A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components

Esther Meenken[a, b], Chris Triggs[a], Hamish Brown[b] and Edmar Teixeira[b]

[a] University of Auckland
[b] The New Zealand Institute for Plant & Food Research Limited
Email: esther.meenken@plantandfood.co.nz

Abstract: We propose a framework for Uncertainty Evaluation, UE, to explore how the components of a crop model contribute to overall model output uncertainty. We develop a state-space structure to classify the types of uncertainty introduced by each model component.   The state-space structure, $r = f(\mathbf{Z}_t, \boldsymbol{\theta}, \mathbf{E}_t, \mathbf{C}_t, \varepsilon)$, predicts the real world *r* as a function *f* of state variables $\mathbf{Z}_t$, input parameters $\boldsymbol{\theta}$, calibration and environmental data $\mathbf{C}_t$ and $\mathbf{E}_t$ and noise ε.

Bio-physical agricultural models, colloquially crop models, are simplified mathematical representations of physiological and physical processes. Such deterministic models are made up of many components that together simulate real-world agricultural systems. Although the models are deterministic, there are many possible sources of uncertainty; incorrect specification of the governing equations, incorrect input parameters, bias, scaling, aggregation, and inherent stochasticity in data.

Our UE framework has seven steps to provide guidelines for the identification and utilisation of the uncertainty inherent in different aspects of crop models. We illustrate the UE framework using the SIRIUS model as a case study for simulating spring-wheat development. In this paper we combine the three aspects; different types of uncertainty, use of state space structure to describe a time-step model; and a framework for UE. This allows us to explicitly describe and allocate each type of uncertainty within a state-space structure and then curate available information *prior* to diagnosing principal sources of uncertainty and setting or adjusting analysis objectives in the light of actions. It consists of seven steps that sequentially: 1. Describe, validate and verify the model, 2. Clearly identify and compartmentalise model components, 3. Curate available information, 4. Identify principal sources on uncertainty in model components, 5. State the objectives of the evaluation, 6. Generate simulation data and 7. Analyse the simulation data. Our framework can link qualitative UE to quantitative analysis by classifying some classical and modern techniques for generating and analysing data from crop models. The UE framework is illustrated via a case study for simulating spring-wheat development with the SIRIUS model.

Keywords: *Uncertainty, computer simulation model, state-space model*

Meenken et al., A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components.

1.   INTRODUCTION

Bio-physical agricultural models, colloquially crop models, are used to inform and support farm-level decision making, agronomic research, breeding strategies and government policy (Rosenzweig et al. 2013). Such models are simplified deterministic representations of physiological and physical processes, and have many components that work together to simulate real-world agricultural systems. There is been increasing recognition that the impact of uncertainty on crop model results needs to be considered (Rotter et al. 2011; Wallach et al. 2014).   The development of methods to evaluate uncertainty in deterministic models is an active area of research and many approaches have been proposed, including multi-model ensembles, sensitivity analysis, and emulators (Asseng et al. 2013; Teixeira et al. 2015).   There are many excellent resources that discuss and describe the different types of uncertainty (Saltelli et al. 2000; Kurowicka and Cooke 2006; O'Hagan 2006; Wallach et al. 2014). Others describe the conceptualization of time-step models in a state-space structure (Gordon et al. 1993; Cressie and Wikle 2011). Others again put forward plans and frameworks for uncertainty evaluation (UE) (Refsgaard et al. 2006; Refsgaard et al. 2007; Uusitalo et al. 2015). However, none combine the three aspects to explicitly describe and allocate each type of uncertainty within a state-space structure and then curate available information *prior* to diagnosing principal sources of uncertainty and setting or adjusting analysis objectives in the light of actions. We propose a framework to facilitate model UE and we link our qualitative UE to quantitative analysis by classifying some classical and modern techniques for generating and analysing data from crop models.

