Article (refereed) - postprint

Van Oijen, Marcel. 2017. **Bayesian methods for quantifying and reducing uncertainty and error in forest models**. *Current Forestry Reports*, 3 (4). 269-280. 10.1007/s40725-017-0069-9

[This paper is published in Current Forestry Reports 3: 269-280. https://doi.org/10.1007/s40725-017-0069-9]

Bayesian methods for quantifying and reducing uncertainty and error in forest models

Marcel van Oijen

Centre for Ecology & Hydrology, Bush Estate, Penicuik EH26 0QB, UK

mvano@ceh.ac.uk

Abstract

Purpose of review
Forest models are tools for analysis and prediction of productivity and other services. Model outputs can only be useful if possible errors in inputs and model structure are recognized. However, errors cannot be quantified directly, making uncertainty inevitable. In this paper, we aim to clarify terminological confusion around the concepts of error and uncertainty, and review current methods for addressing uncertainty in forest modelling.

Recent findings
Modellers increasingly recognize that all uncertainties – in data, model inputs and model structure – can be represented using probability distributions. This has stimulated the use of Bayesian methods for quantifying and reducing uncertainty and error in models of forests and other vegetation. The Achilles' heel of Bayesian methods has always been their computational demand, but solutions are being found.

Summary
We conclude that future work will likely include: (1) more use of Bayesian methods, (2) more use of hierarchical modelling, (3) replacement of model spin-up by Bayesian calibration, (4) more use of ensemble modelling and Bayesian model averaging, (5) new ways to account for model structural error in calibration, (6) better software for Bayesian calibration of complex models, (7) faster MCMC algorithms, (8) more use of model emulators, (9) novel uncertainty visualization techniques, (10) more use of graphical modelling, (11) more use of risk analysis.

Keywords

Calibration, Discrepancy, Ensemble modelling, Hierarchical modelling, Risk analysis, Visualization

Introduction

Forests provide many services to mankind, and models can be used to predict how well they will do so in different circumstances. This ranges from prediction of productivity at one specific site, to prediction of many different forest services worldwide. Model studies are used to optimise forest management [1] and to assess the possible impacts of environmental change [2,3•]. Model structure varies widely too, from simple look-up tables to complex dynamic models [4]. Despite these differences among forest models and their uses, there are methodological problems they all share. In this paper, we focus on errors in models and data, and how they contribute to predictive uncertainty.

Every model can be viewed as a function that converts inputs into outputs. There are two types of input: *drivers* and *parameters*. Drivers are environmental variables (weather, disturbances, management) whereas parameters are constant properties of trees and soil. Errors can be present in all three model components: drivers, parameters, and model structure itself. These errors affect model output, and when we are uncertain about the magnitude of the errors, there will be uncertainty about the quality of the model prediction. Unfortunately, not many modelling studies include a comprehensive analysis of errors and uncertainties, possibly because methods for doing so have not been standardized.

In recent years, there has been a trend toward basing scientific research under conditions of incomplete information (i.e. most of science) on probability theory [5–9]. This approach is taken here as well. We aim to show that defining all uncertainties in forest modelling as probability distributions allows rigorous reduction of those uncertainties when new data become available. The approach that we are presenting is known in the literature by many different names, including Bayesian calibration, data assimilation, model-data fusion and inverse modelling. Whilst the different names refer to different applications of modelling, they all share the idea of specifying probability distributions which are modified according to the rules of probability theory (in particular: Bayes' Theorem) when new data become available. It is this idea that facilitates the comprehensive analysis of errors and uncertainties.

All methodology presented in this paper can be applied to any kind of forest model: simple or complex, site-based or area-based, deterministic or stochastic. However, for reasons of consistency we shall take most of our examples from the literature on deterministic process-based modelling of forest biogeochemistry. This is a class of dynamic models that simulates tree growth as part of the carbon-, water- and nutrient-cycles in the soil-tree-atmosphere system [3•]. Such *process-based models* (*PBMs*) can be used for assessment of multiple ecosystem services delivered by forests [1]. PBMs tend to be parameter-rich and computationally demanding. We shall explain how these characteristics hamper the comprehensive application of the probabilistic approach to uncertainties, and which solutions have been proposed.

The lay-out of the paper is as follows. We begin by defining some key terms that will be used in the paper, including 'error' and 'uncertainty', and the symbols that we use in equations. After that, we explain the Bayesian approach to model parameter estimation and the most common computational problems and solutions. Next, we discuss hierarchical modelling which frees the modeller from the need to define as constants any parameters that, in reality, are variable. We then move on to errors and uncertainty in model structure,

showing different ways of tackling the issue. We then discuss how model uncertainties should be reported, and finish with a general discussion and outlook. Throughout, we keep technical detail to a minimum but point to references that provide examples and further reading, where possible taken from the literature on process-based forest modelling.

Terminology and symbols

In this section, we introduce terminology and symbols that will be used throughout the paper. We distinguish between three terms that are often confused but have different meanings: *error*, *uncertainty* and *variability*. Say we want to measure a forest property such as basal area. Let the true value be z, and our measured value be y. Since every measurement has some error, ϵ_y, we can write:

$$z = y + \epsilon_y.$$

Error is thus defined as the difference between an estimate and the true value. Instead of using measurement, we could also have estimated z by means of a model. We then would write:

$$z = f(x, \theta) + \epsilon_{model},$$

where f is the model, x are the environmental conditions (drivers) for which we run the model, and θ are the model's parameter values. The final term ϵ_{model} represents model error which arises because f, x and θ will all be wrong to some extent. All the terms in the above equations can be scalars (single values) but more commonly they are vectors (multi-dimensional).

Uncertainty is having incomplete knowledge about a quantity. Uncertainty can always be represented by a probability distribution, denoted as $p[\,.\,]$. So the probability distribution $p[\epsilon_y]$ defines which values of measurement error we deem possible and their relative probabilities of occurring. Say our instrument is accurate (unbiased) but has only 50% precision. Then we are highly uncertain about the error and might write $p[\epsilon_y] = N[\epsilon_y; \mu = 0, \sigma = y/2]$, which stands for a normal distribution with mean zero and standard deviation equal to half the observed value. Conditional probability distributions, e.g. for θ given a certain value of y, are denoted in the standard way as $p[\theta|y]$.

Note that we use square brackets for probabilities, probability distributions and likelihoods (e.g. p[..], N[..]) and parentheses for functions such as f(..).

