Tools for model evaluation and parameterization

Haythem Ben Touhami, Romain Lardy, Gianni Bellocchi

AnimalChange

SEVENTH FRAMEWORK PROGRAMME

THEME 2: FOOD, AGRICULTURE AND FISHERIES, AND BIOTECHNOLOGIES

Grant agreement number: FP7- 266018

DELIVERABLE 15.2

Deliverable title: Tools for model evaluation and parameterization

Abstract: The goal of task 15.2 of AnimalChange project is to quantify and reduce uncertainties associated with model parameterization.

Due date of deliverable: M24 Actual submission date: M25

Start date of the project: March 1st, 2011 Duration: 48 months

Organisation name of lead contractor: INRA

Authors: Haythem Ben-Touhami, Romain Lardy and Gianni Bellocchi

Revision: V1

Dissemination level: PU

Table of Contents

1. Introduction

Grassland ecosystems are highly complex and dynamic. The many interactions between herbivores, vegetation, soil and the atmosphere, and the role of management practices make our ability to experiment on these systems extraordinarily limited. Thus, testing scenarios of climate change using ecosystem models which simulate physical, chemical, and biological processes in great detail is an imperative. The models that we consider in AnimalChange are deterministic, that is, running the model with the same inputs always produces the same outputs. Using models allows indeed a greater insight and understanding of these processes and interactions than it can be assumed by just considering experimental evidences. They are therefore widely used in climate change impact projections, especially for long-term analyses (Johnson *et al.*, 2008). Such projections need to account for the uncertainty cascade resulting from multiple sources including climate scenarios, the impact model, local climate, vegetation and soil conditions (Olesen *et al.*, 2007; Soussana *et al.*, 2010) and to consider farming practices adaptation (Tubiello *et al.*, 2007).

A major problem with the use of complex ecological models is incomplete knowledge of input variables as well as model parameters. This means that model estimates are highly dependent on parameter settings, leading to a large uncertainty due to uncertainties in parameter values, driving variables (climate, soil and management) and model structure (Gabrielle *et al.*, 2006). There is thus a need to set (calibrate) parameter values to obtain good agreement between model outputs and observations (Wallach *et al.*, 2011). At the same time, uncertainty associated with model outputs needs to be quantified and documented (van Oijen and Thomson, 2010). A two-step methodology has become increasingly used to (try to) figure out the uncertainties associated with model parameters (for instance, the Good Practice Guidance, Methodological Tier 2, of the Intergovernmental Panel on Climate Change, through http://www.ipcc-nggip.iges.or.jp). First, uncertainties are quantified by expressing them as probability distribution functions (PDFs). Then, representative samples are taken from the PDFs to propagate parameter uncertainty forward through the model calculations. This is an exercise that applies the Bayes theorem (Bayes, 1763) to incorporate as much information as possible into PDFs. The principles of Bayesian predictive inference were set out in Aitchison and Dunsmore (1975). The reader may refer to Kennedy and O'Hagan (2001), van Oijen *et al.* (2005), and Patenaude *et al.* (2008) for more details about Bayesian calibration.

This choice of Bayesian calibration comes from the fact that it applies to models of any type or size and combines model parameterization and uncertainty analysis. The main characteristic of the Bayesian approach is that it quantifies model inputs and outputs in the form of probability distributions, and applies the rules of probability theory to update the distributions when new data are obtained (Sivia, 1996). This approach prevents the generation of uncertainty estimates that are unrealistically high, or even unusable in practice, by reducing uncertainties where possible (*e.g.* limiting calibration to a set of most influential parameters for which prior information about their variability is available), but also by combining direct and indirect information when estimating uncertainties. PDFs are initially the expression of current imprecise knowledge about model parameter values, and this prior probability is then updated with the measured observations into posterior probability distribution using the Bayes' theorem (Makowski *et al.*, 2006; Gallagher and Doherty, 2007). In this way, Bayes' theorem is valuable because it tells us how our uncertainty about the parameters decreases because of that information. Useful information could be measurements of any key output variables of interest. The method therefore not only propagates uncertainty in input variables and parameters to model outputs, but also uses data on output variables to reduce the uncertainty in inputs and parameters. Bayesian calibration yields a posterior (calibrated) distribution for the parameters, which can be

summarized in the form of a maximum likelihood vector and covariance matrix. The uncertainty of the parameters can be quantified by running the model with different parameter settings, sampled from the posterior distribution. The posterior parameter distribution is approximated in the form of a representative sample of parameter values. This is generally achieved by means of Markov Chain Monte Carlo simulation, which is suitable for process-based models because of its simplicity and because it does not require advanced knowledge of the shape of the posterior distribution (van Oijen *et al.*, 2005).

2. Uncertainties in models

The widespread application of impact models is accompanied by a widespread concern about quantifying the uncertainties prevailing in their use. The following is one way of classifying the various sources of uncertainty.