2.   CASE STUDY: THE WHEAT DEVELOPMENT MODEL SIRIUS.

We will illustrate our methodology using the spring wheat sub-model of the crop model SIRIUS (Jamieson et al. 1998). SIRIUS is a dynamic, deterministic computer simulation model for the development of a plant through time as realised by the number of fully extended leaves. It has a discrete nature such that on each day there is a well-defined set of states by which each state variable may either remain in its current state or update according to environmental cues. SIRIUS simulates the plant's development based on mean daily environmental information. State variables Z_t are used to predict the (observable = LN_t) state variable leaf number (ln_t) on day t based on calculations that simulate the rate of leaf development ($phyllochron_t$), the number of organs ($primordia_t$) and the possible final number of leaves, given daylength and developmental progress ($flnt_t$, fln_t). $t = 1$ is the day of sowing. The simulations depend on input parameters θ that describe cultivar-specific characteristics and responses to environmental signals, and observable environmental information E_t, where TT = mean daily thermal time and PP = daily photoperiod.   He et al. (2012) describe how the model components which integrate the effects of thermal time and photoperiod to express vegetative development are implemented in equations (1 - 5). Model components are described next.

$$ln_t = min\left(fln_{t-1}, \left(\frac{TT_t}{phyllochron_t} + ln_{t-1} \right) \right) \tag{1}$$

$$phyllochron_t = r * bp \tag{2}$$

where, $r = 0.75$ if $ln_t \leq 2$, 1 if $2 < ln_t \leq 8$, 1.3 if $8 \leq ln_t$

and *bp* is a cultivar-specific value for baseline phyllochron (the rate at 2-8 leaves).

$$fln_t = lmin + ps((ppsat - PP_t) * s) \tag{3}$$

where $s = 1$ if $PP_t \leq ppsat$, 0 otherwise

and *ps* and *ppsat* are cultivar-specific values for rate of growth in response to photoperiod, and the photoperiod at which full response occurs.

$$primordia_t = pe + pn * ln_t \tag{4}$$

where *pe* is the number of primordia in the grain at sowing and *pn* is the number of primordia present in the meristem on per leaf basis.

$$fln_t \begin{cases} min(fln_t, primordia_t) & \text{if } primordia_t < (fln_{t-1} + pe) \\ fln_{t-1} & \text{otherwise} \end{cases} \tag{5}$$

Meenken et al., A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components.

3. FOUNDATION: A STATE-SPACE STRUCTURE FOR A CROP MODEL

3.1. Rationalisation of State Space Structure

Crop models often update through time whilst responding to environmental information such as weather or nutrient management input. Most are process based, and update at some discrete time-unit (e.g. hourly, daily, weekly). It is natural to adopt a state-space framework to describe the model structure. In Section 2.2 we follow the notation of authors such as Gordon et al. (1993) and Cressie and Wikle (2011).

3.2. Components of a Crop Model

Input parameters
The input parameter θ represents information that does not change during the sequential updating process., e.g. soil type or cultivar in a crop model.

State equations and State variables
The equations linking the state variables Z_t at time step t jointly define the processes that make up the model given the input parameters θ and the input data.. The state equations describe the underlying scientific processes of the model. They are distinct from the input parameters. The model outputs are then a function f of all the model components.

Observation data
The observed data are denoted $D_t = \begin{pmatrix} C_t & E_t \end{pmatrix}^T$. Observation data may be present at each time point t, only at selected time points, or on some more or less precise scale. C_t represents response or calibration data, possibly for multiple variables and/or scenarios. It could also be available only as a single C, e.g. yield at the end of the simulation process. As defined here C_t, is distinct from the data used to construct the state equations or to validate the model during the model *building* stage. Rather, they either act within the modelling process itself upon the estimated real, unknown target quantity r, for example during data assimilation, or are used as an independent validation data set during model *assessment* (Section 4.1). The construction/validation data used during model building are not considered beyond their indirect contribution to Z_t and the validated model. E_t represents updating environmental or managerial inputs such as rainfall/irrigation or temperature.

3.3. Model Notation

A generic model can be formally represented as the functional relationship (O'Hagan 2006):

$$y = f(x) \tag{6}$$

where x is a vector of inputs and y a vector of outputs. The model structure $f(\bullet)$ specifies how the characteristics of y are determined by those of x. Equations (7 - 10) build on (6) to formally compartmentalise the model components. Strong (2012) identified an extra discrepancy term δ as a linear, additive term to quantify the effect of model error on the model's ability to predict the real, unknown target quantity $r : r = f(x) + \delta$. We extend this to include instead a more general term ξ not assuming either linearity or additivity, to encompass any form of uncertainty that cannot be otherwise allocated. We define:

$$r = f(x, \xi) \tag{7}$$

The inputs of the model x have been defined as either input parameters θ or input environmental data E_t:

$$r = f(\theta, E_t, \xi) \tag{8}$$

The form of the model f used in (6) - (8) includes state equations Z_t and their interactions. Thus we can further refine (8):

$$r = f(Z_t, \theta, E_t, \xi) \tag{9}$$

Meenken et al., A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components.

