Environmental Research Letters

CrossMark

OPEN ACCESS

RECEIVED
30 September 2014

REVISED
5 February 2015

ACCEPTED FOR PUBLICATION
13 March 2015

PUBLISHED
8 April 2015

LETTER

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]

[1] Centre for the Study of the Sciences and the Humanities, University of Bergen, Postboks 7805, NO-5020 Bergen, Norway
[2] Enonmental Sciences, Copernicus Institute of Sustainable Developmvirent, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

E-mail: Jeroen.Sluijs@svt.uib.no

Keywords: uncertainty, knowledge quality assessment, sensitivity auditing, assumption hunting, pedigree analysis, NUSAP

Supplementary material for this article is available online

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.

1. Introduction

Over the coming decades societies will simultaneously face further changes in the climate and in socio-economic conditions. This will have impacts on the quantity and quality of ground water and surface water and subsequently on agriculture and nature because these land use functions require certain amounts of fresh water of sufficient quality. In order to enable anticipation and proactive adaptation, decision makers and local stakeholders increasingly seek adequate information and foresight (Dessai and Van der Sluijs 2011). Typically, the spatial scale of spatial planning decisions in The Netherlands is far below the size of a grid cell in present day Global Climate Models, which poses major challenges in linking climate model projections to local decision making The current suite of tools available to Dutch knowledge institutes tasked with supplying information and foresight includes models of ground water, surface water, hydrological impact, agricultural yield, and vegetation. At present, these separate models are not fit for the function of informing local climate adaptation policies. Further, uncertainties in these models have insufficiently been assessed and reported, which limits the reliability of these models as a basis for decision making. A full blown assessment going all the way from global climate change scenarios to long term projections of local hydrological impacts on agriculture and nature requires major steps forward in the integration and linking of a whole chain of sub-models.

To address these challenges, major Dutch knowledge institutes in the fields of climate and hydrology established the Netherlands Models and Data-Centre (NMDC) in 2010. The NMDC initiated an innovation project 'integrated water management: from critical zone towards critical uncertainties' aiming to develop a method to link hydrology models to crop yield and vegetation models across multiple spatial and temporal scales in order to better meet the information needs of local adaptation policy making (Schipper *et al* 2013). Such an integration poses major challenges to the assessment of uncertainty related to model parameters, model structure, model assumptions, and up- and down scaling in the linking of models that run with mismatching time-steps and mismatching grid cell size. In such projects, Knowledge Quality Assessment (Clark and Majone 1985, Funtowicz and Ravetz 1990, Van der Sluijs *et al* 2005, 2008, Risbey *et al* 2005, Maxim and Van der Sluijs 2011) and other forms of sensitivity auditing of the models (Saltelli *et al* 2013) are essential.

The aim of the study presented in this paper is to systematically assess uncertainties and assumptions in a linked chain of models to project local impacts on hydrology, agriculture and nature in an area in the East of the Netherlands known as Baakse Beek which is a catchment area of about 270 km^2. Note that the authors have not been part of the modelling team and that the present paper presents an *ex-post* critical appraisal of assumptions made in the modelling experiment that has been selected as the object of analysis of the present study. The model experiment itself has been published elsewhere (Schipper *et al* 2013).

The Baakse Beek area is representative for the eastern and southern Pleistocene cover landscape which covers about half of the Netherlands. It typically has sandy soils and has a patchy land-use including small to medium scale agriculture (dairy farming with pasture and maize fields) and nature (about 13% of the area). Over the past decades, nature restoration has become a policy issue in the area and for the implementation of the national ecological network, agricultural lands will have to be converted to nature. (Witte *et al* 2014)

Key outputs of the model-chain implemented for this area are (1) hydrological impacts in terms of ground water levels, soil moisture, water fluxes and stream runoff; (2) agricultural impacts in terms of changes in crop yield; (3) impacts on nature in terms of biomass stock, vegetation type, nature value, and N, K, etc in biomass. The main input is formed by climate scenarios for the Netherlands as developed by the KNMI.

Box 1. Steps in the critical appraisal of assumptions in chains of linked models (Kloprogge *et al* 2011).

Analysis

1. Identify explicit and implicit assumptions in the calculation chain;
2. Identify and prioritise key assumptions in the chain;
3. Characterize assumptions and assess their potential value-ladenness;
4. Identify 'weak' links in the calculation chain;
5. Further analyse potential value-ladenness of the key assumptions.

Revision

6. Revise/extend assessment (a) sensitivity analysis to key assumptions; (b) diversification of assumptions; (c) replace weak assumptions by more realistic ones;

Communication

7. Communication of the findings (a) make key assumptions explicit; (b) discuss alternatives and underpinning of choices regarding assumptions made; (c) explain the potential influence of key assumptions on model results; (d) discuss implications in terms of robustness of result.

The central research questions in this paper is: what are the most critical modelling assumptions that underlie the linked chain of models that has been implemented by the NMDC project for the case study Baakse Beek?

To answer this question the following sub questions need to be addressed: (1) What are the implicit and explicit modelling assumptions in the chain of models? (2) What is the strength of scientific underpinning of these assumptions made? and (3) How sensitive are the key model outcomes to these assumptions?

2. Method

We applied the method for critical review of assumptions in model-based assessments, originally developed by Kloprogge *et al* (2005, 2011). This method has been applied to an increasing range of case studies where risk assessment was similarly based on a linked chain of models such as assessment of externalities of nuclear energy (Laes *et al* 2011), quantitative microbial risk assessment and food safety (Boone *et al* 2010, Bouwknegt *et al* 2014), and health risk assessment of overhead power lines (De Jong *et al* 2012). Structured critical analysis of assumptions yields valuable insight into the modelling process and structure, and allows for a prioritization of assumptions based on their relative weakness and influence. The result can then be used to further improve the modelling process.

The method systematically identifies, prioritises and analyses importance and strength of assumptions in calculation chains of linked models (box 1).