1. **PARAMETER UNCERTAINTY** The parameters of a model are the constants of its equations. They are generally partly unknown inputs. Generally, they specify features of a particular application context, but they may also be more global parameters, assumed to have a common value over a range of contexts or even in all contexts. In general, model estimates are highly affected by the uncertainty in parameter values and additional uncertainty is added by the simultaneous occurrence of climate changes and extreme events.
2. **MODEL INADEQUACY** Even if there is no parameter uncertainty, so that we know the true values of all the inputs required to make a particular prediction of the process being modelled, the predicted value will not equal the true value of the process. This discrepancy is model inadequacy that is the difference between the true mean value of the real world process and the model output at the true values of the inputs. This means that there are deficiencies in the model structure in terms physiological and biochemical mechanisms.
3. **OBSERVATION ERROR** In tackling the calibration problem, actual observations are used. The possibility of observation errors adds further uncertainty, although in practice it may not be feasible to separate it.
4. **CODE UNCERTAINTY** The output of a computer code, given any particular configuration of inputs, is in practice not known until we actually run it with those inputs. The output is not really unknown because it is a known function of the inputs. So it is because there is a mathematical model which defines that function. Nevertheless, the relationship is so complex that it needs to be implemented in a computer code, and it is not realistic to say that the output is known for given inputs before the code is actually run and the output is seen. It may not be practical to run the model to observe the output for every input configuration of interest. Therefore, uncertainty about model output needs to be acknowledged.
5. **RESIDUAL VARIABILITY** The model is supposed to predict the value of some real processes under conditions specified by the inputs. In practice, the modelled process may not always take the same value if those conditions are repeated. This may happen when the process is inherently unpredictable and stochastic or where the model lacks detail to discriminate between conditions which actually lead to different process values.
6. **PARAMETRIC VARIABILITY** It is often desired to use the model to predict the process when some of the conditions specified in the inputs are uncontrolled and unspecified. Here, the inputs require more detail than we desire or are able to use. By leaving some of the input parameters unspecified (i.e. set to a default value), the

predicted process value acquires an extra uncertainty, which is called parametric uncertainty.

The goal of this report is focus attention on solutions designed to reduce the first type of uncertainty.

3. Model calibration and uncertainty analysis

The objective of uncertainty analysis is to study the distribution of the model output that is induced by probability distributions on inputs. The input parameter distributions may be formulations of parameter uncertainty, that is, for parameters whose values are unknown or left unspecified. The simplest approach to uncertainty analysis is a Monte Carlo solution in which configurations of inputs are drawn at random from their distribution. The model is then run for each sample input configuration and the resulting set of outputs is a random sample from the output distribution to be studied (Helton, 1993 for a review).

To improve the model results, a method is needed to i) quantify how uncertainty about inputs and model structure causes output uncertainty; ii) efficiently use input and output data to reduce uncertainties. The final topic of our survey is calibration. The conventional way of estimating unknown parameters is by an *ad hoc* search for the best-fitting values. Some account is thereby taken of observation errors, residual variation and model inadequacy, but only implicitly through the measure of the discrepancy in fits. This measure is not generally developed by modelling these error terms in any explicit way and is usually entirely heuristic. Also, since the estimated values are then treated as if they were known, the subsequent predictions do not take into account the remaining parameter uncertainty. In contrast, effectively Bayesian estimation methods do allow fully for parameter uncertainty (*e.g.* Romanowicz *et al.*, 1994). Bayesian calibration applies to models of any type and combines model parameterization and uncertainty analysis. Basically, an initial Monte Carlo sample is drawn from what amounts to the prior distribution of the unknown parameters and is then weighed by a likelihood term. Then, model predictions can be made using all the sample parameter configurations, and the result is a weighed sample from the posterior distribution.

A Bayesian approach calibrates a model by using observations from the real process. Bayesian analyses of the data used to parameterize the model produce precisely the description of input uncertainty that is needed for an uncertainty analysis. Thus, Bayesian calibration is naturally followed by prediction and uncertainty analysis of the process. The posterior summaries take into account for all remaining sources of uncertainty. The source code of the model is basically treated as a black box, and the methods described in this paper are applicable to source codes of arbitrary complexity. The problem of selection of calibration parameters might be dealt with by calibrating all parameters rather than a few selected ones. Non-influential parameters can be handled because the posterior distribution of a totally irrelevant parameter will be the same as the prior distribution. However, with most Monte Carlo methods for Bayesian analysis, the computational burden hardly grows by the addition of irrelevant parameters.

In particular, complex models with a large number of parameters may make technically unfeasible to simultaneously calibrate all of them. It can also be shown that this is logically inconsistent because of the lack of resolution of certain parameters. For instance, cell processes may be tied to model behavior through highly nonlinear functions, the respective parameters may be unobservable and the valid values of them may fall within a very narrow range that is hard to capture. Most of these parameters can then be assumed to be fixed and identical for all model realizations. In addition, there are other parameters that are more suitable to be fitted to the experimental data across a range of conditions. Hence, at least as

a first step, a set of parameters can be selected with respect to the availability of prior information to support calibration and their relevance for the output variables. In this case, the application of a sensitivity analysis method is recommended to screen out insignificant parameters and to identify the most relevant parameters for a given model output. This becomes imperative with complex ecological models with large parameter sets and high temporal resolution, which need a parsimonious screening method to identify the most important parameters in a specific context. The Morris method (Morris, 1991) of sensitivity analysis, changing parameters "one-at-a-time", is considered the best way to reduce computational costs when the number of input parameters is high (several tens or higher) and/or the model runs are computationally expensive. We do not deal explicitly with sensitivity analysis in this report: appropriate techniques are outlined, for instance in Saltelli *et al.* (2004) and Confalonieri *et al.* (2010).

4. Bayesian calibration

Bayesian calibration operates on a single model, whose structure is assumed to be correct, but whose parameter values are not exactly known. The first step in Bayesian calibration is expressing initial "prior" uncertainty about the parameter values in the form of a joint probability density function (PDF), symbolized $P(\vartheta)$. In the absence of specific information, bounded normal or uniform distribution are common choices for $P(\vartheta)$. Bayesian calibration uses data on model outputs to update the PDF for the parameters. This is done by applying Bayes' theorem: $P(\vartheta|D)=P(\vartheta)\cdot P(D|\vartheta)/P(D)$, where $P(\vartheta|D)$ is the posterior distribution for ϑ given the data D, $P(D|\vartheta)$ is the likelihood of the data given model output using parameters ϑ, and $P(D)$ is a normalization constant. The data are only used in the calculation of the likelihood, with measurement error determining how likely any given model-data mismatch is. If the data are informative, i.e. are plenty and have low measurement error, then the posterior PDF $P(\vartheta|D)$ will be narrower, more sharply peaked than the prior $P(\vartheta)$, indicating that parameter uncertainty is reduced.