We further decompose ξ to allow for the uncertainty associated with calibration data C_t separately from the remaining, unidentified uncertainty ε:

$$r = f(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{E}_t, \boldsymbol{C}_t, \varepsilon) \qquad (10)$$

Next we describe sources of uncertainty in models, and allocate them to one of the components of (10).

3.4. Types and sources of uncertainty and their allocation to model components

Although computer simulation models represent detailed scientific understanding of real-world systems they are simplifications of reality and hence imperfect. Imperfections may arise due to either epistemic or aleatory sources of uncertainty as discussed in light of model components next.

Epistemic uncertainty is a potential deficiency in the model that is due to a lack of knowledge. It may also arise from assumptions introduced in the derivation of the mathematical model used or simplifications related to the correlation or dependence between physical processes.

Input parameter uncertainty $\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{E}_t, \boldsymbol{C}_t, \varepsilon)$ is also known as 'state of the world' uncertainty. It refers to uncertainty about the appropriate values to describe the scenario to be modelled.

Data uncertainty $\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{E}_t, \boldsymbol{C}_t, \varepsilon)$, both from environmental input and calibration, may enter because of bias or scaling/aggregation, imperfectly measured observations, inherent stochasticity or, in the case of environmental input data, because of incorrect specification in the model. Scaling/aggregation uncertainty refers to situations in which the model is used on a scale different from that on which it was developed to operate.

Structural uncertainty, $\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{E}_t, \boldsymbol{C}_t, \varepsilon)$, is also known as model inadequacy and refers to discrepancies between the structure of the model relative and the processes it represents. This may be in the nature of the state equations $\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{E}_t, \boldsymbol{C}_t, \varepsilon)$, or the more inclusive $\boldsymbol{f}$ which describes how the model components fit together. The obvious sign of structural uncertainty is the difference between the value of the real-world process, and the model outputs at the true values of the input data and parameters.

Aleatory uncertainty: Aleatory uncertainty $\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{E}_t, \boldsymbol{C}_t, \varepsilon)$ is usually thought of as intrinsic, random variation of a real world process even when the conditions are fully specified. The true process $\boldsymbol{r}$ is then defined as the mean value averaged over this intrinsic variation. It will only be able to be quantified as an aspect of the residual uncertainty in observed calibration data.
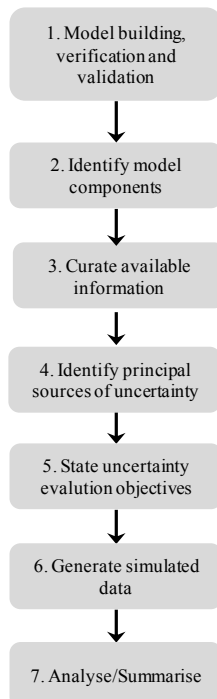
4. A FRAMEWORK FOR UE

4.1. The life of a model

A model's life is a sequence of phases starting with conceptualisation and implementation in code (Model Building), followed by testing (Model Assessment), and use (Model Application). However, adjustments can be made at any phase as understanding of the system changes. Explicit identification of the model phase during UE is necessary when defining the objective of the UE to help to identify the most appropriate techniques. Our UE framework is primarily aimed at enriching the Model Assessment and Application phases.

4.2. Outline of a robust UE

The most appropriate choice of sampling and analysis techniques depends on available resources, and the objectives of the UE. Figure 1 outlines seven steps which can help ensure a robust UE. Figure 1 has some likeness to the conceptual model put forward in Figure 3 of Refsgaard et al. (2006). The UE framework will be expanded upon and illustrated using SIRIUS.

1. Model building, verification and validation
2. Identify model components
3. Curate available information
4. Identify principal sources of uncertainty
5. State uncertainty evaluation objectives
6. Generate simulated data
7. Analyse/Summarise

Figure 1. Seven steps for model UE.