Typically chains of soft-linked computer model calculations are used that start with scenarios for population and economic growth. The models in the chain vary in complexity, and were originally constructed with different purposes in mind. Often, the construction of these calculation chains involves many analysts from several disciplines. Many assumptions have to be made in combining research results in these calculation chains, especially since the output of one computer model often does not fit the requirements of input for the next model (scales, aggregation levels). Assumptions are also frequently applied to simplify parts of the calculations. Assumptions can be made explicitly or implicitly.

Such assumptions can to some degree be value-laden, arbitrary, or weak in other aspects. This method distinguishes four types of value-ladenness of assumptions: value-ladenness in a socio-political sense (e.g., assumptions may be coloured by political preferences of the analyst), in a disciplinary sense (e.g., assumptions are coloured by the discipline in which the analyst was educated), in an epistemic sense (e.g., assumptions are coloured by the approach/methods that the analyst prefers), and in a practical sense (e.g., the analyst is forced to make simplifying assumptions due to time constraints).

In order to characterize the assumptions (box 1, step 3) a so called pedigree matrix (Funtowicz and Ravetz 1990, Van der Sluijs *et al* 2005) has been applied. Pedigree is part of the so called Numeral, Unit, Spread, Assessment, Pedigree (NUSAP) analytical and notational system that aims to jointly address quantifiable and unquantifiable uncertainties in scientific knowledge (Funtowicz and Ravetz 1990, van der Sluijs *et al* 2005, Dankel *et al* 2012). It extends the classic notational system for quantitative scientific information (usually provided as a number, a unit, and a standard deviation) with two additional qualifiers: expert judgement of the reliability (the assessment) and a multi-criteria characterization reflecting the origin and status of the information (the pedigree). The classical notational system does not reveal the distinction between nearly perfect information (such as the speed of light) and highly imperfect information (such as the local rain fall in April in 50 years from now). The two additional qualifiers, assessment and pedigree, attempt to remedy this problem. The pedigree analysis is a qualitative structural process to clarify the knowledge base on which scientists and stakeholders frame their perceptions of a problem, by appraising the information underpinning the numbers, theories and assumptions that form the basis of scientific advice, often model-derived (Dankel *et al* 2012).

Here we used the pedigree matrix for characterizing model assumptions (table 1) as proposed by Kloprogge *et al* (2005, 2011) and modified by Craye *et al* (2009) and Laes *et al* (2011).

The pedigree criteria used here to critically reflect on the assumptions are:

- *Influence of situational limitations*
  The degree to which the choice for the assumption can be influenced by situational limitations, such as limited availability of data, money, time, software, tools, hardware and human resources.
- *Plausibility*
  The degree, mostly based on an (intuitive) assessment, through which the approximation created by the assumption is in accordance with 'reality'.
- *Choice space*
  The degree to which alternatives were available to choose from when making the assumption
- *Agreement among peers*
  The degree to which the choice of peers is likely to coincide with the analyst's choice.
- *Agreement among stakeholders*
  The degree to which the choice of stakeholders is likely to coincide with the analyst's choice
- *Sensitivity to view and interests of the analyst*
  The degree to which the choice for the assumption may be influenced, consciously or unconsciously, by the view and interests of the analyst making the assumption.
- *Influence on results*
  In order to be able to pinpoint important value-laden assumptions in a calculation chain it is not only important to assess the potential value-ladenness of the assumptions, but also to analyse the influence on outcomes of interest of the assessment.

A more elaborated introduction to the pedigree method followed is provided as supplementary text 1 (stacks.iop.org/ERL/10/045005/mmedia).

We identified assumptions and other key sources of uncertainty (drafting of a gross list) by means of document analysis and interviews with six experts involved in the development of the tailored chain of models for the Baakse Beek area. The documents studied are: Gommes (1998), Van der Linden *et al* (2008), Bessembinder *et al* (2011), Geertsema *et al* (2011), Schipper *et al* (2011), Sluiter (2011), Wamelink *et al* (2011), Bakker and Bessembinder (2012a, 2012b), Bessembinder (2012), Groot *et al* (2012), and Van Ek *et al* (2012). All experts were asked to check the completeness and correctness of formulations of the assumptions identified.

During the interviews respondents were explicitly encouraged to think in particular of assumptions related to:

- (over-) Simplifications of reality;
- up/down scaling in the linking of models;
- variables kept constant (in time and space) in the model that vary in reality;

Table 1. Pedigree matrix for characterization of model assumptions.

Score	Influence situational limitations	Plausibility	Choice space	Agreement among peers	Agreement stake-holders	Sensitivity views and interests analyst	Influence on results
4	No such limitations	Very plausible	No alternatives available	Complete agreement	Complete agreement	No 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 schools	Competing perspectives	Moderately sensitive	Important impact in a major step in the calculation
1	Importantly influence	Hardly plausible	Average number of alternatives	Low degree (embryonic stage)	Low degree of agreement	Highly sensitive	Moderate impact on end result
0	Completely influenced	Fictive or speculative	Very ample choice of alternatives	Low degree (controversial)	Controversial	Very highly sensitive	Important impact on end result

