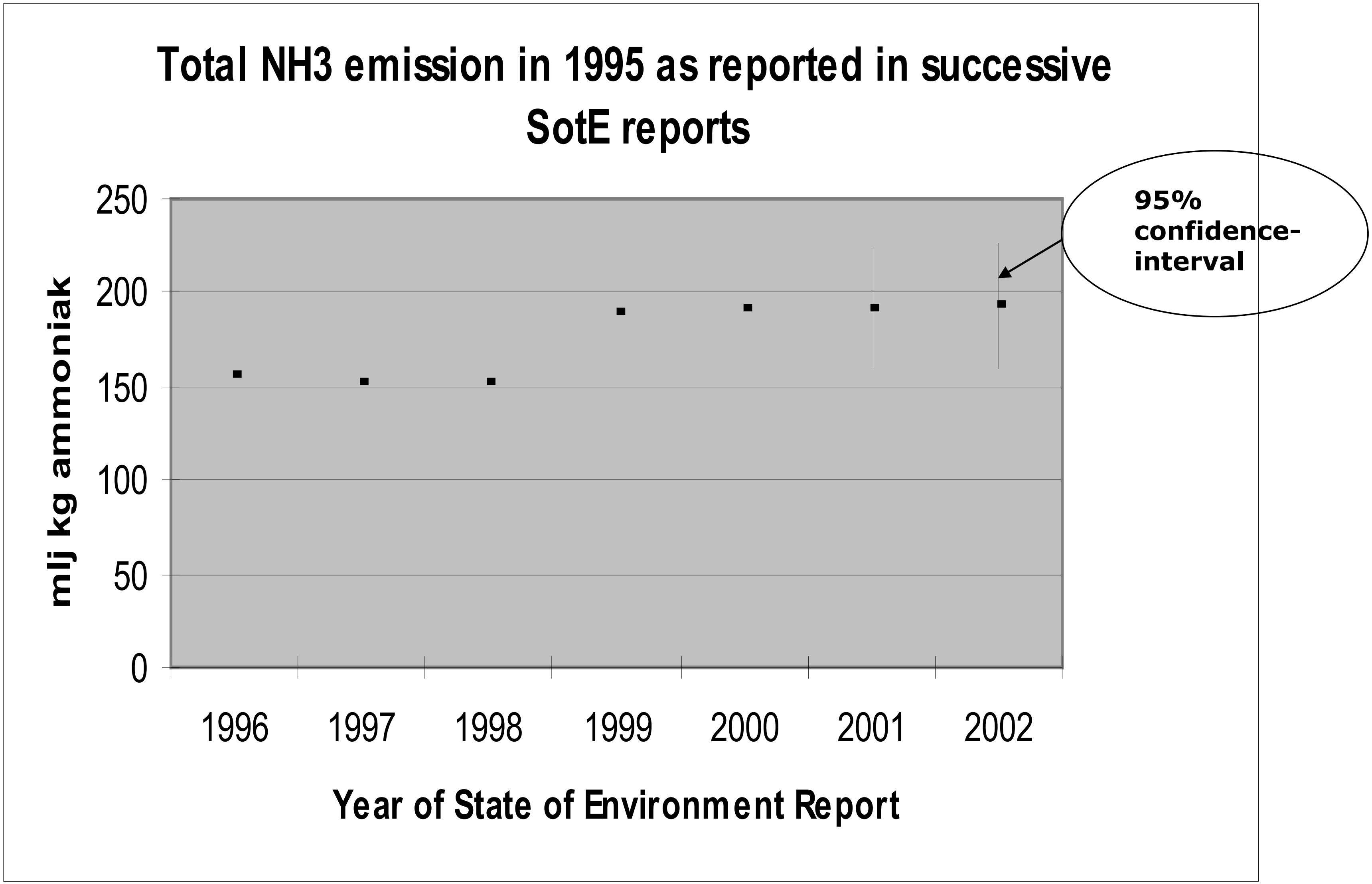


Fig. 1. Successive recommended values of the fine-structure constand α^{-1} (B. N. Taylor *et al.*, 1969, 7)



Total NH3 emission in 1995 as reported in successive SotE reports
95% confidence-interval
mlj kg ammoniak
250
200
150
100
50
0
1996
1997
1998
1999
2000
2001
2002
Year of State of Environment Report

NUSAP: Qualified Quantities

Classic scientific notational system:

- **N**umeral **U**nit **S**pread

For problems in the post-normal domain, add two qualifiers:

- **A**ssessment & **P**edigree

  "Assessment" expresses expert judgement on reliability of numeral + spread

  "Pedigree" expresses multi-criteria evaluation of the strength of a number by looking at:

  - Background history by which the number was produced
  - Underpinning and scientific status of the number