Meenken et al., A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components.

4.3. Carry out the UE

Step 1 - Build, verify and validate the model: As with any modelling exercise, building a computer simulation model begins with observations, from which hypotheses are derived, and then implemented in code. The joint processes of building, verifying and validating models are well documented and are not discussed further.

Step 2 - Identify model components: Each component of a crop model can be formulated to fit into the state-space structure described in Section 2. Table 1 provides a tableau into which each component can be assigned, illustrated for our case study.

Step 3 - Curate available information: Data, expert opinion, the phase of the model's life, components, sources of uncertainty and other relevant information is usually able to be identified about a model. For example, for SIRIUS, information is found in the literature, and data and expert knowledge can be accessed through research organisations in New Zealand. Selected information known about SIRIUS is collected in Table 2. If the model is too large to enable this process easily, a first step should be to identify which sub-component of the model structure is to be evaluated.

Step 4 - Identify principal sources of uncertainty in the model: Given the components identified in Steps 2 and 3, assess which sources of information are likely to be of key importance. These include model structural error, incorrectly specified input parameters, bias in environmental data, and inherent stochasticity. The most important problems can then be identified.

Table 1. Classification of model components filled in for the SIRIUS wheat model as defined above.

| Input Parameter | Observation Variable | | Structural Uncertainty | |
	Calibration Variable	Environmental Variable	Model Form	State Equation
$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$
ps	LN_t	PP_t	*the way in which model components interact at each time step*	$phyllochron_t$
ppsat		TT_t		$primordia_t$
lmin				$flnt_t$
pn				fln_t
pe				ln_t
bp				

Table 2. Information matrix adjusted from (Refsgaard et al. 2007) filled in for the SIRIUS wheat model.

Context	Notation	Data	Expert opinion	Other
Structural uncertainty	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$		The developmental phase between imbibation and emergence may not be correctly specified.	Simulated Data for wheat grown in Southland consistently underestimates time of anthesis
Model Inputs	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$		Data from the literature (Jamieson et al. 2008) the rate of wheat development (*bp*) is widely variable between cultivars. Expert opinion on range values for several cultivars: *bp* 90-110; *lmin* 5-9; *ps* 0.01-0.7	unknown
Scaling, aggregation, sampling or aleatory sources of uncertainty	$r = f(Z_t, \theta, E_t, C_t, \varepsilon)$	Measured day of flag in southland crops for one cultivar. Measured observations of LN at weekly intervals for controlled climate conditions. Measured weather station data in Lincoln, Canterbury, New Zealand from 1960 - present	unknown	unknown

Step 5 - State objectives of evaluation: UE does not occur in a vacuum, and neither can it be generalised. Rather, it is a tailored exploration that will be specific not only to the model but also to the environment/scenario(s) for which it is to be used (Wallach et al. 2014). The key is to state clearly the objective of the analysis, whilst accounting for the information and resources available, as summarised in Tables 1-2. A selection of techniques will probably be helpful and should be combined to provide a heuristic view of the model. The most insightful techniques will vary depending on the phase of the model's life that is under study and on the properties of that specific model. The objectives of the UE could comprise one or more of the following:

1. Assess the size and direction of bias of model-simulated values for *fln* for a new location, potentially to guide new research/calibration efforts $r = f(Z_t, \theta, E_t, C_t \varepsilon)$.

Meenken et al., A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components.

2. Carry out a sensitivity analysis to assess whether the model is also (as the literature suggests the real world is) sensitive to changes in bp $r = f(Z_t, \boldsymbol{\theta}, E_t, C_t, \varepsilon)$.

3. Carry out a sensitivity analysis to assess the impact (in number of days of error in day of flag leaf estimation) of spatial bias in thermal time (TT) input data $r = f(Z_t, \boldsymbol{\theta}, E_t, C_t, \varepsilon)$.

4. Fit a data assimilation model to explore whether the model correctly simulates the development of each leaf through time $r = f(Z_t, \boldsymbol{\theta}, E_t, C_t, \varepsilon)$.

Figure 2. Classification of data generation techniques.