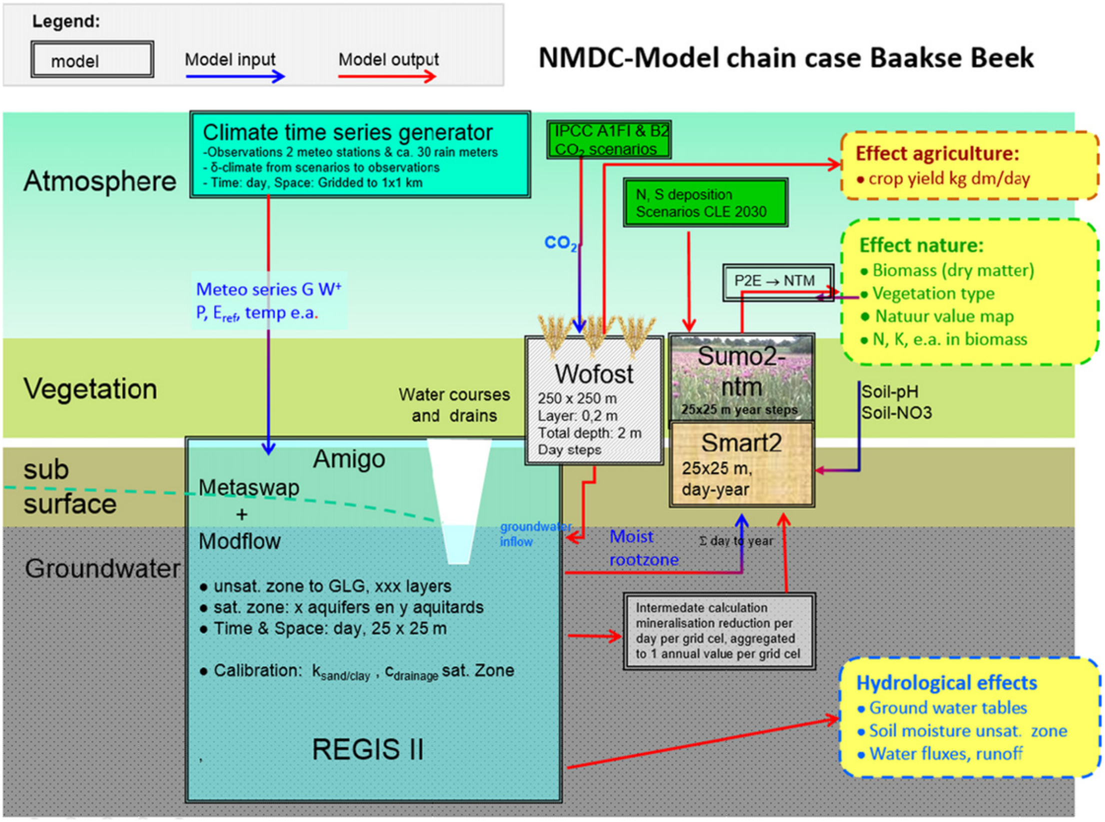

Figure 1. Chain of linked models implemented for the Baakse Beek case area.

- feedbacks excluded in the analysis;
- processes kept outside the system boundary;
- major sources of uncertainty.

In drafting the gross list, strongly overlapping assumptions were merged. In a next step a broader group of experts has been consulted by e-mail with the question to select a top 10 of the most critical assumptions in the list. External experts not familiar with the model chain were provided with a briefing document on the project. In total, 16 experts responded and provided a top 10. The resulting individual top 10 lists of each expert where combined into a group ranking of the gross list of assumptions. From this group ranked gross-list, the group top 10 assumptions were selected for the next step.

Using the open source software LimeSurvey (www.limesurvey.com) we implemented an online survey where respondents could characterize each of the assumptions from the group top-10 using the criteria and 5-point ordinal scoring scales as defined by the pedigree matrix (table 1). Respondents were asked to make explicit their reasoning and arguments for each score they gave. We received responses from 19 experts.

2.1. The chain of models for the Baakse Beek case

Figure 1 shows the models used in the Baakse Beek case and how they have been linked. The end points of the model chain are the long term local impacts of climate change on hydrology, agriculture and nature in the Baakse Beek catchment area. To this end it compares model outcomes for two input scenarios for future climate time series for the period 2035–2065 with outcomes obtained when using the historic time series 1981–2010 as input. We will briefly describe each of the sub models used.

2.1.1. *KNMI Climate scenario's and future weather generator*

The Royal Netherlands Meteorological Institute has developed climate change scenarios for the Netherlands. In this study the 2006 versions of the scenarios were used (Van den Hurk *et al* 2006). These project average changes in temperature (annual, summer, winter), precipitation (summer, winter), summer evaporation and sea level rise in the years 2050 and 2100 for the whole of the Netherlands. To assist local climate impact and adaptation studies that need to account for climate variability, KNMI has developed a tool with which one can generate for each of the four scenarios high resolution virtual time-series of daily values of climate parameters in future years for all Netherlands whether stations and rain fall metres for each of the four National climate scenarios. This is done in such a way that these future time series per weather station are consistent with the national averages in the year 2050 as prescribed by the scenarios. For an individual weather station this is done by taking the 30 year historic observational time-series of data of the period 1981–2010 which is then transformed by the tool to a future time series with the same temporal variability but with a shift in the average that is consistent with the scenario for which the transformation is done (Bakker *et al* 2012a). There are 30 rainfall metres and 2 weather stations In the Baakse Beek case study area.

2.1.2. Subsurface model REGIS II and ground water model AMIGO

The regional hydrology model Actueel Model Instrument Gelderland Oost (AMIGO) covers the area between the River Ijssel and the border with Germany which includes the Baakse Beek area. AMIGO exists of a MODFLOW model (Harbaugh *et al* 2000) describing ground water flows and a model of the unsaturated zone named metaSWAP (van Walsum and Groenendijk 2008) whereas the subsurface is based on the National hydrogeological model REGIS II (Vernes and Van Doorn 2005) that distinguishes 12 layers of respectively aquifers and aquitards in the case study area. REGIS has a spatial resolution of 100 × 100 m. AMIGO has a spatial resolution of 25 × 25 m and time-steps of one day. MetaSWAP distinguishes 2 layers of the unsaturated zone: the root zone and the zone between the root zone and the ground water level.

2.1.3. Crop yield model WOFOST

WOrld FOod Studies (WOFOST; van Walsum and Supit 2012) is a dynamic crop growth model that calculates potential crop yield. Here it uses the hydrology of the unsaturated zone provided by MetaSWAP as input, together with the KNMI climate time series as well as a scenario for future CO_2 concentrations in order to account for the CO_2 fertilization effect on crop growth.