Variability is the presence of differences in a set of values. Variability can lead to sampling error of uncertain magnitude. Say we randomly sample 100 trees from a large forest, measure stem diameter on each, and collect the 100 values in vector y. The standard deviation σ_y is then a measure of the variability in the sample. If all of the measurements are very precise then there will be little uncertainty about that variability. But if we use the sample mean as an estimate for the mean diameter of all trees in the forest, then there will be an unknown sampling error. In this example, basic sampling theory states that the uncertainty about the sampling error may be represented well by a normal distribution with mean zero and standard deviation equal to $\sigma_y/\sqrt{100}$.

Errors and uncertainties in process-based forest modelling

In this section, we give an overview of errors that arise in forest modelling - in drivers, parameters and model structure. We discuss how to quantify uncertainty, i.e. what kind of probability distributions we may want to assign to the errors, and how the uncertainties about errors propagate to uncertain model predictions. However, we leave formal theory to later sections.

Errors and uncertainties in forest model drivers

As described above, drivers are the environmental conditions for which we run our forest model (the x in our equations). Methodologically, drivers are defined as boundary conditions that we do not attempt to calibrate, but that we accept from an external source of information. A typical example would be time series of temperature and precipitation received from a weather station in or near the forest, or predicted by a climate model.

Although drivers are not calibrated, we are interested in any possible errors they may contain, and how uncertain we are about their magnitude. Driver errors come in three types: *gaps*, *measurement error* and *non-representativeness*, which each need to be treated in a different way.

Gaps in driver data are common. When a model requires daily driver data and the weather time series misses some day or days, the gaps need to be filled. Gap-filling methods range from simple linear interpolation to stochastic process modelling, akin to kriging. The latter is the only approach for which uncertainties are readily quantified (as the kriging or Gaussian process variance).

Measurement error is caused by limited precision and accuracy of the measurement instrument, and sometimes also by errors in transferring measured values to data files. This is the simplest kind of error; uncertainty about it is generally well represented as Gaussian noise.

The most complicated errors arise when the driver data are not fully representative of the forest conditions for which we plan to run our model. The weather data may have come from an off-site weather station, or the data may have a different spatiotemporal scale than the model requires. A very common example is using output from gridded climate models as input for forest PBMs: the climatic data will then be grid cell averages, thus ignoring spatial heterogeneity. Such data are not representative for any specific forest within the cell. If the forest model is nonlinear, as most models are, then input averaging leads to errors in model outputs, but methods to estimate that error exist, based on Taylor-expansion of the driver-dependency [10•]. Another option is to downscale the data, although that brings its own errors and uncertainty. No general rule for assigning probability distributions to errors from non-representativeness exist: each case must be examined on its own.

Errors and uncertainties in forest model parameters

Parameters are constants that represent properties of trees and soil. In PBMs, every process that is modelled requires a parameter to define its basic rate. But process rates also depend on internal and external conditions (e.g. tree nitrogen content, air temperature), and each controlling mechanism adds at least one parameter to the model to quantify the strength of

the dependence. The number of parameters in forest PBMs thus tends to range from tens to many hundreds [11–13].

Parameters can be set to incorrect values, making model behaviour unrealistic. Except for some physical constants, a precise definition of parameter error is hard to give. What is the 'true' value of a forest model parameter? Because a model is a simplification of reality, every parameter plays a somewhat different role in the model than its namesake in a true forest. We cannot, therefore, simply go out and measure the true value exactly. This is apart from the fact that measurements have errors too and that no tree or soil property is truly constant in reality. In practice, we say that the correct value of a parameter is that value that makes the model behave most realistically. Parameter error is the difference from that value, and parameter uncertainty is not knowing what that value is.

Although measurements cannot provide true parameter values, they can give first rough estimates. For about thirty different parameters common to three forest PBMs, Levy et al. [14] reviewed measurement values reported in the literature. For each parameter, a wide range of values was reported that could be interpreted as probability distributions representing parameter uncertainty. For many parameters, the distributions were skewed and better represented by beta distributions than by normal distributions [15].

Errors and uncertainties in forest model structure

Every model is a simplification of reality and therefore, to some extent, wrong. But the behaviour of models - their repertoire of outputs for different conditions - can be compared, showing that some models are more realistic than others. Forest model comparisons abound in the literature and tend to show large differences between their predictions [11,12]. Some tentative general conclusions may be drawn: the feedback structure of models is more important than the mathematical form of individual equations [16] and a consistent level of process detail in different parts of the model is desirable [17,18].

Increasingly, forest model comparisons involve assessing simulations against observations, not just of output variables such as productivity, but also of the underlying processes and mechanisms as represented in the models [2,19–21]. Medlyn et al. [22•] refer to this as an "assumption-centred model intercomparison approach". However, these model comparisons do not employ probability theory and therefore cannot quantify the degree of uncertainty about model structural error. Moreover, any advice about model structure based on observations is contingent on the range of environmental conditions for which models were tested and remains heuristic: there is no unique way to derive forest PBMs from first principles, so errors remain inevitable.

This leaves us with only two ways to account, probabilistically, for structural error in a modelling study: extend the PBM with a stochastic error term [23], or use a large ensemble of different models and proceed as if one model in the set should be correct [24]. In the first approach, uncertainty is quantified by assigning a probability distribution to the structural error term. In the second approach, uncertainty is represented by a probability distribution over the set of models, with highest probabilities assigned to models that are considered to be most plausible. Both require taking into account that model performance depends not only on its structure but also on the parameter settings with their own uncertainties. Technical details of both methods will be given in a later section.

Forward propagation of uncertainty to model outputs

Because measurements and models have unknown errors, all we can quantify are uncertainties. Therefore, the common term of 'error propagation' denoting how much the error in inputs contributes to error in outputs is a misnomer. What is propagated is uncertainty, not error. Outputs may have minimal error despite large errors in inputs, if the errors happen to have compensating effects. This is a common occurrence when models are tuned to produce a desired behaviour.