There is no closed-form solution for Bayes' theorem applied to process-based models, because it is always possible to generate a representative sample from the posterior by Markov Chain Monte Carlo (MCMC) methods. MCMC methods are a class of algorithms for sampling from probability distributions based on constructing a Markov chain (mathematical system that undergoes transitions from one state to another) that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution. The quality of the sample improves as a function of the number of steps. Many MCMC methods move around the equilibrium distribution in relatively small steps, with no tendency for the steps to proceed in the same direction. These methods are easy to implement and analyze, but unfortunately it can take a long time for the walker to explore all of the space. Among the random walk MCMC methods, the Metropolis-Hasting method (Metropolis *et al.*, 1953) obtains a sequence of random samples from a probability distribution in which each new point in the chain is found by randomly generating a candidate parameter vector, which can be accepted or rejected. The candidate parameter vector is usually generated by taking a multivariate normal step away from the current vector. This means that a covariance matrix is needed to define this multivariate normal "proposal distribution". Whether a proposed candidate vector is accepted or not depends on the Metropolis ratio, which is the ratio of two products: likelihood times prior for the candidate and likelihood times prior for the current point. The Metropolis ratio is always accepted if it is larger than 1 (i.e. the candidate point has a higher posterior probability then the current point). If the Metropolis ratio is less than 1 (i.e. the candidate is

less probable than the current vector), then the candidate can still be accepted but only with probability equal to the Metropolis ratio. The chain stops when it has converged, i.e. it has explored the posterior parameter space adequately. Some iteration at the beginning of the chain can be rejected as unrepresentative of the chains (burn-in stage, according to van Oijen *et al.*, 2005).

For each dataset D_i, the logarithm of the data likelihood, $P(D|\theta)$, which is the probability of the data given the parameters, is set up as follows:

$$log\ P(D|\vartheta)_i = \sum_{j=1}^{k} \left(-0.5 \left(\frac{f_i\ (\omega; \theta) - D_j}{\sigma_j} \right)^2 - 0.5\, log(2\pi) - log(\sigma_j) \right)$$

where D_j is the data value measured on sampling date j in the dataset D_i, σ_j is the standard deviation of the data on that date, ω is the vector of model input data at the same date, $f_i(\omega;\theta)$ is the model simulation of D_j with the parameter vector (generated by the Metropolis-Hastings MCMC algorithm based on Haario *et al.*, 2001), and K is the total number of observational dates in the dataset.

The following details are from the methodology described by Ben-Touhami *et al.* (2012a):

- Markov chains of length 50,000–100,000 sound reasonable using a multivariate normal PDF to generate candidate parameter vectors
- Running one to three parallel Markov chains may ensure convergence of parameter estimates
- The maximum of a posteriori estimation (as approximated by the maximum likelihood estimate) can be used to obtain a single estimate of each parameter for use in the simulations

5. Example 1: Stubai (Austria) grassland site

Data and model

It is presented here an analysis of Bayesian calibration results from the simulation of a European grassland site located at Stubai, in Austria (47° 05′ North, 11° 11′ East, 1850 m a.s.l.), which is a typical mountain meadow for the Central Eastern Alps. A detailed description of this study is given in Ben-Touhami *et al.* (2012a). The data used were collected over the period 2008-2009, over which the grassland plot is cut once a year, lightly grazed in late summer, and fertilized with manure / slurry roughly every three years. The experimental layout included a control plot in which none of the factors was manipulated and a treated plot received reduced precipitation using an appropriately sized rain shelter. Table 1 shows the set of five variables (D_1, …, D_5) measured and used in this study for the purpose of calibration, with soil moisture and soil temperature taken at three depths (0.05 m, 0.10 m and 0.20 m) and averaged at five-day time step.
The grassland site was simulated with the Pasture Simulation model (PaSim, https://www1.clermont.inra.fr/urep/modeles/pasim.htm, APP ID: 44 IDDN.FR.001.220024.000.R.P.2012.000.10000), a plot-scale, multi-year, complex (~130 parameters for vegetation) biogeochemical model operating on a time step of a 1/50th of a

day and consisting of sub-models for plant, animals, microclimate, soil biology, and soil physics. It simulates water, carbon (C), nitrogen (N) and energy cycles. Photosynthetic-assimilated C is either respired or allocated dynamically to one root and to three shoot compartments (each of which consisting of four age classes). Accumulated aboveground biomass is lost by either cutting or grazing, or enters a litter pool. The N cycle considers three types of inputs to the soil via atmospheric N deposition, fertilizer N addition, and symbiotic N fixation by legumes. The inorganic soil N is available for root uptake and may be lost through leaching, ammonia volatilization and nitrification/denitrification, the latter processes leading to nitrous oxide (N_2O) gas emissions to the atmosphere. Management includes N fertilization, mowing and grazing and can either be set by the user or optimized by the model). The model has been tested at European sites for the ability to simulate forage yields, soil C stocks, soil NH_x and NO_3^- concentrations, CO_2, CH_4, N_2O and NH_3 emissions, energy fluxes and animal performances.