2.1.4. Nature model: SMART2/SUMO2

To assess the hydrological impacts on nature, the soil chemistry model SMART2 and vegetation model SUMO2 were used (Wamelink *et al* 2011). SUMO2 models the development of vegetation over time, including the ecological process of succession. Its output is biomass and vegetation structure whereas the input includes annual average temperature, annual average precipitation, annual average CO_2 concentration, nutrient availability and vegetation history. The spatial resolution is 25 × 25 m and the time step is one year. SUMO2 is implemented as a subroutine of the soil chemistry model SMART2 that models amongst others soil pH and key nutrients such as NO_3. Together with the ground water levels calculated by AMIGO, SMART2/SUMO2 feeds into a nature-value model named P2E NTM that produces a nature-value map of the area. To this end P2E NTM applies so called Ellenberg indicators (Ellenberg *et al* 1991), a widely used approach to characterize nature value based on six gradients: soil acidity, soil productivity or fertility, soil humidity, soil salinity, climatic continentality and light availability. The resulting nature value map shows plant diversity and potential for red list species.

3. Results

We identified 51 assumptions (full list in Van der Sluijs *et al* 2012). From this gross list, 16 respondents each selected a top-10 of the most critical assumptions. 46 assumptions of the gross list figured in the top-10 of at least one respondent; 35 were in the top-10 of at least 2 respondents; 27 were in the top-10 of at least 3 respondents; 7 were in the top-10 of at least 8 (=50%) of the respondents. This shows that there is a tendency of convergence in the group of respondents.

The 16 individual top-10 lists of assumptions ware combined into a group top-10 by giving 10 points to the number 1 in each list, 9 points to number 2, etc. The resulting group ranking is presented in table 2.

A broader group of experts was asked to critically appraise and characterize each of the assumptions in the group top-10 using the pedigree criteria and ordinal scoring scales from table 1. Table 3 presents the pedigree scores obtained for the first assumption in the list. The full set of scores is presented in supplementary text 2 (stacks.iop.org/ERL/10/045005/mmedia). In total 18 experts responded and provided pedigree assessments. Some respondents skipped some assumptions on which they had insufficient expertise and/or criteria they could not judge, which explains why the sum of number of experts in table 3 and in the frequency tables in the supplementary text is not always 18.

Here we discuss the main weaknesses of the assumptions as identified by the respondents.

3.1. Land use and land management constant over time

In making this assumption, situational limitations were important while the plausibility is limited. It could well be the case that climate impacts on nature will differ substantially under different land use and land management scenarios. In theory it is possible to include future land use and land management scenarios in the Baakse Beek model chain but this was not feasible with the available time and money. For the hydrology, the assumption is defendable because earlier hydrological studies have identified climate change as a much stronger factor than land use. But for the impacts on agriculture and nature this may differ. Further, there can be a feedback of climate change on the behaviour of farmers and nature managers (autonomous adaptation). Modelling techniques exists by which such feedbacks can be explored such as complex adaptive systems (CAS) and Agent Based Modelling. Such techniques may provide insight in the systems

Table 2. Group top-10 of model assumptions.

Rank	Assumption	In top 10 of # respectively	∑ points
1	Land use and land management have been kept constant over time. Spatial developments in the area are neglected. The ways in which humans make use of the water system is kept constant, it is assumed that agricultural practices and nature management do not adapt to the changing climate.	11	78
2	Drought-stress within one model year is neglected in the impact on nature. In years that are on average not dry, there can be sequences of dry days that are long enough to cause damage to crops and vegetation. The SMART SUMO module has of a year, thus the vegetation only responds to annual average ground water levels. Years with big differences in highest and lowest ground water level could damage the vegetation via drought stress, even while the average level does not reflect this.	11	66
3	Feedbacks on agriculture via market effects have not been considered. When converted from kg to Euro, changes in crop yields can show a complete different pattern. It is expected that market effects are climate dependent and will thus lead to different outcomes for different climate scenarios (crop failures in other parts of the world impact crop prices on world market).	9	62
4	Model linking of AMIGO and SMART2/SUMO2 around the root zone. The way this linking is now implemented produces major uncertainties about quantity and quality of ground water that reaches the root zone. The quantity can be partially derived from Amigo output but the lateral flows are omitted. SMART2/SUMO2 bases the water quality on an obsolete map with outdated data. For the calculation of soil pH (and thus mineralization) these factors are crucial.	10	62
5	Incompleteness of the range of climate scenarios used. The ranges spanned up by the two scenarios selected (G en W+) insufficiently covers the uncertainty and variability of climate around the year 2050.	6	47
6	Linking of vegetation back to hydrology missing. The model chain has no feedback loop from the projected changes in the vegetation as calculated by SMART2/SUMO2 back to the hydrology in AMIGO. In reality such feedbacks occur, for instance if the area dries out, the vegetation structure shifts to a more open landscape. On the sandy soils in the case area this would impact the soil hydrology, but this mechanism is not included in the model linking.	8	46
7	Feedbacks via plant diseases, pests and weeds have been omitted. Plant diseases, pests and weeds can reduce vegetation development and damage crops and can change nutrient composition (particularly N) and protein content of plants. This in turn can change the attractiveness to herbivores. Shifting climate zones may make the area attractive to (plague) insects and herbivores presently not abundant in the region. The way the model chain is implemented implicitly assumes that these growth limiting factors (plant diseases, pests and weeds) are constant over time and insensitive to the climate.	7	34
8	Developments in agricultural technology and practices has been omitted. It is thus implicitly assumed that these factors are constant over time and insensitive to changes in the climate. This leads to an underestimation of future crop yield.	6	32
9	Hydraulic conductivity of subsurface (Kh and Kv) in the model is too homogeneous compared to reality.	4	28
10	Aggregation algorithm of daily AMIGO hydrology into yearly values of mineralization reduction feeding into SMART2/SUMO2. To link AMIGO to SMART an aggregation algorithm was implemented which calculates a so called 'mineralization reduction factor' to account for the joint impact of soil moisture and pH on the mineralization of nutrients in soil organic matter. For each grid cell, the daily soil moisture projections from AMIGO are combined with the annual average soil pH from SMART2 to produce a daily value for the 'mineralization reduction factor' which is subsequently aggregated to a yearly average 'mineralization reduction factor' which feeds into SUMO2.	6	26

dynamics but they can hardly be validated, thereby limiting their predictive power.