The techniques, rather than the name, associated with 'error propagation' may well be used to quantify output uncertainty, provided the model is simple enough that partial derivatives of output with respect to inputs can be analytically calculated. However, PBMs tend to be too complex for such approaches, so uncertainty is mostly quantified by Monte Carlo methods: sampling from the probability distributions of model inputs (and structures if we have an ensemble of models) to generate a representative set of possible model outputs. In the study by Levy et al. [14], mentioned above, Monte Carlo sampling was used to quantify the contribution of parameter and model structural uncertainty to uncertainty about the carbon sink of a coniferous forest in southern Sweden. They concluded that the sink uncertainty was 92% due to parameter uncertainty and only 8% due to structural uncertainty . The key parameter uncertainties were for allocation of carbon to leaves, stems and roots. These results were, of course, contingent on the three chosen forest models and the single application site. Reyer et al. [25] also demonstrated the importance of parameter uncertainty - in their case for prediction of future forest productivity - but in a comparison with uncertainty about climatic drivers rather than model structure. Sutton et al. [26] showed the importance of uncertainties about model drivers (in particular atmospheric nitrogen deposition) and model structure for predictions of forest productivity across Europe. Minunno et al. [27] found that uncertainty about soil conditions (water availability and fertility) mainly determined the predictive uncertainty of a growth model for *Eucalyptus globulus* in Portugal. In all these studies, the reported predictive uncertainties depended not only on uncertainty in inputs and model structure, but also on the conditions for which the models were run. In less water-limited conditions than Portugal, for example, outputs of the model that was used in [27] might have been less affected by uncertainties in water availability.

Reducing parameter uncertainties: The Bayesian approach

This section introduces the Bayesian approach to reducing uncertainty. Uncertainty reduction is slightly more complicated than uncertainty propagation, which was discussed in the previous section. However, the theory is old and well-established, dating back to the 18th century [28], and we shall only give the essential details here. Longer expositions can be found in textbooks [29,30•] and tutorial papers [5,31].

We initially focus on the model's parameter values. As noted above, we always have some uncertainty about the proper values of a model's parameters, and we can express this uncertainty as the probability distribution $p[\theta]$. When new data y arrive, we may want to use those data to reduce our uncertainty about θ. Bayes' Theorem tells us how to change a

prior parameter distribution, $p[\theta]$, into a *posterior* distribution for "θ given y" denoted as $p[\theta|y]$:

$$p[\theta|y] \propto p[\theta]p[y|\theta] \qquad (\textit{Bayes' Theorem}).$$

So according to Bayes' Theorem, the posterior distribution is proportional to the prior and to $p[y|\theta]$. This last term is called the *likelihood function* of the parameters, as it expresses how likely the data are for any value of the parameters, and is often written as $L[\theta]$. Parameter estimation using the product of prior and likelihood is called 'Bayesian calibration' [23,31]. At this point, we briefly remind the reader that both the parameter vector θ and the data vector y can be (highly) multi-dimensional. In other words, the posterior distribution for the parameters, $p[\theta|y]$, is a *joint* probability distribution, and it is likely that most parameter-pairs in θ will show some degree of correlation.

A simple example of Bayesian calibration

Let us assume, for now, that model error ϵ_{model} is much smaller than measurement error and can be ignored. We can then combine the first two equations of this paper to derive the likelihood function:

$$p[y|\theta] = p[f(x,\theta) - y = \epsilon_y].$$

If we would further assume that ϵ_y has a zero-mean normal distribution and a coefficient of variation of 50% (we used the same example before), then the likelihood function would simplify to $p[y|\theta] = N[f(x,\theta) - y; \mu = 0, \sigma = y/2]$. Say that we have also defined the prior distribution for the parameters, $p[\theta]$, based on literature review and expert opinion. Then the final step, as prescribed by Bayes' Theorem, is to find the product of the prior and the likelihood. In principle, a representative sample from the posterior distribution can easily be generated by Monte Carlo sampling: take a large sample from the prior and use the likelihoods as weights in deciding which parameter values to keep. More sophisticated methods such as MCMC will be explained in the next section. This simple example captures the three essential steps in Bayesian calibration: (1) specify a prior, (2) specify $p[\epsilon_y]$ and from it the likelihood function, (3) apply Bayes' Theorem.

Likelihood

In practice, the hardest step in Bayesian calibration is formulating an appropriate likelihood function, $p[y|\theta]$. The likelihood function is a conditional probability density function. It is the answer to the question: what is the probability of measuring y if the true value is $f(x,\theta) + \epsilon_{model}$. The likelihood is a powerful concept: y is usually multi-dimensional and may consist of measurements over time on several different soil and tree variables, yet $p[y|\theta]$ can always be defined. It is through the likelihood function that Bayesian calibration has the capacity to use highly heterogeneous data sets in parameter estimation. Levy et al. [32] used Bayesian calibration to reconcile eddy-covariance measurements of N_2O emissions with chamber measurements of the same fluxes. Patenaude et al. [33] combined remote sensing data from satellites with field-based data on Corsican pine stands in the UK, in Bayesian calibration for the parameters of forest PBM 3-PG. Höglind et al. [34] combined measurements on 10 different variables from 5 grassland sites in their likelihood function. With such rich data

sets, it becomes important to assess whether measurement errors for different variables have correlations that should be expressed in the likelihood function.

Formulating the likelihood function can be difficult, even when model error can be ignored and different variables are measured independently, because measurements can be wrong in three different ways, as summarized in the next equation:

$$\epsilon_y = \epsilon_{y,stochastic} + \epsilon_{y,systematic} + \epsilon_{y,representativeness}.$$

The first of the three data error terms, $\epsilon_{y,stochastic}$, quantifies random measurement noise, independent for each data point in y, which we could represent with a zero-mean normal distribution [8]. This is often the only data error recognized by modellers, but it is unlikely to be the most important one. The second error term, $\epsilon_{y,systematic}$ represents measurement bias which could shift whole time series of data up- or downward. The final term, $\epsilon_{y,representativeness}$, is generally the hardest to quantify. It refers to the possibility of our data being derived from other conditions than our model is designed for. If the observed forest has a hidden growth limitation, say phosphorus deficiency, that is not expressed in the model, then the data - from the point of view of model parameter estimation - will be underestimating growth. In contrast to the other two types of measurement error, the scope for $\epsilon_{y,representativeness}$ cannot be reduced by more careful measurement or greater sampling intensity.

For every data set, the modeller and data expert need to distinguish the three types of possible data error, and assign probability distributions to the respective error uncertainties. The parameters of these probability distributions, such as the standard deviation of the stochastic noise, can be fully specified a priori (e.g. $\sigma_{y,stochastic} = y/2$), or they can be estimated in the calibration together with the regular tree and soil parameters. The latter method would be preferable in this example if we have little information about the precision of our measurement instrument. Van Oijen et al. [9] calibrated the degree of systematic error in chamber-measurements of soil-fluxes of CO_2, N_2O and NO using four different forest PBMs - and all four calibrations suggested that the measured CO_2 emission values had been unrealistically high.