The set of 17 most influential parameters of the model (Table 2) was subjected to Bayesian calibration, using the Metropolis-Hastings MCMC-algorithm to estimate the posterior probability distribution for each parameter. Table 2 gives, for each parameter, the description and the prior information collected from literature review, expert knowledge and previous experimental data. The set of parameters used consists of five vegetation parameters (θ_1, …, θ_5), 10 soil parameters (θ_6, …, θ_{15}) considered as site-specific, and two parameters related to management practices (θ_{16}, θ_{17}). In the absence of information on the form of the PDFs, a uniform distribution with bounds set to the minimum (min) and the maximum (max) of the *a priori* known range of values was taken as a prior distribution of the parameters (Table 2).

Table 1. Output variables (D_i, i=1, …, 5) used for Bayesian calibration.

Variable (D_i)		Unit	Measurement dates	
			Control	Precipitation reduction
D_1	Leaf area index	$m^2\ m^{-2}$	15.08.2008, 29.07.2009	29.07.2009
D_2	Yield	$g\ m^{-2}$	15.08.2008, 29.07.2009	29.07.2009
D_3	Specific leaf area	$m^2\ kg^{-1}$	15.08.2008, 29.07.2009	29.07.2009
D_4	Soil moisture	$m^3\ m^{-3}$	2008-2009 (daily measurements)	2008-2009 (daily measurements)
D_5	Soil temperature	K	2008-2009 (daily measurements)	2008-2009 (daily measurements)

Table 2. PaSim parameters (θ_i, i=1, …, 17) used for Bayesian calibration.

Parameter				Default value	Prior probability distribution	
θ_i	Symbol	Description	Unit		θ_i min	θ_i max
θ_1	TypeA*	Fraction of grassland functional type with high specific leaf area, high digestibility, short leaf lifespan and early reproductive growth	-	0.1	0.01	1
θ_2	TypeB*	Fraction of grassland functional type with medium specific leaf area and high digestibility, long leaf lifespan and medium-to-late reproductive growth	-	0.1	0.01	1
θ_3	TypeC*	Fraction of grassland functional type with low specific leaf area, medium digestibility, long leaf lifespan and medium-to-late reproductive growth	-	0.1	0.01	1
θ_4	TypeD*	Fraction of grassland functional type with low specific leaf area, low digestibility, very long leaf lifespan and very late reproductive growth	-	0.95	0.01	1
θ_5	slaMax	Maximum specific leaf area	$m^2\ kg^{-1}$	28	21	35
θ_6	zsb	Soil depth	mm	700	500	900
θ_7	thetassatb	Saturated volumetric soil water content	$m^3\ m^{-3}$	0.6	0.4	0.7
θ_8	thetasfc	Field capacity fraction of the saturated soil water content	-	0.56	0.5	0.9
θ_9	thetasfpwp	Permanent wilting point fraction of the saturated soil water content point	-	0.4	0.2	0.49
θ_{10}	psieb	Air-entry water potential	mm	-228	-20	-400

θ_{11}	bb	Campbell's *b* parameter	-	7.9	2	14
θ_{12}	ksatb	Saturated hydraulic conductivity	mm d^{-1}	165	20	400
θ_{13}	b_root_profile	Parameter of the relative root dry matter distribution in different soil layers	-	10	4	20
θ_{14}	thetasbspring	Water content of the soil boundary layer in spring	m^3 m^{-3}	0.35	0.15	0.3
θ_{15}	thetasbautomn	Water content of the soil boundary layer in autumn	m^3 m^{-3}	0.35	0.2	0.39
θ_{16}	wshtotcutinit	Shoot dry matter after cutting	kg m^{-2}	0.15	0.09	0.25
θ_{17}	lcutinit	Leaf area index after cutting	m^2 m^{-2}	0.4	0.3	1.7

*FA= TypeA/(TypeA + TypeB + TypeC + TypeD)

The preparation of data files, the generation and analysis of the Markov chains and the uncertainty analysis (including graphical provisions) were carried out with the R language and environment for statistical computing (http://www.r-project.org). During the calibration, PaSim was invoked and launched by the R-code for calibration. The chains of calibration were running on high computing cluster with 22 nodes dual processor AMD Opteron Quadcore, with 12 GB of RAM per machine under CentOS 5.5. The server used to compute 64-bit parallel or sequential jobs is at the Laboratory of Computer Modeling and Optimization Systems (LIMOS, http://limos.isima.fr) in Clermont-Ferrand (France).

The models' predictive ability resulting from prior and posterior parameters was evaluated using the root mean square error (RMSE) and the coefficient of residual mass (CRM), both based on the differences between the measured (O_i) and the simulated (P_i) values (n = number of samples). RMSE measures the size of the differences between simulated and measured values, and varies between zero (agreement between estimates and data) and positive infinity. CRM shows the tendency of the model to under- (positive values) or over-estimate (negative values) observations, zero being the optimal value.

Results

Table 3 contrasts the prior and posterior means and coefficients of variation (CV) of the most influential parameters of PaSim for C treatment. Except for psieb (Air-entry water potential), smaller CV values of posterior parameter distributions show that incorporation of more precise likelihood information reduced uncertainty. For instance, the CV value of posterior probability for the soil depth (zsb) was nearly fully reduced, and the results were similar for the saturated volumetric soil water content (thetassatb), the maximum specific leaf area (slaMax), and the parameter of shoot dry matter after cutting (wshtotcutinit).

Table 3. Prior and posterior distributions for a sample of 13 parameters and the changes in CV values.