3.2. Drought-stress within one model year is neglected in the impact on nature

The fact that SMART2/SUMO2 calculates in time steps of 1 year implicitly implies that drought stress on the vegetation that occurs within one model year (e.g. during dry spells in warm summers) during a critical period of the year but that is not visible in the years' average hydrology is neglected. This is of particular concern in the Baakse Beek area because its sandy soils with superficially rooting vegetation renders it sensitive to such stress. The present model chain may thus underestimate the impacts on nature. It would not be easy to reduce the time step size of SMART2/SUMO2

Table 3. Pedigree scores for the assumption that land use and land management are constant over time.

		Number of experts that gave score 4,3,2,1 or 0						
Criteria		4	3	2	1	0		Median score
a. Influence of situational limitations	No such limitations	**1**	**1**	**1**	**11**	**4**	Completely influenced	1
b. Plausibility	Very plausible	**0**	**1**	**4**	**9**	**4**	Fictive or speculative	1
c. Choice space	No alternatives available	**2**	**4**	**6**	**4**	**2**	Very ample choice of alternatives	2
d. Agreement among peers	Complete agreement	**0**	**6**	**4**	**4**	**4**	Highly controversial	2
e. Agreement among stakeholders	Complete agreement	**0**	**1**	**10**	**4**	**2**	Highly controversial	2
f. Sensitivity to views and interests analyst	Not sensitive	**4**	**2**	**5**	**4**	**2**	Very highly sensitive	2
g. Influence on result	Negligible influence	**0**	**2**	**3**	**7**	**6**	Important impact on end result	1

to solve this problem and doing such would add substantial complexity to the model for which the necessary data for model calibration are not available either. As a work-around one could perhaps construct an extra drought-stress indicator in the step where daily hydrology as projected by AMIGO is aggregated into the annual aggregate hydrology numbers that feed into SMART2/SUMO2.

3.3. Feedbacks on agriculture via market effects have not been considered

The model chain projects potential crop yield in kg of harvested products for the climate scenarios considered. Feedbacks of prices, subsidies and the world crop market on the farmer's behaviour and crop choice is not considered in the model but can well occur in reality. Under different climate scenarios this can differently affect the crop yield. At present, the expertise required to model such feedbacks does not exists and the market effects are highly unpredictable. This major source of uncertainty seem irreducible and can only be superficially explored by speculative scenarios.

3.4. Model linking of AMIGO and SMART2/SUMO2 around the root zone

The developers of SMART2/SUMO2 are well aware of the problems due to the way in which the model is currently linked to AMIGO around the root zone and the soil-chemical process of mineralization (the process by which nutrients in organic matter are converted into plant accessible forms). In principle this can be resolved but the data required for this are not available and time and resources are lacking to implement the conceptual models into operational models that are calibrated to local data of the case study area. The uncertainty regarding water quality in the root zone could be reduced by a dedicated measurement campaign. Uncertainties remain regarding the role of pH in the process of mineralization because the modelled pH often does not match the observed pH.

3.5. Incompleteness of the range of climate scenarios used

The study only used two (G and W+) of the four KNMI 2006 scenarios. It might have been better to use all four. However, the goal of this pilot project was to gain experience in linking models in order to support local adaptation decision making. For that purpose it was not deemed necessary to include all four scenarios.

3.6. Linking of vegetation back to hydrology missing

The model chain misses a feedback from the vegetation as projected by SUMO2 onto the hydrology in AMIGO and vice versa. This is a major model structure error. For instance, an increase in open landscape would lead to increased ground water supplementation and less evaporation and evapotranspiration. In theory the linking could be implemented in the model chain but it would substantially complicate the model calculations. AMIGO requires a lot of calculation time and the linking has to be implemented real-time two-way while at present it is a soft linking. A further challenge in achieving real-time linking is the gap in time step that has to be bridged (AMIGO runs in daily time steps while SMART2/SUMO2 has time steps of a year). This would unavoidably require additional assumptions that can hardly be validated. Respondents indicate that they are not sure how one should proceed with implementing these feedbacks and that it may require the development of a new, more integrated model instead of linking the existing (sub-optimally matching) models.

3.7. Feedbacks via plant diseases, pests and weeds omitted

Plant diseases, pests and weeds that can limit crop yield and vegetation development have been kept outside the model, mainly for pragmatic reasons. So many different diseases, pests and weed exist that there is no obvious way to account for them in the model chain, especially given the long time horizon considered. Knowledge about the impacts of climate change of plant diseases, pests and weeds is scarce; this is an underexplored area. We have not yet reached a stage where conceptual models exist, let alone operational models that can be implemented for the case study area. Accounting for this feedback by introducing speculative assumptions might add more uncertainty to the model than it resolves. Still we have to keep in mind that this feedback could significantly influence the model outcomes and that this impact may be different under different climate scenarios.

3.8. Developments in agricultural technology omitted

Developments in agricultural technologies are hard to predict but such developments may significantly change future crop yield under different climate scenarios. One could easily include trend extrapolation but doing so would overlook discontinuities and surprises that are common in technological innovation. Two directions of improvement are at play here: those that increase crop yield and those that increase drought resistance of crops. The first factor will not differ substantially across different climate scenarios whereas the second will indeed lead to different outcomes. Overall there seems no non-speculative way to incorporate this factor in the model chain.

3.9. Hydraulic conductivity of subsurface in the model is too homogeneous compared to reality

In the making of this modelling assumption, the key issue has been data-availability. To map the sub surface in high resolution, a large number of soil drillings would be required to obtain the data but such drillings would perturb and impact the permeability of the same subsurface which poses a major dilemma. Respondents indicated that local data exists that have not yet been utilized in the model, so there is some potential for improvement. However this is a labour intensive process for which resources are lacking. Alternatively, a stochastic model of the subsurface could be used, such as the one developed by TNO.