Computational solutions: MCMC and emulation

As mentioned above, Bayesian calibration tends to be implemented through sampling algorithms, thus increasing the computational demand of forest modelling. In this section, we present two methods for reducing computational demand: MCMC and model emulation.

MCMC

The Bayesian approach to parameter estimation requires modellers to make a major mental shift: we no longer aim to find a single 'best' parameter vector - instead we aim to determine the posterior probability distribution for the parameters. Only the full distribution adequately represents our state of knowledge. Although this shift has made rigorous uncertainty quantification possible, it has also created computational problems: when models have many parameters, the distribution will be high-dimensional and difficult to

sample from. A solution for this problem was provided by Metropolis et al. [35] who introduced the so-called Markov Chain Monte Carlo (MCMC) method.

MCMC consists of a walk through parameter space in such a way that the visited points together form a representative sample from the posterior distribution. The method requires that at each step we decide to accept or reject a proposed new point (= vector of parameter values) depending on the prior probability times the likelihood for that parameter vector. In this way, the chain gradually moves to the region of highest posterior probability in parameter space. Many different variants of MCMC now exist following that general plan. A simple introduction to the method, with an example for a forest model, was provided by Van Oijen et al. [31] and some frequently asked questions about MCMC are answered by Van Oijen [36].

Free software exists that makes it easy to encode priors and likelihood functions and run the MCMC. These include, in historical order, the packages BUGS (and more recently WinBUGS and OpenBUGS), JAGS and Stan. These packages are mainly suited for simple stochastic models and not for complex PBMs. However, in recent years, two impressive pieces of software have been developed that facilitate the Bayesian calibration of PBMs using MCMC: BayesianTools [37•] can run different MCMC algorithms, and it has been incorporated in PEcAn [38] which allows the user to choose from a suite of different PBMs and driver conditions.

MCMC is more efficient than MC (e.g. random sampling from the prior followed by filtering out low likelihood vectors) because it focuses the sampling effort on the region of highest posterior probability; in most circumstances a chain of 10^4 to 10^5 steps is long enough even for parameter-rich models. However, because of the way MCMC works, it does require evaluation of the forest model for each new proposed parameter vector in order to calculate the likelihood. And if the data that are compared with model output in the likelihood function are derived from n different sites (each with different drivers x), then we need to run our model n times at each step. So despite the efficiency of MCMC, computational demand can still be considerable, all the more so for forest PBMs with short time steps and a long forest rotation to simulate. One proposed solution has been to include only a fraction of the model's parameters in the calibration [39], a practice referred to as 'parameter screening'. Parameter screening speeds up calibration by reducing the dimensionality of the parameter space that the MCMC needs to traverse. However, it thereby underestimates parameter uncertainty and should if possible be avoided.

The next section discusses a method for speeding up the forest model itself rather than the MCMC.

Emulators

An *emulator* is a replacement for the original model f. It is also a function of x and θ, but its output is not $f(x,\theta)$ itself but a probability distribution for it. So the emulator is used to probabilistically forecast the output from our model f without actually running the model. The emulator $p_g[..]$ is usually derived from a small training set of N modelling results $\{x_i, \theta_i, f(x_i,\theta_i)\}_{i=1:N}$:

$$p_g[g|x,\theta] = p[f(x,\theta) = g|\{x_i,\theta_i,f(x_i,\theta_i)\}].$$

Emulators are designed to compute faster than the original model, which may make sampling from the posterior distribution (e.g. using MCMC) more feasible. However, emulator speed comes at the cost of only producing a probabilistic answer. This makes the likelihood function more complicated. Say our uncertainty about measurement error is captured by the parameter σ_y. Then the likelihood function for the original model would be $L[\theta] = p[\epsilon_y = y - f(x,\theta)|\sigma_y]$, but when using the emulator we would have to integrate out our uncertainty for g:

$$L'[\theta] = \int_{g=-\infty}^{\infty} p\,[\epsilon_y = y - g|\sigma_y] p_g[g|x,\theta] dg.$$

In Bayesian calibration, the posterior would then be proportional to $p[\theta]L'[\theta]$. To keep calculation of the likelihood-integral as well as $p[\theta]L'[\theta]$ manageable, emulators are often selected for their mathematical properties. In particular, Gaussian process emulators are often used [23], which approximate the original model using the multivariate normal distribution and use convenient functions for calculation of covariances.

From uncertainty reduction to identifying structural error

A major benefit of Bayesian calibration of the model's parameters is that it makes it easier to discern structural errors in the model. After calibration, any remaining differences between observations and model outputs will be mainly due to data error and model structural error. It is impossible to fully disentangle these two errors, but some heuristic advice can be given [9].

The first recommended step after Bayesian calibration is to inspect the marginal posterior pdf's, i.e. the probability distribution for each individual parameter in the model. When any of these probability distributions is highly skewed, creeping up to its lower or upper prior bound, that may be an indication that the process affected by that parameter, or related processes, are incorrectly implemented. For example, if the litter decomposition rate constant a posteriori is higher than expected, that may indicate that other decomposition processes were overlooked.

The second step is decomposition of the mean squared deviance (MSD) of time series [40]. Forest PBMs simulate the dynamics of forest biogeochemistry, so after calibration we can compare observed and simulated time series of, for example, gas fluxes. MSD-decomposition allows us to split the MSD into three additive terms that quantify differences between the time series in their mean, variance and phase. For example, if model and data mean differ, the model may be missing a process; if the variances differ, the model may be missing a feedback mechanism; if the phases differ, the model may be missing a linked state variable e.g. in a decomposition chain. MSD-decomposition was applied to the mismatch between observed and simulated gas flux time series for four forest models calibrated for a Norway spruce site in Germany [9] and to two models for growth and respiration of *E. globulus* in Portugal [27].

The third step is to check any assumptions of parameter universality. Did we carry out a calibration using data from multiple forest sites assuming that parameter values should be generic, i.e. the same everywhere? Then we could redo the calibration with genotype- or site-

specific parameter estimation. However, this approach pits two extreme ideas against each other: generic vs. site-specific parameters. A better approach would be to allow parameters to vary between sites or genotypes but not completely independently - and let the calibration reveal the appropriate degree of parameter independence. Such a more flexible, hierarchical approach is becoming common in environmental and ecological modelling [41] and is explained in the next section.