Model parameter	Prior		Posterior		% in CV change
	Mean	CV*	Mean	CV	
TypeA	0,18	0,62	0,92	0,07	-88,68
TypeB	0,10	1,60	0,40	0,07	-95,31
TypeC	0,10	1,58	0,01	0,30	-81,11
zsb	697	0,10	897	~0,00	-96,65
thetassatb	0,60	0,08	0,51	~0,00	-95,68
thetasfc	0,56	0,12	0,73	0,01	-93,54
thetasfpwp	0,40	0,12	0,23	0,10	-12,77
psieb	-231	0,27	-251	0,34	29,16
ksatb	165	0,38	245	0,34	-12,82

b_root_profile	9,88	0,27	13	0,24	-11,48
wshtotcutinit	0,15	0,19	0,09	~0,00	-98,49
lcutinit	0,40	0,58	1,55	0,02	-97,36
slaMax	28	0,09	35	~0,00	-99,56

*The coefficient of variation (CV) is the ratio of the standard deviation to the arithmetic mean. The sample size is 1000.

Examples of output variables are provided as resulting from Bayesian calibration. Figure 1 compares the uncertainties of prior and posterior estimates of soil moisture, with precipitation reduction, at 0.05 m depth. Despite the lack of data in some periods, the posterior curve fits well the observations and the uncertainty obtained in the Bayesian process was much reduced compared to the prior uncertainty. Similar results were obtained when soil moisture at 0.10 m depth was simulated (Figure 2). In that case, some inaccuracies are due to soil moisture initialization, indeed a certain time is needed before the model reach the average value the time needed by the model to stabilize its states at the beginning of the simulation.

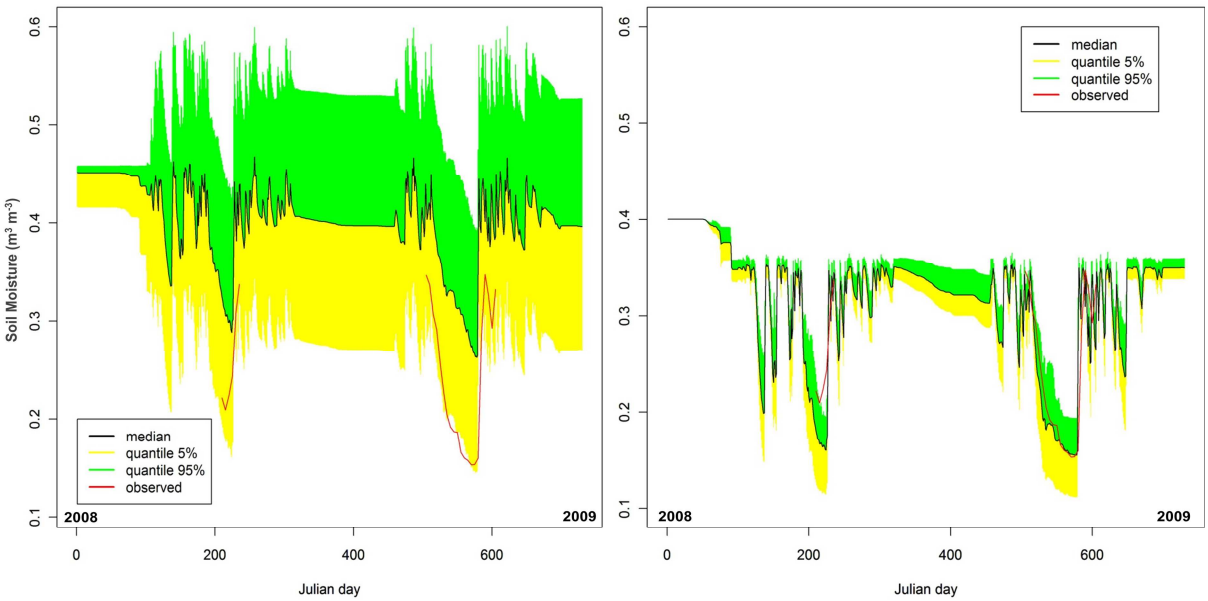

Figure 1. Uncertainty analysis for prior (left graph) and posterior (right graph) estimates of soil moisture at 0.05 m depth with precipitation reduction.