Step 6 - Generate simulation data: about how the model responds to complex combinations of inputs. It is not synonymous with real-world, observed data. Simulation data can be used either to explore $y = f(Z_t, \boldsymbol{\theta}, E_t, C_t, \varepsilon)$ (sensitivity analysis) or $r = f(Z_t, \boldsymbol{\theta}, E_t, C_t, \varepsilon)$ (calibration). The difference lies
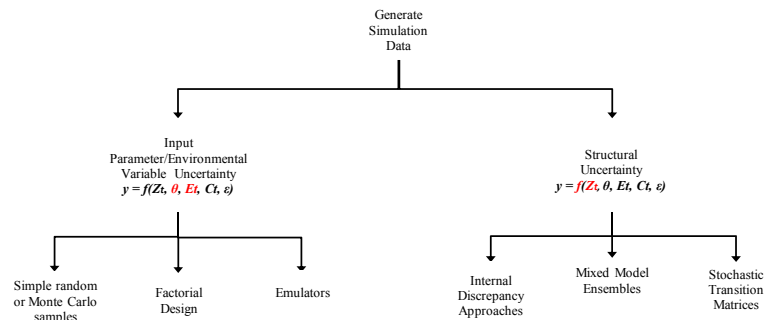
in whether there are real-world data C_t is available to provide information about r and ε. The technique used to generate data will influence the direction of the analysis toward evaluating uncertainty due either to input parameters, environmental inputs, or to structural uncertainty. A selection of sampling techniques to facilitate exploration of structural or input parameter/environmental data uncertainty is shown in Figure 2. Simulated data arising from any of these data generation techniques could be used for most of the analysis techniques shown in Figure 3. For the case study, we focus on objective 2). We will carry out a sensitivity analysis to explore the importance of incorrectly specifying the base phyllochron of a wheat cultivar in the model, in comparison to other input parameter and environmental input variables. The model output was

Table 3. Input Vector.

Term	Levels			
	1	2	3	4
Year	1980	1990	2000	
Day	10	69	160	252
lmin	5	9		
bp	90	130		
ps	0.01	0.07		
Factorial	=3*4*2*2*2		96 observations	

sampled using a complete factorial design over a selection of years and sowing date combinations. Non-linearity in responses may be explored via experimental data and model outputs. This does not allow for probabilistic information to be included. The ranges of values for input weather information and input parameters *ps*, *lmin* and *bp* are recorded with the rest of the information in Table 2. This information is used to develop a simple factorial design to use as an input vector for a set of model simulations shown in Table 3.
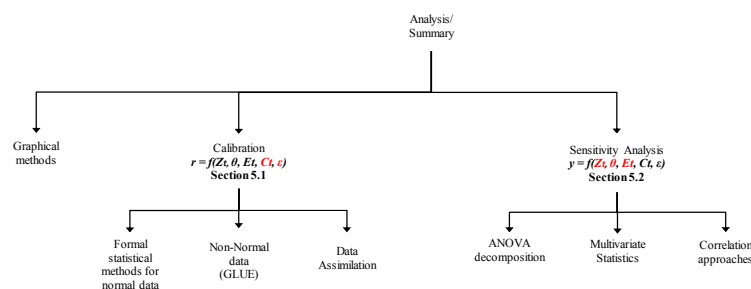

Step 7 - Analyse/Summarise data: Once the simulation data have been generated, analysis and summary of the information can begin. Depending on the objectives defined in step 5, the data will be analysed either to identify areas in need of further research (Model Assessment), or to predict/smooth with confidence

Figure 3. Analysis and summary techniques with and without observed data.

ranges representing the desired sources of information (Model Application). Figure 3 shows a classification analysis techniques based on real-world observation data (calibration) or not (sensitivity analysis). Graphical methods should always be a part of UE. The data resulting from the factorial design implemented in step 6 is trivial to analyse. We use sums of squares from ANOVA decompositions to give the first order sensitivity index for the main factor effects (S_i). The total sensitivity index (T_i) is calculated by adding the interaction component to each main factor (Saltelli et al. 2000) to show the relative contribution of each term to the overall variation in model outputs (Figure 4). After the effects of weather have been accounted for, *bp* accounts for the largest source of variation in model outputs. Improving information about the phyllochron in most varieties is likely to affect the ability of SIRIUS to accurately simulate spring wheat development.

Meenken et al., A framework for uncertainty evaluation of agricultural computer simulation models with a focus on allocation of uncertainty to model components.