Bayesian Hierarchical Modelling (BHM)

Hierarchical modelling constitutes a simple but powerful idea: do not assume that all the model's parameters are unknown constants, but allow for some or all of them to have variability - and capture that variability in another set of parameters, the so-called *hyperparameters* [8,41,42]. For example, modellers need not assume that all trees in a region have the same wood density, but can allow for spatial variation represented by two hyperparameters: mean wood density and its standard deviation.

In recent years, forest modellers have given increased attention to how various tree traits - usually represented as parameters in PBMs - vary with environmental conditions [43]. Hierarchical modelling allows for great flexibility in such cases: each parameter can range from being completely site- or condition-specific (large standard deviation) to being completely generic (zero standard deviation). When we carry out Bayesian calibration of the parameters and hyperparameters in a hierarchical model, statisticians refer to the approach as *Bayesian Hierarchical Modelling* [44]. In BHM, we specify a prior for the hyperparameters, $p[\xi]$, which then automatically implies a prior for the regular parameters, $p[\theta|\xi]$ or $p[\theta|x,\xi]$. The likelihood function does not change in this hierarchical setup, because model output $f(x,\theta)$ only depends on the regular model parameters, not the hyperparameters. So, in BHM, the posterior distribution is defined as:

$$p[\theta,\xi|y] \propto p[\xi]p[\theta|\xi]L[\theta]$$

Note that the use of BHM implies greater data needs. If the hyperparameters in the BHM stand for spatial variability, then calibration cannot be effective without data from dispersed sites.

A good introduction to hierarchical modelling in the context of forest research was provided by Dietze et al. [45]. They quantified variability in parameters for tree allometry (power law relationships for height vs. crown area) both within and between species in southeastern US forests, and showed how even species for which few data were available could be fit reliably within the hierarchical model.

Discrepancy and ensembles

In this section, we focus on how model structural error, also called the *discrepancy* [46,47] - we use the terms interchangeably - can complicate Bayesian parameter calibration. We first discuss the treatment of discrepancy in case we have only one forest model in our study, and thereafter consider the case where we have an ensemble of models.

Treatment of discrepancy in single-model calibration

As explained, Bayesian calibration combines prior information with the likelihood, and it is in the latter term that model outputs are compared with data. There are three different ways in which the discrepancy can be treated in the likelihood: (1) ignore it, (2) lump it with data error, (3) account for data error and discrepancy separately. We now discuss the merits of the three methods.

First, we may choose not to represent the discrepancy at all, in which case our calibration is formally conditional on the model being structurally correct. For ease of exposition, let us assume that error uncertainties are represented by zero-mean Gaussians. Then our likelihood function would read:

$$p[y|\theta] = N[y - f(x,\theta); \mu = 0, \sigma^2 = \sigma_y^2],$$

where we have the usual choice of pre-specifying the value of σ_y^2 or adding it to the parameters to be calibrated, in which case we must specify a prior distribution for it. Ignoring the discrepancy in this way is quite common; it is akin to linear regression without considering the possibility of a nonlinear function. It is a simple method that underestimates predictive uncertainty of the model because it can only propagate input uncertainty, not structural uncertainty. This method has been used in forest modelling [31], but we may want to move away from it.

The second approach also does not recognize the discrepancy explicitly but estimates the sum of data and model structural error, i.e. the overall mismatch between data and model outputs. Again using Gaussians, the likelihood would then read:

$$p[y|\theta] = N[y - f(x,\theta); \mu = 0, \sigma^2 = \sigma_{mismatch}^2],$$

where we make no attempt to decompose $\sigma_{mismatch}^2$ into contributions from data error and discrepancy. In this method, the $\sigma_{mismatch}^2$-term cannot be pre-specified - because we cannot foresee how much model outputs will deviate from data - and must be added as a parameter to be calibrated. This method was used by, amongst others, Kavetski et al. [48]. Despite being simple, this method is also not to be advised, as it obscures the difference between errors in data and model [46]. It also leads to overestimation of predictive uncertainty: we do not want to predict future measurements ($f(x,\theta) + \epsilon_{mismatch} = f(x,\theta) + \epsilon_y + \epsilon_{model}$) but future true values ($f(x,\theta) + \epsilon_{model}$).

This leaves the third method, rarely used but to be recommended: explicitly distinguishing errors in data from those in model structure:

$$p[y|\theta] = N[y - f(x,\theta); \mu = 0, \sigma^2 = \sigma_y^2 + \sigma_{model}^2],$$

where we have to specify a prior for σ_{model}. In this approach, we can account for all sources of model predictive uncertainty ($p[x]$, $p[\theta|y]$, $p[\epsilon_{model}]$). And information we may have about the precision of our data (σ_y^2) can still be used, in contrast to the 'mismatch' method. Although this third method is in principle superior to any other, it may be difficult to specify the prior for the discrepancy because simple Gaussians are unlikely to be adequate. However, the method has been sketched out in ground-breaking papers [23,46] and should be developed further.

Treatment of discrepancy in multi-model calibration

Uncertainty with respect to model structure itself is more difficult to visualize and represent formally than parameter uncertainty. For parameters, as we have seen, we can simply define a probability distribution $p[\theta]$, but how would we specify a probability distribution $p[f]$? A practical solution to this issue is to collect multiple forest models and assume that the set forms a representative sample from 'model space'. Such *ensemble modelling* is becoming increasingly common [47]; for examples with forest models see e.g. [13,17]. It allows us to define $p[f]$ as a discrete probability distribution over the models in the set. There are then two ways to proceed, based on how we treat the discrepancy. We may make the additional assumption that one of the models is exactly correct, i.e. has no discrepancy. That simplifies the treatment considerably because we can then find data and use Bayes' Theorem to reduce the uncertainty about which model is correct:

$$p[f|y] \propto p[f]p[y|f],$$

where the model likelihood, $p[y|f]$, requires integrating out parameter uncertainty:

$$p[y|f] = \int p[y|\theta, f]p[\theta]d\theta.$$

In this approach, predictive uncertainty can make use of the whole model ensemble by weighting each model's contribution by its posterior probability $p[f|y]$. This is referred to as Bayesian Model Averaging [49], and an example using six forest models was given by Van Oijen et al. [13]. Remarkably, BMA-predictions tend to be superior to those of any single model in the ensemble.

A more advanced method would be to recognize that each model in the ensemble is wrong and try to estimate any biases shared by all models, as well as discrepancies specific to individual models. This method was initially proposed by Chandler [47] for climate model ensembles, based on the realisation that there is always some convergence in model development, leading to errors that are shared by all models in the ensemble. Although this approach is in principle the most comprehensive and realistic, it requires assigning priors to the biases, for which no heuristics yet exist.