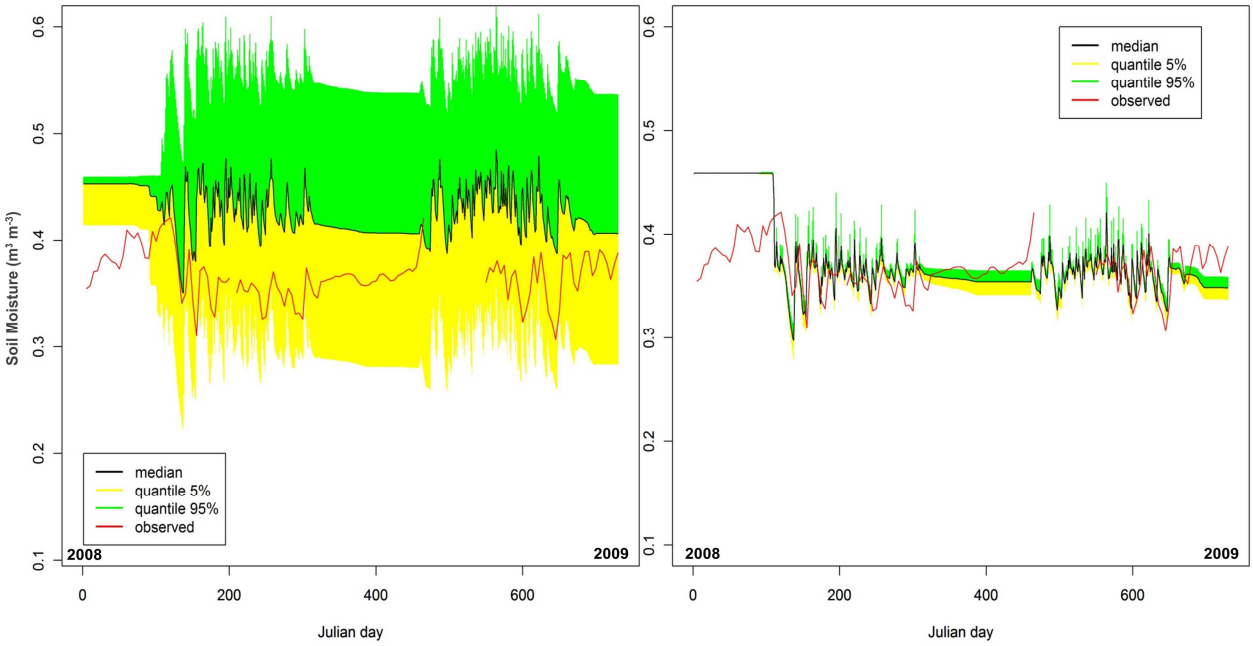

Figure 2. Uncertainty analysis for prior (left graph) and posterior (right graph) estimates of soil moisture at 0.10 m depth with precipitation reduction.

For leaf area index (Figure 3), both control and treatment showed a global improvement in the posterior probability estimates, though uncertainty (bars of standard deviation) was not reduced with the latter.

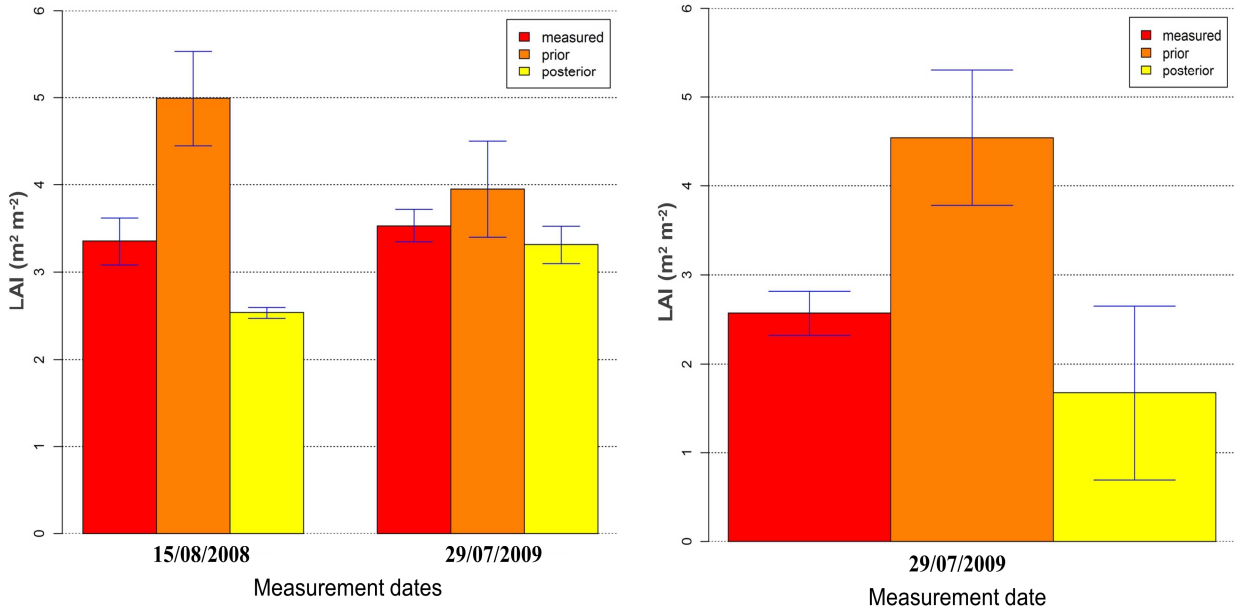

Figure 3. Uncertainty analysis for prior and posterior estimates of leaf area index without (left graph) and with (right graph) precipitation reduction.

For the yield in control (Figure 4), Bayesian calibration reduced uncertainty in 2008 with lower accuracy, while it improved the accuracy in 2009 without reducing uncertainty. However, the only few data of LAI and yield available do not allow deriving fair conclusions.

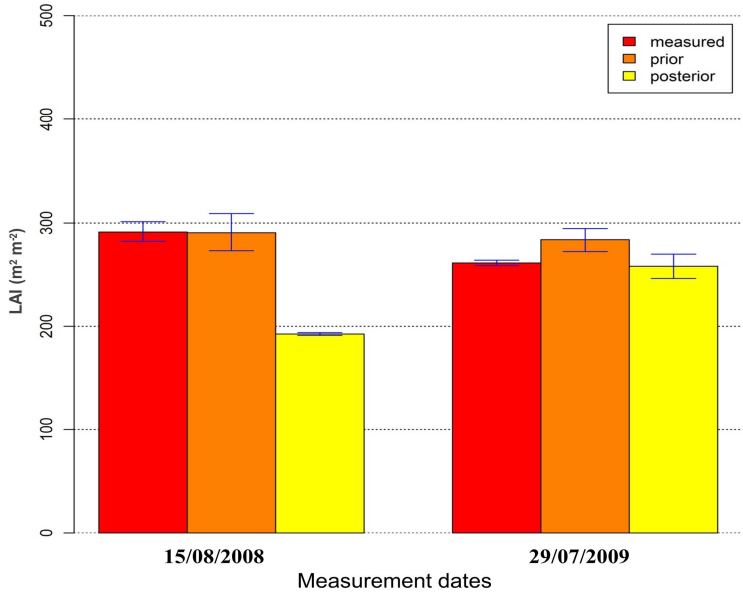

Figure 4. Uncertainty analysis for prior and posterior estimates of simulated yield without precipitation reduction.