Example Pedigree matrix parameter strength

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

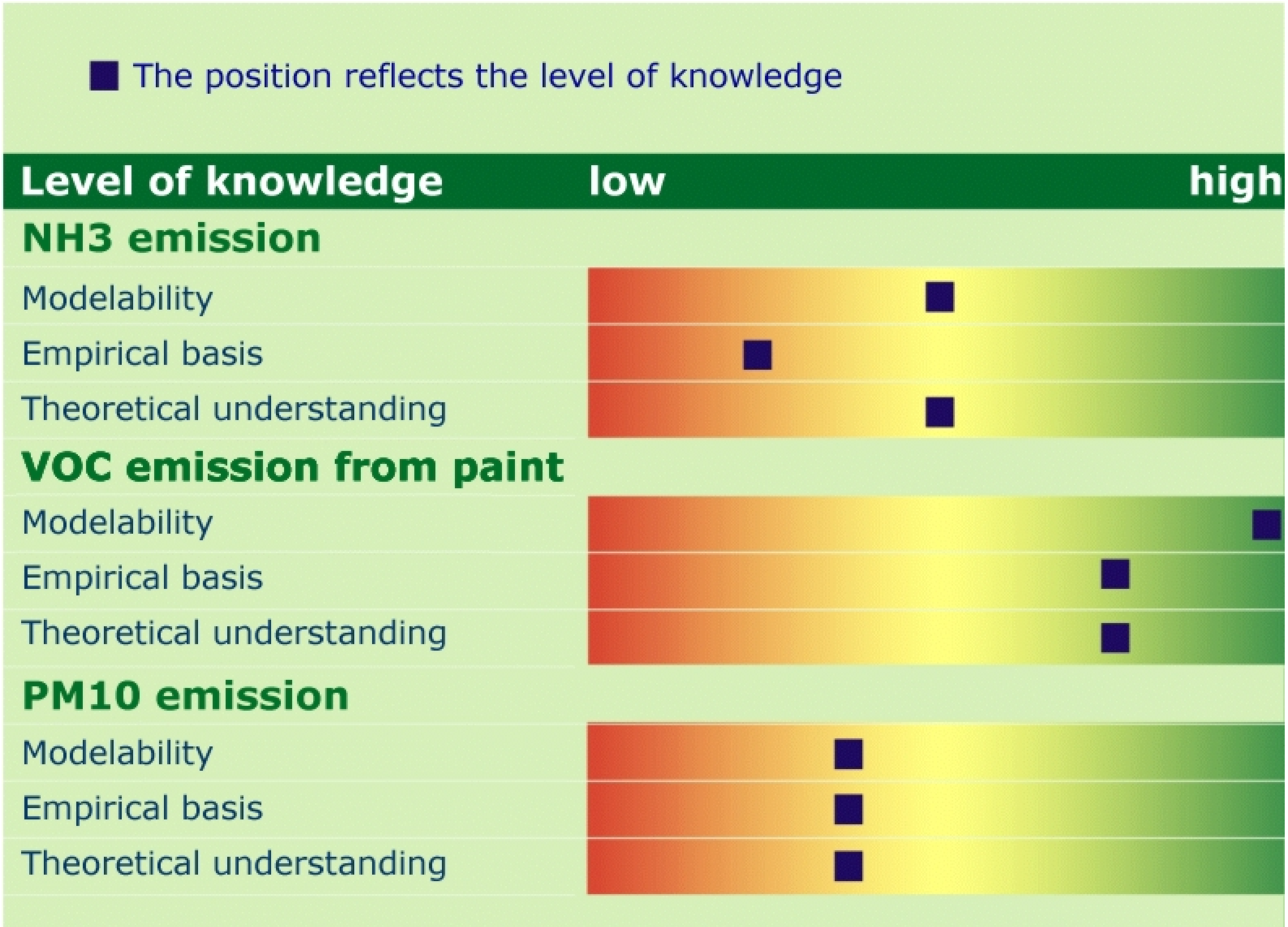
Example Pedigree results

	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

This example is the case of VOC emissions from paint in the Netherlands, calculated from national sales statistics (NS) in 5 sectors (Ship, Building & Steel, Do It Yourself, Car refinishing and Industry) and assumptions on additional thinner use (Th%) and a lump sum for imported paint and an assumption for its VOC percentage. See full research report on www.nusap.net for details.

Example: Air Quality



Energy Research & Social Science 46 (2018) 332–344

Assessing qualitative and quantitative dimensions of uncertainty in energy modelling for policy support in the United Kingdom

Steve Pye[a,e,*], Francis G.N. Li[a], Arthur Petersen[b], Oliver Broad[c], Will McDowall[c], James Price[a], Will Usher[d]

A B S T R A C T

Strategic planning for the low carbon energy transition is characterised by a high degree of uncertainty across many knowledge domains and by the high stakes involved in making decisions. Energy models can be used to assist decision makers in making robust choices that reflect the concerns of many interested stakeholders. Quantitative model insights alone, however, are insufficient as some dimensions of uncertainty can only be assessed via qualitative approaches. This includes the strength of the knowledge base underlying the models, and the biases and value-ladenness brought into the process based on the modelling choices made by users. To address this deficit in current modelling approaches in the UK context, we use the NUSAP (Numeral Unit Spread Assessment Pedigree) approach to qualify uncertainty in the energy system model, ESME. We find that a range of critical model assumptions that are highly influential on quantitative model results have weaknesses, or low pedigree scores, in aspects of the knowledge base that underpins them, and are subject to potential value-ladenness. In the case of the UK, this includes assumptions around CCS deployment and bioenergy resources, both of which are highly influential in driving model outcomes. These insights are not only crucial for improving the use of models in policy-making and providing a more comprehensive understanding of uncertainty in models, but also help to contextualise quantitative results, and identify priority future research areas for improving the knowledge base used in modelling. The NUSAP approach also promotes engagement across a broader set of stakeholders in the analytical process, and opens model assumptions up to closer scrutiny, thereby contributing to transparency.