Uncertainty communication

Our theme in this paper has been that uncertainties are represented by probability distributions. This raises the issue: how to communicate probability distributions to others in, for example, a scientific publication or a report for a forest manager? This is a complicated issue, exacerbated by the fact that the posterior predictive uncertainty for different output variables is typically represented by a high-dimensional joint probability distribution with possibly many non-zero correlations.

Formally, only the full joint probability distribution captures all information about predictive uncertainty. However, summarization is inevitably required, and the way to report uncertainty should depend on both the shape of the distribution and the needs of the persons who are expected to read the report [50]. If the distribution is symmetric and bell-shaped, then providing the mean value and the standard deviation may be sufficient. For more complicated distributions, reporting quantiles may be preferable and strong

correlations should be identified [31]. However, we note that the field of uncertainty communication is still in flux, and several researchers recommend the use of different graphical representations of uncertainty over verbal or numerical reporting [50,51•]. Examples of reporting posterior uncertainties following Bayesian calibration of forest models are provided by, amongst others, Van Oijen et al. [9,13,31] and a map showing how predictive uncertainty of carbon sequestration in UK forests varies spatially is provided by Van Oijen & Thomson [15].

The Bayesian method facilitates uncertainty analysis, and this in turn makes it easy to calculate *risk* as the probabilistic expectation of loss - in terms of a model output variable such as stem volume or primary productivity - due to some type of hazardous event. A reader might prefer to see depictions of risk rather than the underlying probability distribution for the variable. Recently, a novel method was proposed for decomposing risk into *hazard probability* and *ecosystem vulnerability*, and this was used with six vegetation models (including the forest model BASFOR) to analyse future drought risks to vegetation in Europe [52•] - showing highest risks for Mediterranean forests. Risk, vulnerability and hazard probability were depicted visually on maps and as bar charts.

In its simplest form, the Bayesian method formalizes the flow of information from data to model parameters and ultimately model output. However, we have shown that the method can be expanded to a much more comprehensive framework involving multiple models, model emulators, uncertain drivers, uncertain parameters and hyperparameters combined in a Bayesian hierarchical model. Whilst powerful, the complexity of the framework may obscure a reader's grasp of the flows of information between different components. In this paper, we opted for verbal descriptions and mathematical equations to explain the methodology, but a visual alternative exists in the form of *graphical models*, such as Bayesian Belief Networks [53]. BBNs typically depict all uncertain quantities in separate circles or boxes, linked by arrows representing conditional probabilities. For example, an arrow from a circle labeled θ to one labeled y would represent the likelihood $p[y|\theta]$. The use of graphical models might thus be considered as a tool for introducing the Bayesian approach to uncertainties in forest modelling to a new audience, although equations remain essential to communicate the exact details of the approach.

Discussion

Advantages of the Bayesian approach

Bayesian methods are increasingly used by modellers, including forest modellers. The approach offers a consistent way of quantifying and reducing uncertainties, rigorously based in probability theory. The concepts are simple, consisting of no more than defining and combining priors and likelihoods. This provides a straightforward procedure for assessing the reliability of models and model predictions. Moreover, the approach can be used to evaluate different measurement strategies. For example, when considering to replace a piece of equipment with a higher-precision tool that is more expensive, we could first run Bayesian calibrations with virtual data of differing precision in order to assess how much the expensive tool is likely to reduce predictive uncertainty [31].

Bayesian calibration may also resolve an issue typical for biogeochemistry PBMs: the need for model *spin-up*. When measurements of soil carbon and nutrients are used as model drivers, to initialise the soil pools, the PBM may be forced out of equilibrium, with the first simulated years or decades showing unrealistic trends in soil composition and gas fluxes when the model slowly reverts to equilibrium. Modellers tend to prevent these early unrealistic dynamics by discarding soil data and instead initializing the model with equilibrium values for the soil pools - which are found after long spin-up runs. However, this amounts to replacing unrealistic dynamics with unrealistic initial pools, and cannot be recommended. Yeluripati et al. [54] showed a superior approach where soil measurements were not discarded but used together with flux-data in a Bayesian calibration of the parameters and initial values of the model. This way, spin-up was avoided.

The increased uptake of Bayesian methods by the modelling community has been stimulated by the appearance of excellent introductory books [29,30•] and computational tools (BUGS, JAGS, Stan, BayesianTools).

The need for improved methodology

Bayesian methods are computationally demanding and their recent uptake by many modellers has only been made possible by the advent of fast computational methods for sampling from distributions, in particular MCMC. The new computational tools have stimulated interest in application of Bayesian methods to complex, slow models and to complex uncertainty assessments, e.g. those where multiple models are considered or where parameter uncertainty is defined hierarchically. In such cases, computational demand becomes a bottleneck again, and we expect that the search for new MCMC-algorithms and model emulators will continue.

One issue that remains is how to treat the discrepancy between model behaviour and real system dynamics. Whereas this was either ignored or treated superficially before in Bayesian calibration, we expect intensified methodological research on how to represent discrepancy and prior uncertainty about it. Ensemble modelling, which likewise addresses model structural uncertainty, is also likely to remain of key interest.

So far, the main applications of Bayesian hierarchical modelling have been to fairly simple stochastic models, especially in the area of spatiotemporal modelling [44]. BHM allows for the possibility that parameters measured in different locations may differ yet preserve a degree of family resemblance - which the BHM represents using hyperparameters. We expect that the modelling of, say, productivity of a tree species across Europe would benefit from a hierarchical approach, with hyperparameters representing genetic variance between tree provenances in parameters that affect light-use efficiency and drought tolerance. But the scope for using BHM goes well beyond representing spatial or taxonomic differences: the many different sources of error reviewed in this paper can all be represented explicitly, with their own hyperparameters, in a hierarchical statistical model [41]. For example, BHM facilitates the use of different data sets in one Bayesian calibration whereby differences in measurement precision and accuracy are represented by data set-specific and generic error-parameters [42].

Uncertainties quantified and reduced using Bayesian methods must be communicated to end-users of the modelling studies. There seems to be no standard approach or clear trend visible in how uncertainties are being reported. However, research on uncertainty

communication is continuing and graphical modelling may play an increasingly prominent role. Graphical models have been used to convey uncertainty about the impact of management decisions on the different ecosystem services that a single ecosystem can provide [53]. These issues may acquire even more significance in the future with the expected increase in frequency and intensity of extreme climatic events, and their concomitant risks to forests [3•]. Risk analysis is a form of uncertainty analysis and will therefore benefit from Bayesian methods applied to forest and climate models in order to provide reliable estimates of risk, vulnerability and hazard probability.