5. CONCLUSIONS

Understanding the effect of uncertainty on the outputs of models is highly topical in the crop modelling literature. This paper proposes a framework to help researchers quickly identify an appropriate method for the uncertainty quantification goal at hand. Based on a state-space structure, tools to help to curate information, to diagnose the most important sources of uncertainty, and to identify uncertainty evaluation objectives have been developed. It then provides guidelines for the identification of approaches to explore uncertainty.
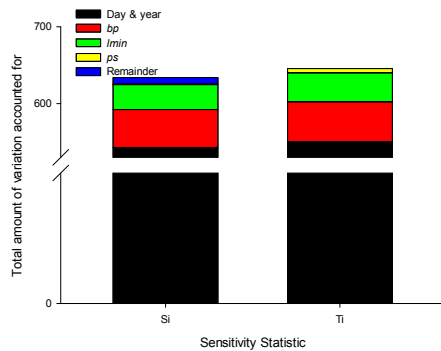

Figure 4. S_i + remainder (interactions) and T_i

REFERENCES

Asseng S., Ewert F., Rosenzweig C., Jones J., Hatfield J., Ruane A., Boote K., Thorburn P., Rötter R., Cammarano D. 2013. Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* 3:827-832

Cressie N., Wikle C.K. 2011. Statistics for Spatio-Temporal Data. Hoboken, New Jersey: John Wiley & Sons.

Gordon N.J., Salmond D.J., Smith A.F.M. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *Radar and Signal Processing, IEE Proceedings F* 140:107-113

He J., Le Gouis J., Stratonovitch P., Allard V., Gaju O., Heumez E., Orford S., Griffiths S., Snape J.W., Foulkes M.J., Semenov M.A., Martre P. 2012. Simulation of environmental and genotypic variations of final leaf number and anthesis date for wheat. *European Journal of Agronomy* 42:22-33. doi:10.1016/j.eja.2011.11.002

Jamieson P., Brooking I., Zyskowski R., Munro C. 2008. The vexatious problem of the variation of the phyllochron in wheat. *Field Crops Research* 108:163-168. doi:10.1016/j.fcr.2008.04.011

Jamieson P.D., Semenov M.A., Brooking I.R., Francis G.S. 1998a. Sirius: a mechanistic model of wheat response to environmental variation. *European Journal of Agronomy* 8:161-179

Kurowicka D., Cooke R. 2006. Uncertainty Analysis. Chichester: John Wiley & Sons.

O'Hagan A. 2006. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety* 91:1290-1300. doi:DOI: 10.1016/j.ress.2005.11.025

Refsgaard J.C., van der Sluijs J.P., Brown J., van der Keur P. 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* 29:1586-1597. doi:http://dx.doi.org/10.1016/j.advwatres.2005.11.013

Refsgaard J.C., van der Sluijs J.P., Højberg A.L., Vanrolleghem P.A. 2007. Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software* 22:1543-1556. doi:http://dx.doi.org/10.1016/j.envsoft.2007.02.004

Rosenzweig C., Jones J., Hatfield J., Ruane A., Boote K., Thorburn P., Antle J., Nelson G., Porter C., Janssen S. 2013. The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies. *Agricultural and Forest Meteorology* 170:166-182

Rotter R.P., Carter T.R., Olesen J.E. 2011. Crop–climate models need an overhaul. *Nature Climate Change* 1:175-177

Saltelli A., Chan K., Scott E.M. (eds). 2000. Sensitivity Analysis. Wiley, New York

Strong M., Oakley J.E., Chilcott J. 2012. Managing structural uncertainty in health economic decision models: a discrepancy approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 61:25-45

Teixeira E.I., Brown H.E., Sharp J., Meenken E.D., Ewert F. 2015. Evaluating methods to simulate crop rotations for climate impact assessments–A case study on the Canterbury plains of New Zealand. *Environmental Modelling & Software*

Uusitalo L., Lehikoinen A., Helle I., Myrberg K. 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software* 63:24-31. doi:http://dx.doi.org/10.1016/j.envsoft.2014.09.017

Wallach D., Makowski D., Jones J.W., Francois B. (eds). 2014. Working with Dynamic Crop Models. Second Edition edn. Elsevier.