https://doi.org/10.1016/j.erss.2018.07.028

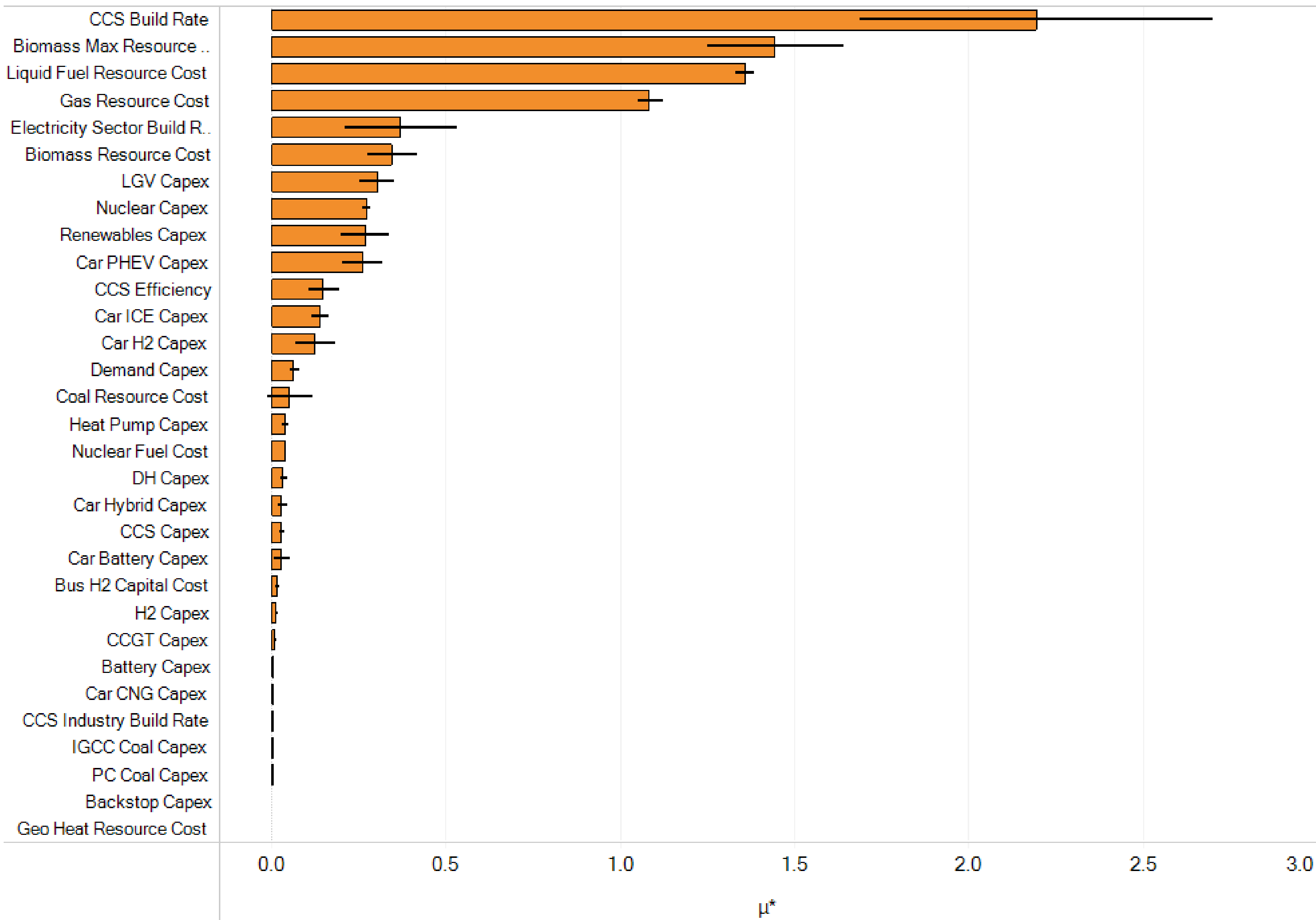


Figure E1. Results of Morris Method analysis, ranking the most important input parameters in relation to the variance in the output metric discounted system costs

Criteria	Criteria scores				
	4	3	2	1	0
Proxy	Exact representation	Good representation	Moderate or acceptable representation	Weak representation	Poor representation
Empirical basis	Observation	Mix of observations and model-based estimates	Model estimates only	Educated guess	Crude speculation
Rigour	Best available practice	Reliable method; very few concerns	Acceptable method but questions on reliability	Preliminary or experimental methods with no clear view of reliability	No discernible rigour in the method, grave concerns
Validation	Huge database of reliable sources	Compared with numerous reliable sources	Limited validation with only a few reliable sources	Weak validation, questions on reliability of sources	No validation
Theoretical understanding	Universally agreed theory	Accepted theory	Accepted theory but lack of consensus	Preliminary theory	Crude speculation
Choice space	No alternatives	Only a few acceptable/plausible alternatives	Small or limited range of alternatives	Moderate range of alternatives	Extremely wide range of alternatives
Justification	Fully justified	Strong justification	Acceptable justification	Weak justification	Completely speculative
Agreement amongst peers	Complete or near-complete agreement	High degree of agreement, with some variation	Some disagreement possible, there are a few competing schools of thought	Low degree of agreement, contentious subject	No agreement or almost no agreement, extremely controversial