Conclusions

We conclude by listing the major changes we are expecting in the future treatment of errors and uncertainties in forest models. The list can also be taken as a point-wise summary of the methods discussed in this paper. We expect:

- more use of Bayesian methods in general,
- more use of Bayesian hierarchical modelling,
- replacement of model spin-up by Bayesian calibration,
- more use of ensemble modelling followed by Bayesian model averaging,
- novel ways to account for model structural error in calibration,
- further development of general software for Bayesian calibration of complex models,
- faster MCMC algorithms,
- more use (and novel types) of model emulators,
- novel uncertainty visualization techniques,
- more use of graphical modelling,
- more use of risk analysis as a follow-up to Bayesian calibration.

Acknowledgements I thank Peter Levy (CEH-Edinburgh) for his comments on this paper.

Conflict of Interest There are no conflicts of interests to declare.

Human and animal rights This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as: • Of importance; •• Of major importance.

1. Mäkelä A, Río M del, Hynynen J, Hawkins MJ, Reyer C, Soares P, et al. Using stand-scale forest models for estimating indicators of sustainable forest management. Forest Ecology and Management 2012;285:164–78.
2. Hickler T, Rammig A, Werner C. Modelling CO_2 Impacts on Forest Productivity. Current Forestry Reports 2015;1:69–80.
3. • Reyer C. Forest Productivity Under Environmental Change - a Review of Stand-Scale Modeling Studies. Current Forestry Reports 2015;1:53–68. *This paper is useful for forest modellers aiming to provide global assessments of the impacts of environmental change: it identifies those forest types and parts of the globe for which data are at present lacking.*
4. Fontes L, Bontemps J-D, Bugmann H, Van Oijen M, Gracia C, Kramer K, et al. Models for supporting forest management in a changing environment. Forest Systems 2010;3:8–29.
5. Hartig F, Dyke J, Hickler T, Higgins SI, O'Hara RB, Scheiter S, et al. Connecting dynamic vegetation models to data - an inverse perspective. Journal of Biogeography 2012;39:2240–52.
6. Jaynes ET. Probability theory: The logic of science. Cambridge university press; 2003.
7. Sivia D, Skilling J. Data Analysis: A Bayesian Tutorial. 2 edition. Oxford: Oxford University Press, U.S.A. 2006.
8. Ogle K, Barber JJ. Bayesian Data in Plant Physiological and Ecosystem Ecology. In: Lüttge U, Beyschlag W, Murata J, editors. Progress in Botany. Springer Berlin Heidelberg; 2008. pages 281–311.
9. Van Oijen M, Cameron D, Butterbach-Bahl K, Farahbakhshazad N, Jansson P-E, Kiese R, et al. A Bayesian framework for model calibration, comparison and analysis: Application to four models for the biogeochemistry of a Norway spruce forest. Agricultural and Forest Meteorology 2011;151:1609–21.
10. • Van Oijen M, Cameron D, Levy PE, Preston R. Correcting errors from spatial upscaling of nonlinear greenhouse gas flux models. Environmental Modelling & Software 2017;94:157–65. *This paper reviews and revives the powerful yet neglected idea that errors caused by spatial aggregation of the inputs of nonlinear models can be estimated well using a Taylor-expansion of the model, together with information on spatial variability of the inputs.*
11. Cameron DR, Van Oijen M, Werner C, Butterbach-Bahl K, Grote R, Haas E, et al. Environmental change impacts on the C- and N-cycle of European forests: A model comparison study. Biogeosciences 2013;10:1751–73.
12. Van Oijen M, Ågren GI, Chertov O, Kellomäki S, Komarov A, Mobbs D, et al. Evaluation of past and future changes in European forest growth by means of four process-based models. Causes and Consequences of Forest Growth Trends in Europe: Results of the Recognition Project 2008;21:183-99.
13. Van Oijen M, Reyer C, Bohn F, Cameron D, Deckmyn G, Flechsig M, et al. Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. Forest Ecology and Management 2013;289:255–68.
14. Levy PE, Wendler R, Van Oijen M, Cannell MG, Millard P. The effect of nitrogen enrichment on the carbon sink in coniferous forests: Uncertainty and sensitivity analyses of three ecosystem models. Water, Air, & Soil Pollution: Focus 2005;4:67–74.

15. Van Oijen M, Thomson A. Toward Bayesian uncertainty quantification for forestry models used in the United Kingdom Greenhouse Gas Inventory for land use, land use change, and forestry. Climatic Change 2010;103:55–67.
16. Hyvönen R, Ågren GI, Linder S, Persson T, Cotrufo MF, Ekblad A, et al. The likely impact of elevated [CO_2], nitrogen deposition, increased temperature and management on carbon sequestration in temperate and boreal forest ecosystems: A literature review. New Phytologist 2007;173:463–80.
17. Rollinson CR, Liu Y, Raiho A, Moore DJP, McLachlan J, Bishop DA, et al. Emergent climate and CO_2 sensitivities of net primary productivity in ecosystem models do not agree with empirical data in temperate forests of eastern North America. Global Change Biology 2017;23:2755–67.
18. Van Oijen M, Cannell MGR, Levy PE. Modelling biogeochemical cycles in forests: State of the art and perspectives. Towards the sustainable use of European forests-Forest ecosystem and landscape research: scientific challenges and opportunities, edited by: Andersson, F., Birot, Y., and Päivinen, R., European Forest Institute, Joensuu, Finland 2004;157–69.
19. Clark JS, Iverson L, Woodall CW, Allen CD, Bell DM, Bragg DC, et al. The impacts of increasing drought on forest dynamics, structure, and biodiversity in the United States. Global Change Biology 2016;22:2329–52.
20. Johnson MO, Galbraith D, Gloor M, De Deurwaerder H, Guimberteau M, Rammig A, et al. Variation in stem mortality rates determines patterns of above-ground biomass in Amazonian forests: Implications for dynamic global vegetation models. Global Change Biology 2016;22:3996–4013.
21. Schlesinger WH, Dietze MC, Jackson RB, Phillips RP, Rhoades CC, Rustad LE, et al. Forest biogeochemistry in response to drought. Global Change Biology 2016;22:2318–28.
22. •• Medlyn BE, Zaehle S, De Kauwe MG, Walker AP, Dietze MC, Hanson PJ, et al. Using ecosystem experiments to improve vegetation models. Nature Climate Change 2015;5:528–34. *This very clear paper can be seen as a counterpoint to the present one in that it does not mention probability theory as a tool to reduce uncertainty, but instead focuses on consistency with mechanistic understanding.*
23. Kennedy MC, O'Hagan A. Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 2001;63:425–64.
24. Fu YH, Campioli M, Van Oijen M, Deckmyn G, Janssens IA. Bayesian comparison of six different temperature-based budburst models for four temperate tree species. Ecological Modelling 2012;230:92–100.
25. Reyer CPO, Flechsig M, Lasch-Born P, Oijen M van. Integrating parameter uncertainty of a process-based model in assessments of climate change effects on forest productivity. Climatic Change 2016;137:395–409.
26. Sutton MA, Simpson D, Levy PE, Smith RI, Reis S, van Oijen M, et al. Uncertainties in the relationship between atmospheric nitrogen deposition and forest carbon sequestration. Global Change Biology 2008;14:2057–63.
27. Minunno F, van Oijen M, Cameron D, Cerasoli S, Pereira J, Tomé M. Using a Bayesian framework and global sensitivity analysis to identify strengths and weaknesses of two process-based models differing in representation of autotrophic respiration. Environmental Modelling & Software 2013;42:99–115.