3.10. Aggregation algorithm from daily AMIGO hydrology into yearly values of mineralization reduction feeding into SMART2/SUMO2

The way in which the difference in time step of the AMIGO model (day) and the SMART2/SUMO2 model (year) is bridged is based on an implicit assumption that seems to underestimate the mineralization in the soil. The role of soil pH here is also clouded by uncertainties and imperfections. In theory SMART2/SUMO2 could be rebuild to a model that takes smaller time steps (season, month, decade or day) but that poses dilemmas. It would require the inclusion of several other processes that can no longer be neglected at that level of temporal resolution and it will go at the expense of calculation time of the model runs which will limit the usability of the model. An intermediate solution could be to change the order of aggregation where for each day the mineralization is calculated and next the calculated daily mineralization is aggregated into a yearly number (in the present model chain the daily hydrology is aggregated into a yearly hydrology and next the yearly mineralization is calculated). Doing so would unavoidably introduce new assumptions that are also imperfect in numerous ways.

Figure 2 presents the aggregated pedigree scores of the top-10 assumptions. The scores are plotted as so called Box and Whisker plots. This depicts for each assumption the joint spread in scores over the six criteria and over all respondents. The lower the pedigree score is, the weaker the underpinning of the assumption is. This is indicated with a colour gradient running from red (weak) to green (strong).

4. Discussion

A complicating circumstance in the assessment of assumptions presented here was that it ran in parallel with the development of the very object of this study: the Baakse Beek model chain. In the phase where we made an inventory of the assumptions, major design choices as to how the modellers would implement the precise linkings between the sub models were still in the making. On the other hand this had the advantage that the modelling team could immediately benefit from the insights generated by the critical appraisal of their work into which we engaged them.

In the prioritization step some of the respondents noted the ambiguity in the question we asked: 'What is the top 10 of assumptions that according to your expert judgement have the highest influence on the outcomes of the model chain for the Baakse Beek case: the simulated impacts of climate change on agriculture, nature and hydrology around the year 2050'. It refers to impact sectors (agriculture, nature, hydrology) rather than actual quantitative indicators as calculated by the model, so respondents may have had diverging interpretations of this question. However, the goal of the critical appraisal of model assumptions is to gain insight in how these assumptions limit the quality and reliability of the model outcomes. Each respondent will have made a different (implicit) weighing of the three impact sectors. Given the high diversity in disciplinary background of the respondents we do not expect any serious bias towards one of the impact sectors in the resulting group top-10.

The ambiguity in the prioritization question is related to the ambition of the NMDC model linking project, namely to create a single decision support tool to support a wide range of policy questions and challenges. This tendency—which is not unique for this model integration project—is in big contrast with the notion put forward by Beck *et al* (1997) that model quality (in terms of fitness for function) requires that models must change when the question put to them changes.

Many respondents indicated that the gross list of assumptions in itself provided an insightful result and that they considered the systematic critical reflection that was facilitated by our method as very useful and instructive.

5. Conclusions

The most important assumptions in the linked model chain to project local climate impacts for the Baakse Beek case area have been identified, prioritized and characterized by means of a qualitative uncertainty assessment. From a gross list of 51 assumptions a group of experts selected a top-10 which were subsequently critically appraised by pedigree analysis to assess their strengths and weaknesses. The most problematic assumption turned out to be the assumption that land use and land management were kept constant over time. This assumption ended highest in the group top-10 and received the lowest score in the pedigree analysis (median score of 1 on an ordinal scale 0–4). It must be noted that this is not so much an uncertainty in the structure of the model chain itself, but rather an incomplete exploration of input scenarios by considering only one ('constant over time') scenario for this factor. We recommend