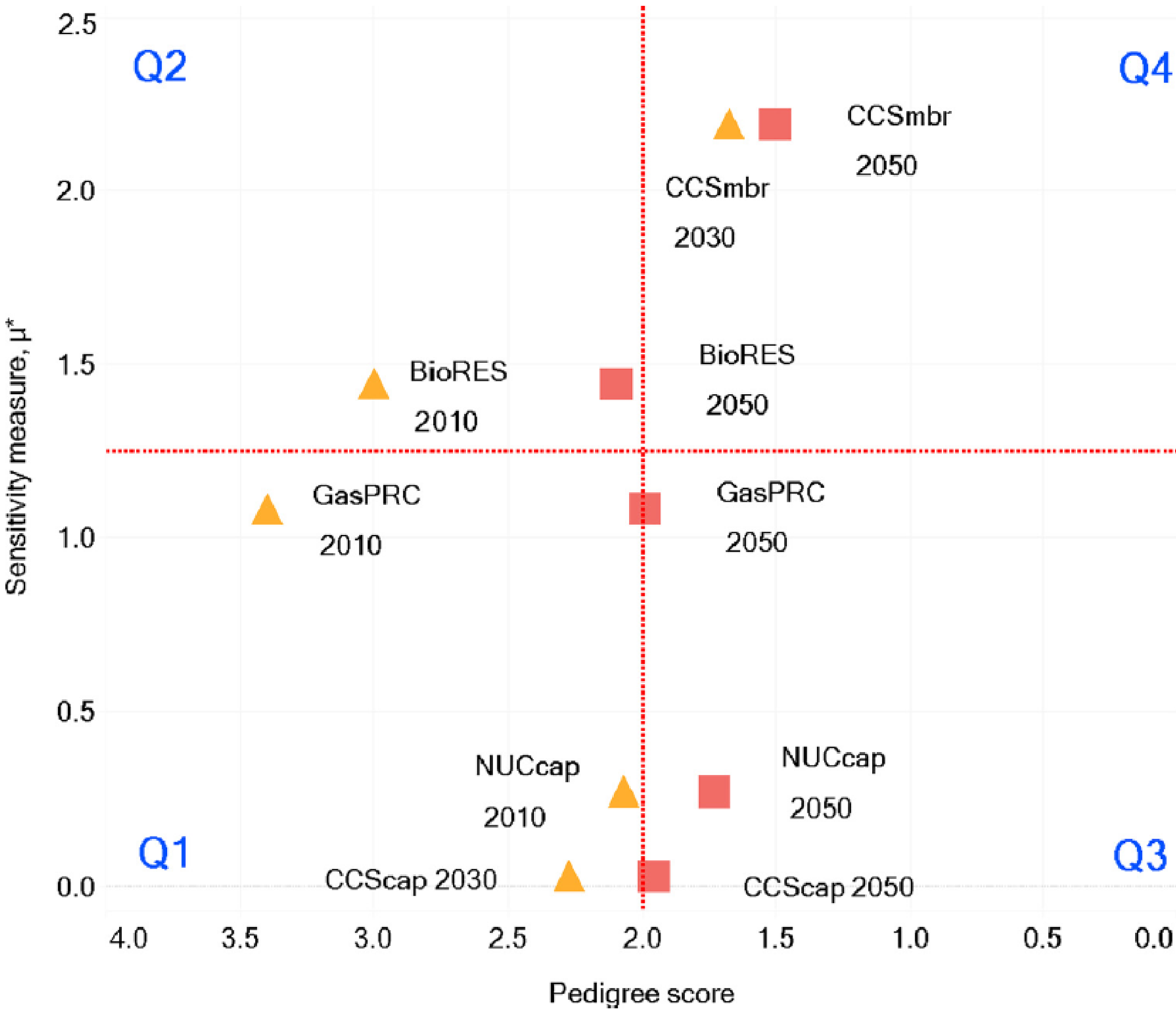


Fig. 6. Diagnostic diagram to compare qualitative (pedigree scores) against quantitative uncertainties (sensitivity measure). The sensitivity measure (based on the Morris Method approach) highlights the influence of the modelled uncertainty on the variance across the model objective function, the total discounted system costs.

Environ. Res. Lett. **10** (2015) 045005

Critical appraisal of assumptions in chains of model calculations used to project local climate impacts for adaptation decision support—the case of Baakse Beek

Jeroen P van der Sluijs[1] **and J Arjan Wardekker**[2]

Abstract

In order to enable anticipation and proactive adaptation, local decision makers increasingly seek detailed foresight about regional and local impacts of climate change. To this end, the Netherlands Models and Data-Centre implemented a pilot chain of sequentially linked models to project local climate impacts on hydrology, agriculture and nature under different national climate scenarios for a small region in the east of the Netherlands named Baakse Beek. The chain of models sequentially linked in that pilot includes a (future) weather generator and models of respectively subsurface hydrogeology, ground water stocks and flows, soil chemistry, vegetation development, crop yield and nature quality. These models typically have mismatching time step sizes and grid cell sizes. The linking of these models unavoidably involves the making of model assumptions that can hardly be validated, such as those needed to bridge the mismatches in spatial and temporal scales. Here we present and apply a method for the systematic critical appraisal of model assumptions that seeks to identify and characterize the weakest assumptions in a model chain. The critical appraisal of assumptions presented in this paper has been carried out ex-post. For the case of the climate impact model chain for Baakse Beek, the three most problematic assumptions were found to be: land use and land management kept constant over time; model linking of (daily) ground water model output to the (yearly) vegetation model around the root zone; and aggregation of daily output of the soil hydrology model into yearly input of a so called 'mineralization reduction factor' (calculated from annual average soil pH and daily soil hydrology) in the soil chemistry model. Overall, the method for critical appraisal of model assumptions presented and tested in this paper yields a rich qualitative insight in model uncertainty and model quality. It promotes reflectivity and learning in the modelling community, and leads to well informed recommendations for model improvement.

http://dx.doi.org/10.1088/1748-9326/10/4/045005

Assumptions in model chains

Analysis

1. Identify explicit and implicit assumptions in the calculation chain.

2. Identify and prioritize key-assumptions in the chain.

3. Assess the pedigree of key-assumptions.

4. Identify 'weak' links in the calculation chain.

5. Further analyze the potential value-ladenness of key assumptions.

Revision

6. Revise/extend assessment:

– sensitivity analysis of key-assumptions;

– diversification of assumptions;

– different choices in chain.

Communication

7. Communication:

– key-assumptions;

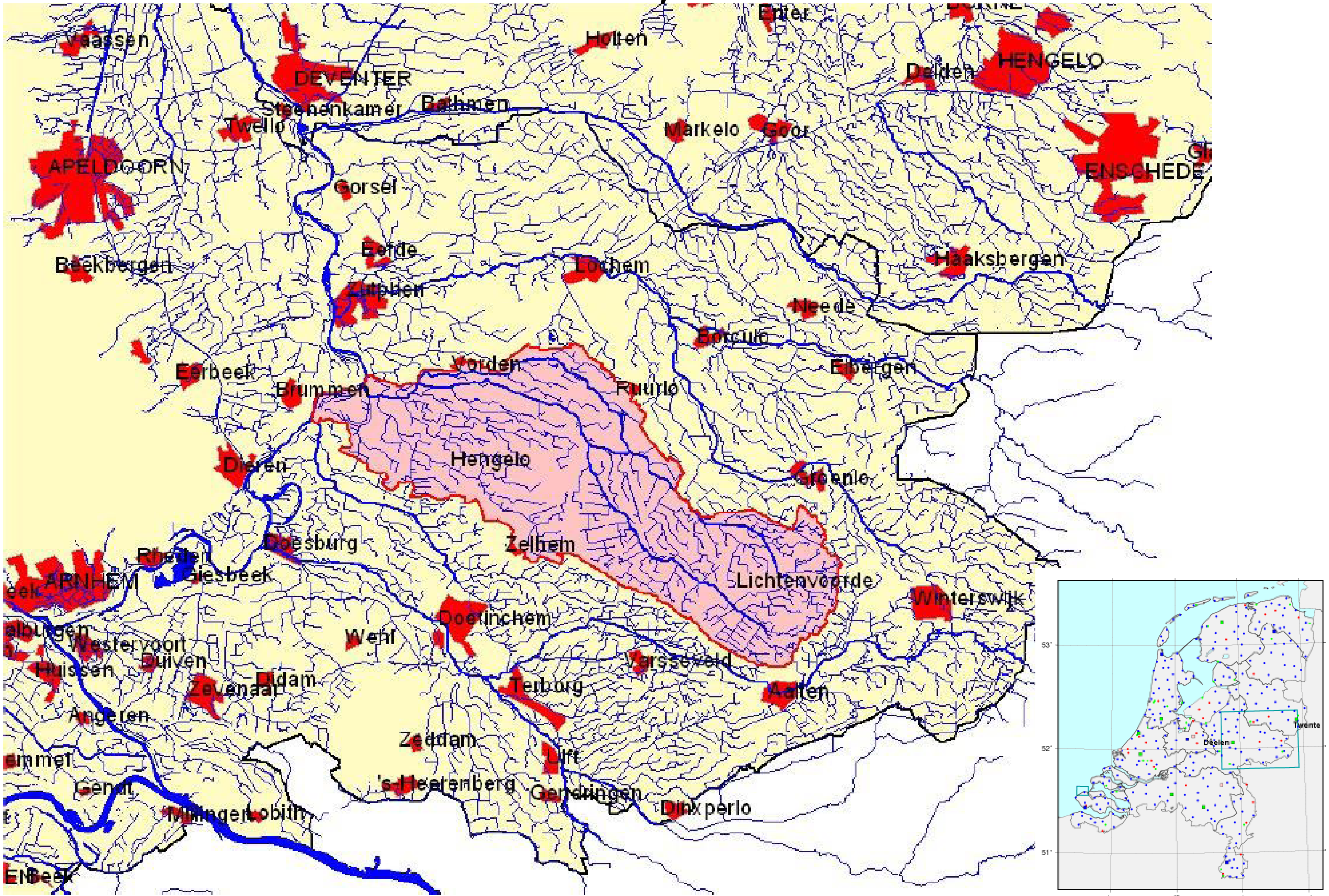
– alternatives and underpinning of choices regarding assumptions made;