28. Bayes T. An Essay towards Solving a Problem in the Doctrine of Chances. Philosophical Transactions 1763;53:370–418.
29. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis, Third Edition. 3 edition. Boca Raton: Chapman and Hall/CRC; 2013.
30. • McElreath R. Statistical Rethinking: A Bayesian Course with R Examples. 2016. *Perhaps the best introduction to Bayesian methods for absolute beginners, explaining both concepts and methods with impressive clarity.*
31. Van Oijen M, Rougier J, Smith R. Bayesian calibration of process-based forest models: Bridging the gap between models and data. Tree Physiology 2005;25:915–27.
32. Levy PE, Cowan N, Van Oijen M, Famulari D, Drewer J, Skiba U. Estimation of cumulative fluxes of nitrous oxide: Uncertainty in temporal upscaling and emission factors. European Journal of Soil Science 2017;
33. Patenaude G, Milne R, Van Oijen M, Rowland CS, Hill RA. Integrating remote sensing datasets into ecological modelling: A Bayesian approach. International Journal of Remote Sensing 2008;29:1295–315.
34. Höglind M, Van Oijen M, Cameron D, Persson T. Process-based simulation of growth and overwintering of grassland using the BASGRA model. Ecological Modelling 2016;335:1–15.
35. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. The Journal of Chemical Physics 1953;21:1087–92.
36. Van Oijen M. Bayesian Calibration (BC) and Bayesian Model Comparison (BMC) of process-based models: Theory, implementation and guidelines. 2008.
37. •• Hartig F, Minunno F, Paul S, Cameron D. Package 'BayesianTools'. 2017. https://cran.r-project.org/web/packages/BayesianTools/BayesianTools.pdf. *This document introduces a major software development: an R-package that facilitates the use of Bayesian calibration by means of MCMC for complex process-based models of forests and other vegetation.*
38. Dietze MC, Lebauer DS, Kooper R. On improving the communication between models and data. Plant, Cell & Environment 2013;36:1575–85.
39. Minunno F, van Oijen M, Cameron DR, Pereira JS. Selecting Parameters for Bayesian Calibration of a Process-Based Model: A Methodology Based on Canonical Correlation Analysis. SIAM/ASA Journal on Uncertainty Quantification 2013;1:370–85.
40. Kobayashi K, Salam MU. Comparing simulated and measured values using mean squared deviation and its components. Agronomy Journal 2000;92:345–52.
41. Cressie N, Calder CA, Clark JS, Hoef JMV, Wikle CK. Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. Ecological Applications 2009;19:553–70.
42. Ogle K. Hierarchical Bayesian statistics: Merging experimental and modeling approaches in ecology. Ecological Applications 2009;19:577–81.
43. Simpson AH, Richardson SJ, Laughlin DC. Soilclimate interactions explain variation in foliar, stem, root and reproductive traits across temperate forests. Global Ecology and Biogeography 2016;25:964–78.
44. Cressie N, Wikle CK. Statistics for Spatio-Temporal Data. Wiley; 2011.

45. Dietze MC, Wolosin MS, Clark JS. Capturing diversity and interspecific variability in allometries: A hierarchical approach. Forest Ecology and Management 2008;256:1939–48.
46. Rougier J. Probabilistic Inference for Future Climate Using an Ensemble of Climate Model Evaluations. Climatic Change 2007;81:247–64.
47. Chandler RE. Exploiting strength, discounting weakness: Combining information from multiple climate simulators. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 2013;371:20120388–8.
48. Kavetski D, Kuczera G, Franks SW. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resources Research 2006;42.
49. Kass RE, Raftery AE. Bayes Factors. Journal of the American Statistical Association 1995;90:773–95.
50. Spiegelhalter D, Pearson M, Short I. Visualizing Uncertainty About the Future. Science 2011;333:1393–400.
51. • Milne AE, Glendining MJ, Lark RM, Perryman SAM, Gordon T, Whitmore AP. Communicating the uncertainty in estimated greenhouse gas emissions from agriculture. Journal of Environmental Management 2015;160:139–53. *A thorough assessment of six different ways of communicating uncertainty, both verbal and visual.*
52. •• Van Oijen M, Balkovič J, Beer C, Cameron DR, Ciais P, Cramer W, et al. Impact of droughts on the carbon cycle in European vegetation: A probabilistic risk analysis using six vegetation models. Biogeosciences 2014;11:6357–75. *This paper shows how risk can be formally decomposed as the product of two terms: hazard probability and ecosystem vulnerability. The method is applied to the results of six vegetation models, and predicts future drought risk to increase mainly in Mediterranean vegetation because of changes in hazard probability.*
53. Smith R, Dick J, Trench H, Van Oijen M. Extending a Bayesian Belief Network for ecosystem evaluation. In: Conference paper of the 2012 Berlin Conference of the Human Dimensions of Global Environmental Change on“ Evidence for Sustainable Development”, 5-6 October 2012, Berlin, Germany. 2012.
54. Yeluripati JB, van Oijen M, Wattenbach M, Neftel A, Ammann A, Parton W, et al. Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. Soil Biology and Biochemistry 2009;41:2579–83.