Table 4 summarizes the RMSE and CRM values of four output variables. For posterior probability estimates, the two indices were closer to zero than the prior probability for LAI and soil moisture in both treatments. For specific leaf area, improvements were only observed with precipitation reduction. No improvement was observed on yield estimates, for which the data were not enough to make a conclusion. Performance indices for the prior and posterior estimates of soil temperature, not shown here, were very similar.

Table 4. Evaluation of prior and posterior probability estimates of four output variables obtained with and without (Control) precipitation reduction. RMSE, root mean square error; CRM, coefficient of residual mass.

Treatment	Information	Leaf area index (m^2 m^{-2})		Yield (g m^{-2})		Specific leaf area (m^2 kg^{-1})		Average soil moisture (0.05 m - 0.10 m) (m^3 m^{-3})	
		RMSE (m^2 m^{-2})	CRM	RMSE (g m^{-2})	CRM	RMSE (m^2 kg^{-1})	CRM	RMSE (m^3 m^{-3})	CRM
Control	Prior	1.20	-0.30	15.77	-0.03	1.68	-0.07	0.08	-0.23
	Posterior	0.59	0.15	22.93	0.18	4.34	-0.24	0.03	-0.01
Precipitation reduction	Prior	1.97	-0.76	5.57	-0.02	11.13	-0.71	0.07	-0.17
	Posterior	0.89	0.35	16.10	0.68	7.73	-0.49	0.03	0.06

Conclusions

The above results show how a complex grassland model (PaSim) can work within a Bayesian framework to (try to) answer the emerging issues of climate change research. The reduced uncertainty in posterior model outputs indicates the potential of Bayesian calibration to reduce uncertainty under conditions of altered climate (represented here by precipitation reduction). Similar results (Ben-Touhami *et al.*, 2012b) were also obtained at two grassland sites in Switzerland (Frübüel, 47° 06′ North, 08° 32′ East, 982 m a.s.l.; Chamau, 47° 12′ North, 08° 24′ East, 393 m a.s.l.), on which the same experimental protocol was applied for precipitation reduction.

The above scheme (Bayesian calibration applied at single sites) is helpful to reduce uncertainty under specific conditions. However, it is unsuitable to update model parameter values for use over large areas (*e.g.* region to continent). For this, model calibration over

multiple sites representative of weather and management conditions of the study-area is of interest. An important question is about the choice of design points because physical observation sites are often limited. The example below shows the Bayesian calibration of Pasim, as run over multiple sites to provide a new parameterization for Europe.

6. Example 2: long-term observational grassland sites in Europe

Data and model

Based on Ben Touhami *et al.* (2012a), the Pasture Simulation model (PaSim) described in Section 5 was used (Ben Touhami *et al.*, 2013) to update prior parameter distribution to achieve a posterior distribution (30,000 iterations with acceptance rate of about 30%) by incorporating the information contained in the measured data of seven multi-year observational grassland sites in Europe (Amplero, Italy; Bugac, Hungary; Easter-Bush, United Kingdom; Früebüel and Oensingen, Switzerland; Laqueuille intensive and extensive plots, France) mainly derived from the FLUXNET network (http://fluxnet.ornl.gov).

The nine most relevant PaSim vegetation parameters (chosen from Europe-wide sensitivity analysis, data not shown) were calibrated using a set of soil (temperature, water content), vegetation (leaf area index, harvested biomass) and atmospheric (NEE) measured variables. The calibrated model was used to assess CO_2 (NEE, g C m^{-2} d^{-1}) and CH_4 (g CH_4-C m^{-2} d^{-1}) fluxes based on the eddy covariance measurements (in place since 2002) of Laqueuille in France (45° 38' N, 02° 44' E, 1040 m a.s.l.). Two paddocks were continuously grazed by heifers from May to October with two management options (Klumpp *et al.*, 2011): the intensive management paddock included significant amounts of N fertilization (three times per year for a total of ~200 kg N ha^{-1}) and annual average stocking rate of 1.1 LSU ha^{-1}; the extensive management paddock had no fertilization and 0.6 LSU ha^{-1}.

Results

The improvement of simulation after parameters calibration is reflected in the posterior estimates (thanks to maximum likelihood) of NEE and CH_4 daily values, which are closer to observations than using the prior distribution. For NEE from multiple years (2004-2008), regression lines (Figure 5) show the improvement obtained with posterior parameter values (higher R^2; slope and intercept closer to 1 and 0, respectively), with no difference between managements. Daily CH_4 observed values were limited to May-October 2010 in the intensive system.