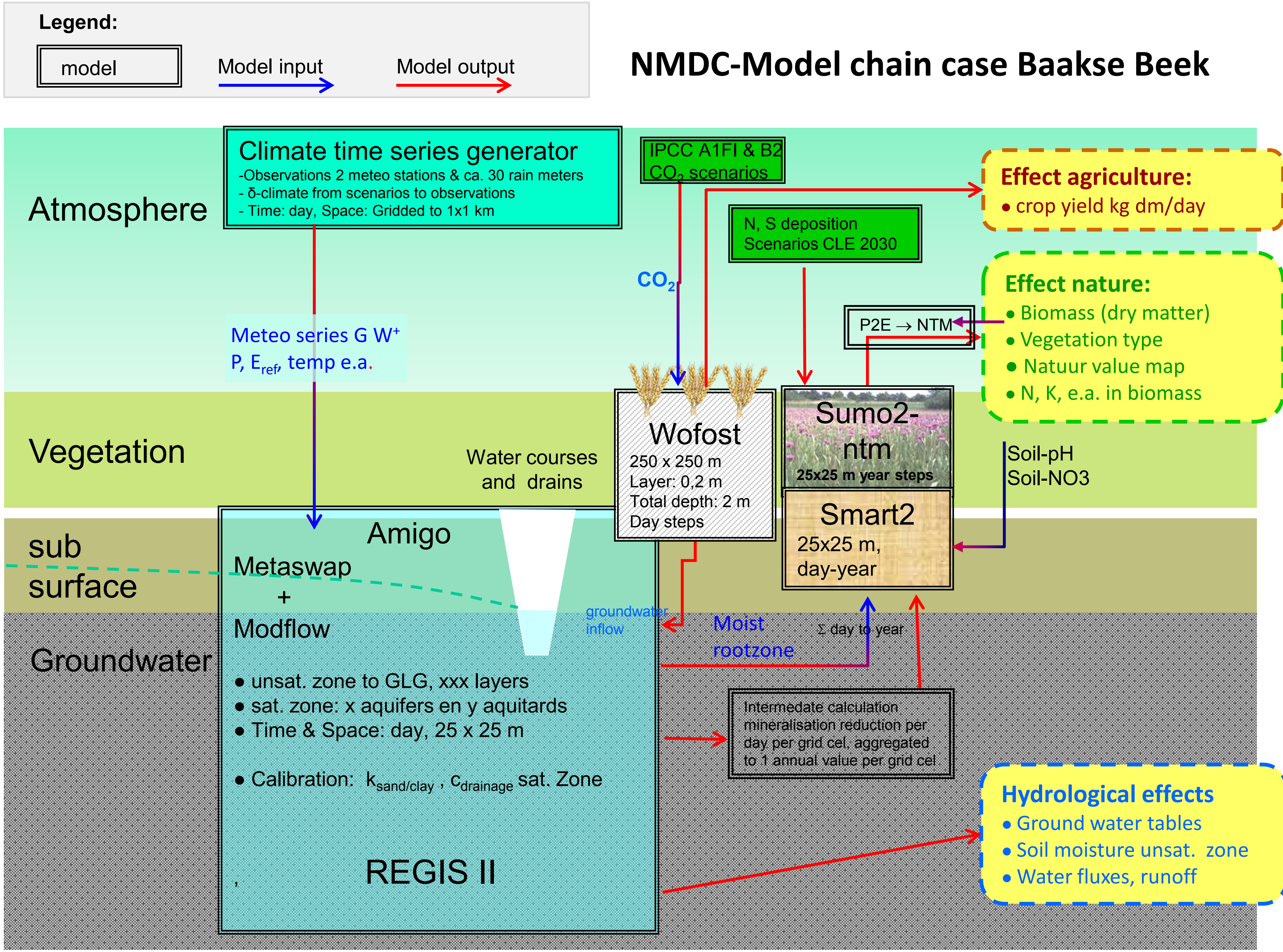
– influence of key-assumptions on results;