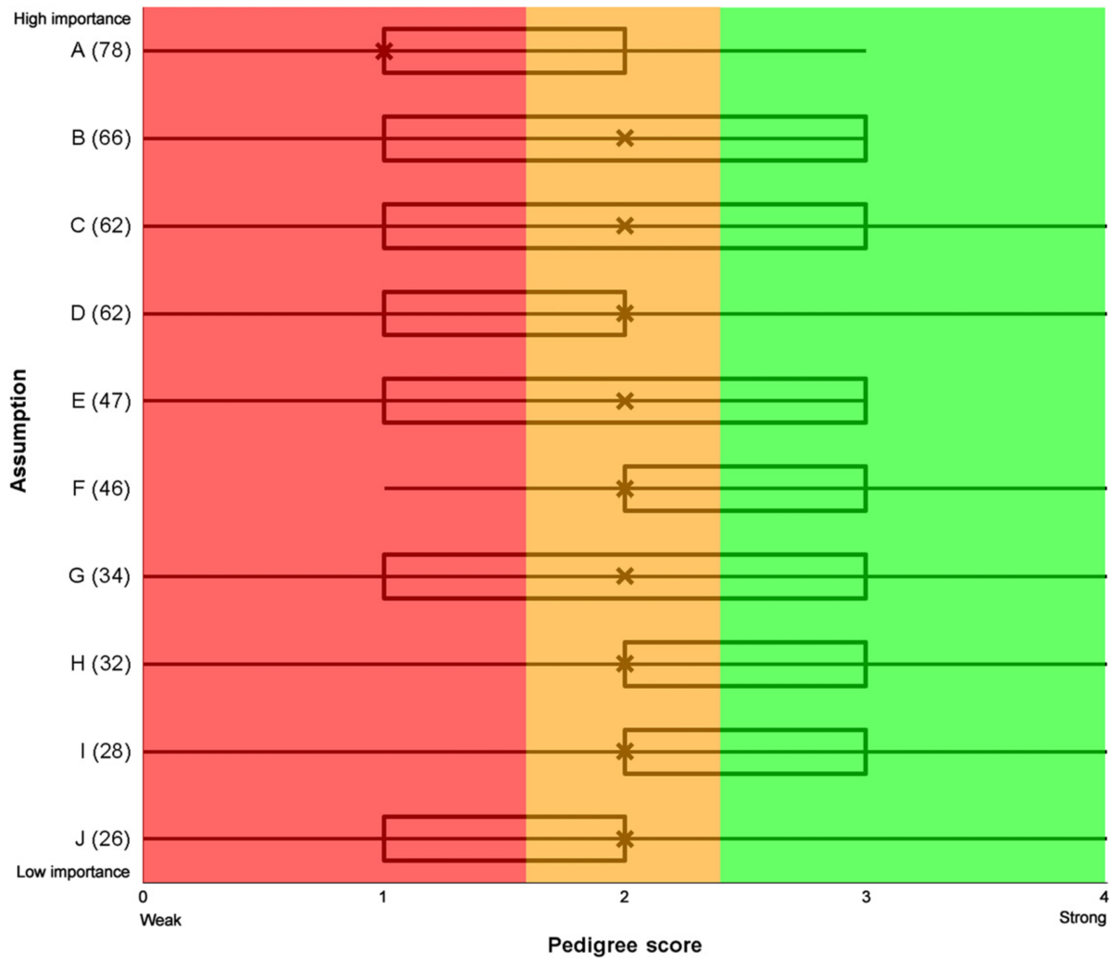

Figure 2. 'Box and Whisker' plots of aggregated pedigree scores of all respondents over all criteria. The *x* indicates the median score, the box represents the interquartile range (25th to 75th percentile) and the line spans the full range from lowest to highest score. On the vertical axis between brackets the sum of points given by respondents during the prioritization of assumptions. The higher the number, the higher the assumption figured on average in the respondents top-10 lists (theoretical maximum: 16 experts\x 10 points = 160 points). Legend for assumptions; A: land use constant; B: drought stress within one year does not impact nature; C: feedbacks via market effects excluded; D: Model linking AMIGO-SMART2/SUMO2 around root zone; E: completeness of range of climate scenarios; F: linking vegetation and hydrology; G: feedbacks via pests, weeds and plant diseases; H: developments agricultural technology omitted; I: conductivity of sub surface too homogeneous in the model; J: aggregation of daily values Amigo-hydrology to annual number for mineralization reduction in SMART2/SUMO2. Colour legend: red = weak pedigree, amber = medium pedigree, green = strong pedigree. Note that both axis use ordinal scales.

that future studies include more relevant land use and land management scenarios which can also be presented as policy options. In doing so, one should be aware that within a given scenario for land use and land management there still can be autonomous adaptation resulting from a feedback of climate change on the behaviour of farmers and nature managers. In theory this could be incorporated in the model chain using so called CAS models or Agent Based Modelling techniques, but such models are very difficult to calibrate and validate for a real life case.

Two other assumptions show a major part of the interquartile range of the pedigree scores in the lower half of the scale (figure 2), indicating major weaknesses in their tenability. These are both related to the linking of the AMIGO soil hydrology model to the root zone of the SMART2/SUMO2 vegetation and soil chemistry model. It is the assumption on the model linking of AMIGO and SMART2/SUMO2 around the root zone (quality and quantity of ground water flux to the root zone) and the aggregation algorithm of daily AMIGO hydrology into yearly values of mineralization reduction feeding into SMART2/SUMO2. By giving priority in the further model development to resolving the weaknesses resulting from these model assumptions, the highest gain in terms of overall model quality is expected.

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, promotes reflectivity and learning in the modelling community, and leads to well informed recommendations for model improvement.

Acknowledgments

We thank all the respondents for their contributions to this study: Alexander Bakker, Janette Bessembinder, Patrick Bogaart, Jerry van Dijk, Remco van Ek, Marta Faneca, Bart van den Hurk, Gijs Janssen, Peter Janssen, Joop Kroes, Janet Mol, Gualbert Oudeessink, Pytrik Reidsma, Peter Schipper, Robert Smit, Iwan Supit, Ab Veldhuizen, George van Voorn, Wieger Wamelink, Joost Wolf. We thank two anonymous reviewers for constructive comments. This study has partly been made possible by the Netherlands National Research Program Knowledge for Climate (http://knowledgeforclimate.climateresearchnetherlands.nl), co-funded by the Netherlands Ministry of Infrastructure and Environment.

References

Bakker A and Bessembinder J J E 2012a *Integraal Waterbeheer—Kritische Zone and Onzekerheden. NMDC Innovatieproject, Deelresultaat Referentie Meteo* (De Bilt: KNMI)

Bakker A M R and Bessembinder J J E 2012b *Time Series Transformation Tool: Description of the Program to Generate Time Series for the Future for the KNMI'06 Climate Scenarios* (De Bilt: KNMI)

Beck M B, Ravetz J R, Mulkey L A and Barnwell T O 1997 On the problem of model validation for predictive exposure assessments *Stoch. Hydrol. Hydraul.* **11** 229–54

Bessembinder J J E 2012 *Uncertainties in Climate Variables: Sources and Relevance* (De Bilt: KNMI)

Bessembinder J J E, Bakker A M R, Leander R and Feijt A 2011 Time series transformation tool: development and use in CCsP projects *National Research Programme Climate Changes Spatial Planning (Amsterdam)*

Boone I, Van der Stede Y, Dewulf J, Messens W, Aerts M, Daubee G and Mintiensa K 2010 NUSAP: a method to evaluate the quality of assumptions in quantitative microbial risk assessment *J. Risk Res.* **13** 337–52

Bouwknegt M, Knol A B, Van der Sluijs J P and Evers E G 2014 Uncertainty of population risk estimates for pathogens based on QMRA or epidemiology: a case study of Campylobacter in the Netherlands *Risk Anal.* **34** 847–64

Clark W C and Majone G 1985 The critical appraisal of scientific inquiries with policy implications *Sci. Technol. Hum. Values* **10** 6–19

Craye M, Laes E and Van der Sluijs J P 2009 Re-negotiating the role of external cost calculations in the Belgian nuclear and sustainable energy debate *Science for Policy* ed A Pereira Guimaraes and S Funtowicz (Oxford: Oxford University Press) pp 272–90

Dankel D J, Aps R, Padda G, Rockmann C, Van der Sluijs J P, Wilson D C and Degnbol P 2012 Advice under uncertainty in the marine system *ICES J. Mar. Sci.* **69** 3–7

De Jong A, Wardekker J A and Van der Sluijs J P 2012 Assumptions in quantitative analyses of health risks of overhead power lines *Environ. Sci. Policy* **16** 114–21