– implications in terms of robustness of results

Kloprogge et al., 2011 http://dx.doi.org/10.1016/j.envsoft.2009.06.009

Baakse Beek area, The Netherlands





Legend:
model
Model input
Model output
NMDC-Model chain case Baakse Beek
Atmosphere
Vegetation
sub surface
Groundwater
Climate time series generator
-Observations 2 meteo stations & ca. 30 rain meters
- δ-climate from scenarios to observations
- Time: day, Space: Gridded to 1x1 km
IPCC A1FI & B2
CO_2 scenarios
CO_2
N, S deposition
Scenarios CLE 2030
Effect agriculture:
• crop yield kg dm/day
Effect nature:
• Biomass (dry matter)
• Vegetation type
• Natuur value map
• N, K, e.a. in biomass
P2E → NTM
Meteo series G W+
P, E_{ref}, temp e.a.
Water courses and drains
Wofost
250 x 250 m
Layer: 0,2 m
Total depth: 2 m
Day steps
Sumo2-ntm
25x25 m year steps
Smart2
25x25 m,
day-year
Soil-pH
Soil-NO3
Amigo
Metaswap
+
Modflow
groundwater inflow
Moist rootzone
Σ day to year
• unsat. zone to GLG, xxx layers
• sat. zone: x aquifers en y aquitards
• Time & Space: day, 25 x 25 m
• Calibration: $k_{sand/clay}$, $c_{drainage}$ sat. Zone
REGIS II
Intermediate calculation mineralisation reduction per day per grid cel, aggregated to 1 annual value per grid cel
Hydrological effects
• Ground water tables
• Soil moisture unsat. zone
• Water fluxes, runoff

Identifying assumptions: think of…

- (over-) Simplifications of reality;
- Up / down scaling in the coupling of models;
- Variables kept constant (in time and space) in the model that vary in reality;
- Feedbacks excluded in the analysis;
- Processes kept outside the system boundary;
- Major sources of uncertainty.

Prioritization & critical appraisal

- 52 assumptions identified: “Gross list”
- 16 respondents each selected a top 10
- Aggregated into a “group top 10”

- Pedigree analysis (“strengths and weaknesses in the underpinning”) of each assumption in “group top 10”

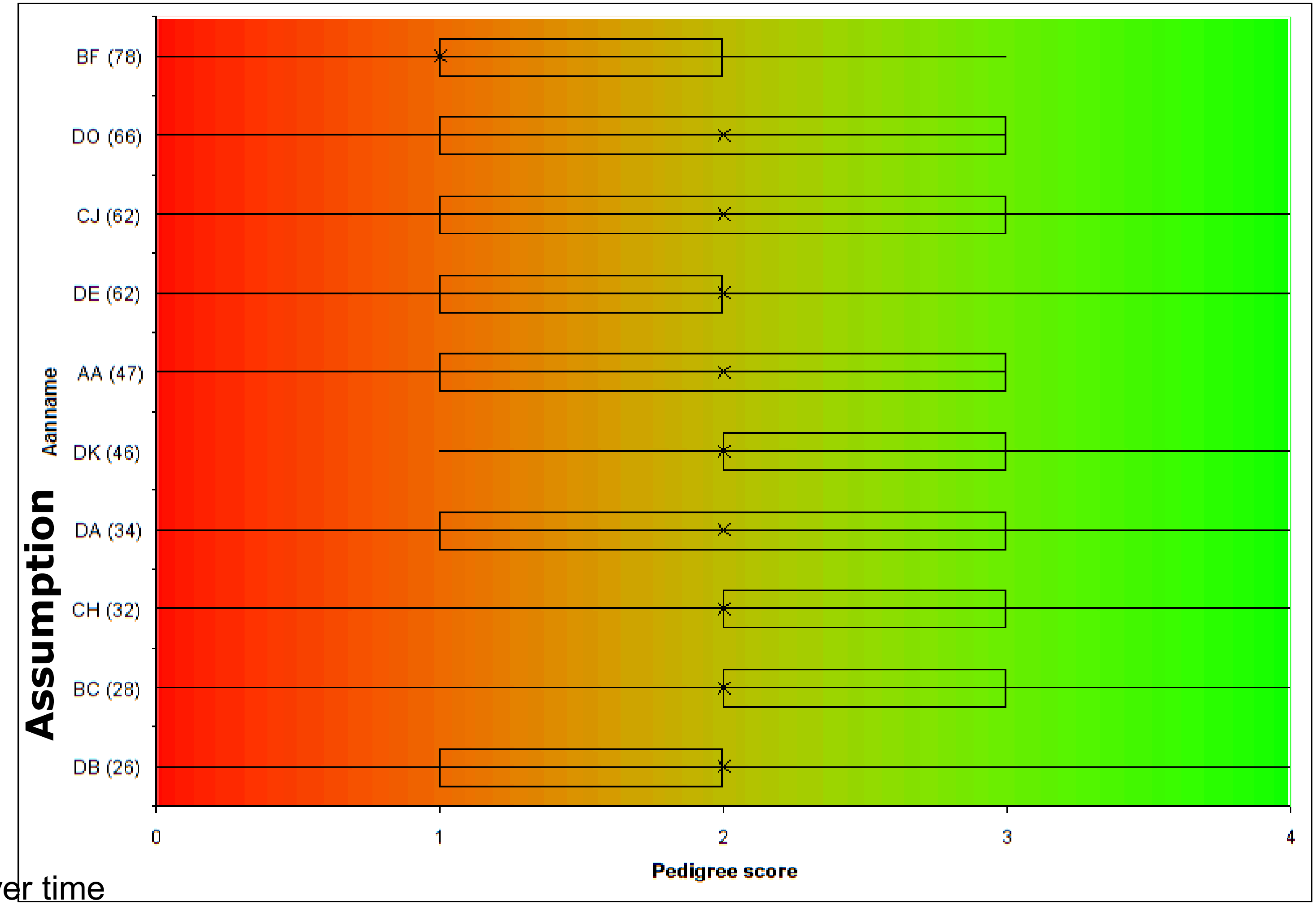
Table 1
Pedigree scheme used to assess assumptions during the workshop.

Score	Influence of situational limitations	Plausibility	Choice space	Agreement among peers	Agreement among stakeholders	Sensitivity to views of analyst	Influence on results
4	No such limitations	Very plausible	No alternatives available	Complete agreement	Complete agreement	Not sensitive	Little or no influence
3	Hardly influenced	Plausible	Very limited number of alternatives	High degree of agreement	High degree of agreement	Hardly sensitive	Local impact in the calculations
2	Moderately influenced	Acceptable	Small number of alternatives	Competing perspectives	Competing perspectives	Moderately sensitive	Important impact in a major step in the calculation
1	Importantly influenced	Hardly plausible	Average number of alternatives	Low degree of agreement	Low degree of agreement	Highly sensitive	Moderate impact on end result
0	Completely influenced	Fictive or speculative	Very ample choice of alternatives	Controversial	Controversial	Extremely sensitive	Important impact on end result

Example result pedigree scores

for one of the assumptions

Criteria		Number of votes for pedigree score						Median
		4	3	2	1	0		
a. Influence situational limitations	No such limitations			2	3	1	Completely influenced	1
b. Plausibility	Very plausible		1	4	1		Fictive or speculative	2
c. Choice space	No alternatives		1	4	1		Very ample choice of alternatives	2
d. Agreement among peers	Complete agreement					6	Low degree (controversial)	0
e. Agreement stake-holders	Complete agreement			6			Controversial	2
f. Sensitivity views and interests analyst	Not sensitive					6	Very highly sensitive	0
Total median pedigree score								1.5
g. Influence on results	Little or no influence					6	Important impact on end result	0



BF: land use constant over time
DO: drought stress within one year does not impact nature
CJ: feedbacks via market effects excluded
DE: Model coupling AMIGO-SMART2/SUMO2 around root zone
AA: Completeness of range of climate scenario's
DK: Coupling vegetation and hydrology
DA: Feedbacks via pests, weeds and plant diseases
CH: Developments in crop growth technologies not accounted for
BC: Conductivity of sub surface too homogeneous in the model
DB: Aggregation of daily values Amigo-hydrology to annual number for mineralisation reduction in SMART2/SUMO2

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.

Uncertainty matrix

UNCERTAINTY MATRIX			Level of uncertainty (from 'knowing for certain' (deterministic knowledge) to 'not even knowing what you do not know' (total ignorance))			Nature of uncertainty		Qualification of knowledge base (backing)			Value-ladenness of choices		
Location ↓			Statistical uncertainty (range+ chance)	Scenario uncertainty (range indicated as 'what-if' option)	Recognised ignorance	Knowledge-related uncertainty	Variability-related uncertainty	Weak –	Fair 0	Strong +	Small –	Medium 0	Large +
Context		Assumptions on system boundaries and ecological, technological, economic, social and political context											
Expert judgement		Narrative; storyline; advice											
Model	Model structure	Relations											
Model	Technical model	Software and hardware implementation											
Model	Model parameters												
Model	Model inputs	Input data; driving forces; input scenarios											
Data (in a general sense)		Measurements; monitoring; surveys											
Outputs		Indicators; statements											

Reporting

- Make uncertainties explicit
- Assess robustness of results
- Discuss implications of uncertainty findings for different settings of burden of proof
- Relevance of results to the problem
- Progressive disclosure of information -> traceability and backing

Insights on uncertainty

- More research tends to increase uncertainty
  - reveals unforeseen complexities
  - Complex systems exhibit irreducible uncertainty (intrinsic or practically)
- Omitting uncertainty management can lead to scandals, crisis and loss of trust in science and institutions
- In many complex problems unquantifiable uncertainties dominate the quantifiable uncertainty
- High quality $\neq$ low uncertainty
- Quality relates to **fitness for function** (robustness, PP)
- Shift in focus needed from reducing uncertainty towards reflective methods to explicitly cope with uncertainty and quality

AFTERNOON PRACTICUM

Group exercise

Assess the pedigree of the model used in the following study:

Extinction risk from climate change

(Thomas *et al.*, *Nature*, 8 January 2004)

Main message of this paper:

- In 2050, 15-37% of species ‘committed to extinction’ due to climate change for a mid-range climate scenario

Extinction risks from climate change

Species-Area relationship:

- numbers of species that become extinct or threatened by habitat loss from climate change

$$S = c\,A^{z}$$

S = number of species

A = area,

c = constant

$z \approx 0.25$

Ratio of number of species that can live in a habitat of area *A* before (0) and after (t) climate change ‘predicts’ extinction rate:

$$\frac{S_t}{S_0} = \frac{c\,A_t^{\ z}}{c\,A_0^{\ z}} = (A_t / A_0)^z$$

Species committed to extinction

Climate scenario 2050	universal dispersal	no dispersal
> 2.0 °C	21–32%	38–52%
1.8–2.0 °C	15–20%	26–37%
0.8–1.7 °C	9–13%	22–31%

(Thomas et al., 2004)

Rule of thumb

Warming rate 1°C / century corresponds to:

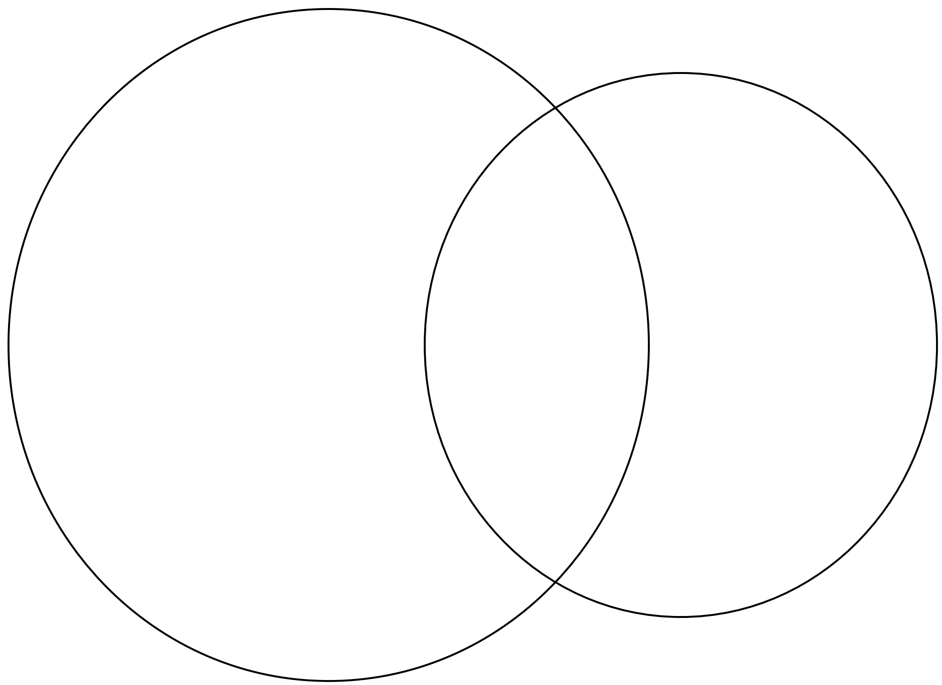
- ± 20 cm sea level rise
- ± 100 km shift of climate zone / century
- ± 150 m upward shift alpine climate zone/century