Dessai S and Van der Sluijs J P 2011 Modelling climate change impacts for adaptation assessments ed S Senn, P Dawid, M Christie and K A Cliffen 2011 *Simplicity, Complexity and Modelling* (Chichester: Wiley) pp 83–102

Ellenberg H, Weber H E, Dull R, Wirth V, Werner W and Paulisen D 1991 Zeigerwerte von pflanzen in mitteleuropa [indicator values of plants in central Europe] *Scr. Geobot.* **18** 1–248

Funtowicz S O and Ravetz J R 1990 *Uncertainty and Quality in Science for Policy* (Dordrecht: Kluwer)

Geertsema W, Runhaar H, Spek T, Steingröver E and Witte J P M 2011 Klimaatadaptatie droge rurale zandgronden—Gelderland *National Research Programme Knowledge for Climate (Utrecht)*

Gommes R 1998 *FAO-WMO Roving Seminar on Crop-Yield Weather Modelling. Lecture Notes and Exercises* (Rome: FAO)

Groot A, Van Baaren E, Werners S and van Ek R 2012 *NMDC Kritische Zone En Onzekerheden, Werkpakket 1 Stakeholderanalyse* (Wageningen, UR: Alterra)

Harbaugh A W, Banta E R, Hill M C and McDonald M G 2000 MODFLOW 2000, The US geological survey modular ground water model user guide to modularization concepts and the ground water flow process (Reston, VA: US Geological Survey)

Kloprogge P, Van der Sluijs J P and Petersen A C 2005 A method for the analysis of assumptions in assessments, exploring the value-ladenness of two indicators in the fifth dutch environmental outlook *Netherlands Environmental Assessment Agency* (Utrecht: Bilthoven and Copernicus Institute, Utrecht University)

Kloprogge P, Van der Sluijs J P and Petersen A C 2011 A method for the analysis of assumptions in model-based environmental assessments *Environ. Modelling Softw.* **26** 289–301

Laes E, Meskens G and Van der Sluijs J P 2011 On the contribution of external cost calculations to energy system governance: The case of a potential large-scale nuclear accident *Energy Policy* **39** 5664–73

Maxim L and Van der Sluijs J P 2011 Quality in environmental science for policy: assessing uncertainty as a component of policy analysis *Environ. Sci. Policy* **14** 482–92

Risbey J S, Van der Sluijs J P, Kloprogge P, Ravetz J R, Funtowicz S O and Corral Q S 2005 Application of a checklist for quality assistance in environmental modelling to an energy model *Environ. Model. Assessment* **10** 63–79

Saltelli A, Guimarães Pereira Â, Van der Sluijs J P and Funtowicz S O 2013 What do I make of your latinorum? sensitivity auditing of mathematical modelling *Int. J. Foresight Innov. Policy* **9** 213–34

Schipper P *et al* 2011 *Integraal Waterbeheer—van Kritische Zone Tot Kritische Onzekerheden NMDC-Innovatieproject* (Wageningen, UR: Alterra)

Schipper P N M *et al* 2013 *Integraal Waterbeheer—Kritische Zone En Onzekerheden—Integraal Hoofdrapport* (Wageningen, UR: Alterra)

Sluiter R 2011 Dataset Release Notes, 18-10-2011. KNMI, De Bilt. (http://opendap.knmi.nl/knmi/datadocs/wdscenarios/Release_Notes_interpolscenarios_WD_V0002_20110421.pdf)

Van den Hurk B J J M *et al* 2006 *KNMI Climate Change Scenarios 2006 for the Netherlands* (De Bilt: KNMI)

Van der Linden W, Berendrecht W, Veldhuizen A, Massop H, Blonk A, Heuven A and Zaadnoordijk W J 2008 *Amigo Actueel Model Instrument Gelderland Oost* (Delft: Deltares/TNO)

Van der Sluijs J P, Craye M, Funtowicz S O, Kloprogge P, Ravetz J R and Risbey J S 2005 Combining quantitative and qualitative measures of uncertainty in model based environmental assessment: the NUSAP system *Risk Anal.* **25** 481–92

Van der Sluijs J P, Petersen A C, Janssen P H M, Risbey J S and Ravetz J R 2008 Exploring the quality of evidence for complex and contested policy decisions *Environ. Res. Lett.* **3** 024008

Van der Sluijs J P, Wardekker J A and Janssen P H M 2012 Integraal waterbeheer—kritische zone and onzekerheden nmdc-innovatieproject, deelresultaat kwalitatieve onzekerheidsanalyse baakse beek *PBL Netherlands Environmental Assessment Agency* (Utrecht: Bilthoven and Utrecht University)

Van Walsum P E V and Supit I 2012 Influence of ecohydrologic feedbacks from simulated crop growth on integrated regional hydrologic simulations under climate scenarios *Hydrol. Earth Syst. Sci.* **16** 1577–93

Van Ek R *et al* 2012 *NMDC-Innovatieproject van Kritische Zone Tot Kritische Onzekerheden: Case Studie Baakse Beek* (Delft: Deltares)

Van Walsum P E V and Groenendijk P 2008 Quasi steady-state simulation of the unsaturated zone in groundwater modeling of lowland regions *Vadose Zone J.* **7** 769–81

Vernes R W and Van Doorn Th H M 2005 *Van Gidslaag naar Hydrogeologische Eenheid. Toelichting op de Totstandkoming van de Dataset REGIS II* (Utrecht: TNO)

Wamelink G W W, Akkermans L M W, Brus D J, Heuvelink G B M, Mol-Dijkstra J P and Schouwenberg E P A G 2011 Uncertainty analyses of SMART2-SUMO-P2E-MOVE4 *The Nature Planner Soil and Vegetation Model Chain* (Wageningen, UR: Alterra)

Witte J-P M, Bartholomeus R P, Van Bodegom P M, Cirkel D G, Van Ek R, Fujita Y, Janssen G M C M, Spek T J and Runhaar H 2014 A probabilistic eco-hydrological model to predict the effects of climate change on natural vegetation at a regional scale *Landscape Ecol.* at press