Climate tolerances of ecosystems

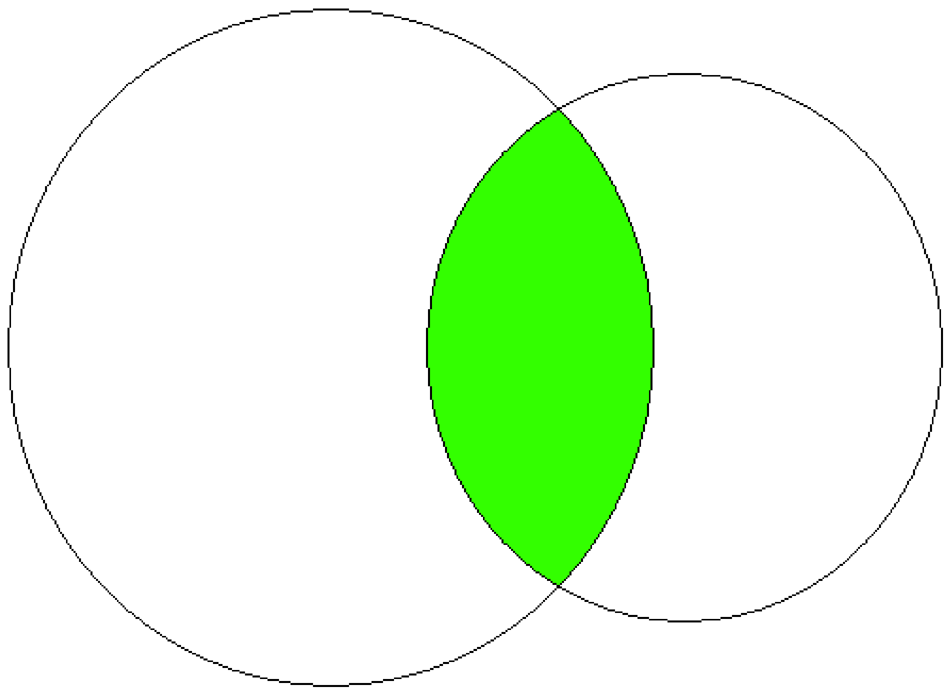
Ecosystem	Climate tolerance (°C/century)
Alpine ecosystem	0
Oak forest	0.12
Mangrove forest	0.50
Coastal wetlands	0.75
Coral reefs equator	1
Coral reefs N/S borders	5

(Hinkley, 1997)

Habitat before Climate Change **Habitat after Climate Change**

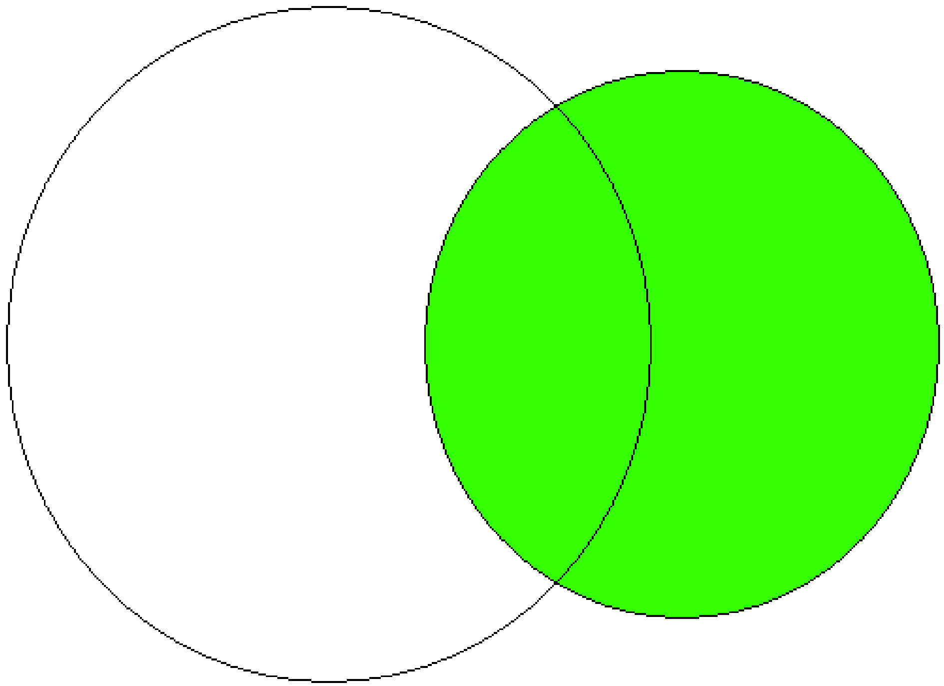


Habitat before Climate Change **Habitat after Climate Change**



Assumption: No dispersal

Habitat before Climate Change **Habitat after Climate Change**



Assumption: Full dispersal

Pedigree matrix for evaluating models

Score	Supporting empirical evidence		Theoretical understanding	Representa-tion of understood underlying mechanisms	Plausibility	Colleague consensus
	Proxy	Quality and quantity				
4	Exact measures of the modelled quantities	Controlled experiments and large sample direct measurements	Well established theory	Model equations reflect high mechanistic process detail	Highly plausible	All but cranks
3	Good fits or measures of the modelled quantities	Historical/field data uncontrolled experiments small sample direct measurements	Accepted theory with partial nature (in view of the phenomenon it describes)	Model equations reflect acceptable mechanistic process detail	Reasonably plausible	All but rebels
2	Well correlated but not measuring the same thing	Modelled/derived data Indirect measurements	Accepted theory with partial nature and limited consensus on reliability	Aggregated parameterized meta model	Somewhat plausible	Competing schools
1	Weak correlation but commonalities in measure	Educated guesses indirect approx. rule of thumb estimate	Preliminary theory	Grey box model	Not very plausible	Embrionic field
0	Not correlated and not clearly related	Crude speculation	Crude speculation	Black box model	Not at all plausable	No opinion

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 pedigree criterion at a time
- If you feel you cannot judge on of the pedigree criteria, leave it blank