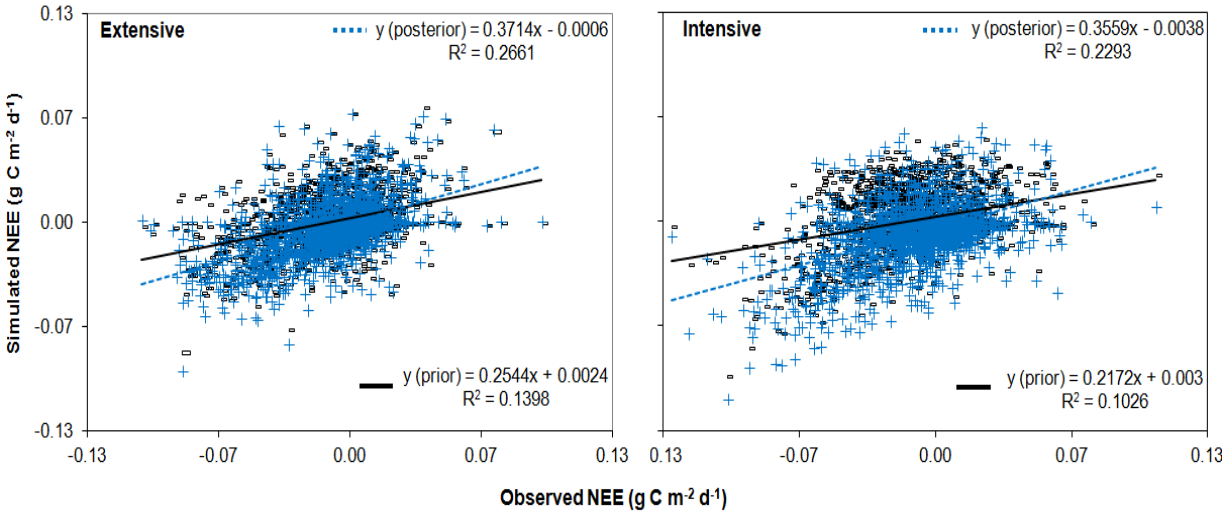

Figure 5. Scatterplots of simulated (prior [-] and posterior [+]) and observed NEE values (g C m^{-2} d^{-1}) at Laqueuille (2004-2008), with regression lines for extensive and intensive management.

Figure 6 shows the improvement obtained with the posterior parameterization but also that the model is not properly simulating the fluctuation in CH_4 values. It is noteworthy that, with posterior simulation, the system emits enteric CH_4 fluxes in summer because enough grass biomass is available and grazing may occur. This approximates what happens in reality, which is not the case with prior parameterization.

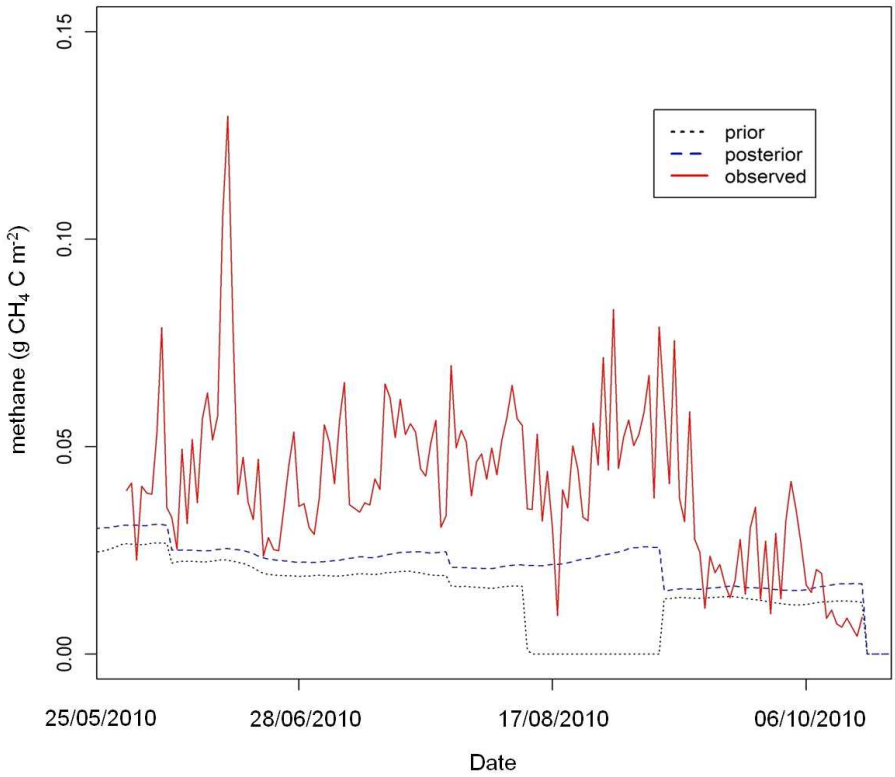

Figure 6. Simulated (prior and posterior) and observed CH_4 values (g C m^{-2} d^{-1}) at Laqueuille (intensive management) over May-October 2010.

Conclusions

These results show that the parameterization of PaSim obtained via Bayesian calibration at multiple European sites has improved simulation of both CO_2 (NEE) and enteric CH_4 fluxes at Laqueuille (France), though without compensating for limitations in the model structure. This means that the modelling of GHG fluxes from grasslands in Europe merits further investigation. This is a non-trivial task, not only because of unsolved theoretical questions but also because fluxes are affected by large observational uncertainties.

Work is ongoing to evaluate posterior estimates of an array of model outputs at different European sites.

7. Implementation

Parameterization and evaluation of model uncertainty in AnimalChange will be grounded to the general methodology for Bayesian calibration illustrated in this deliverable.
An exemplary code excerpt for implementing the Bayesian calibration is given below (Annex I) as employed to calibrate the grassland model PaSim by using the R language and environment for statistical computing (http://www.r-project.org). This is merely illustrative of the steps the user must take to generate posterior samples of the parameters. More details are provided with the proceedings of MS62 ("A training session on Bayesian methods").

8. Annex I – Bayesian calibration algorithm

Required: (1) A model: M

(2) Data: D

(3) A prior probability distribution for the parameters: $P(\theta)$

(4) A likelihood function: $P(D|\theta)$ which is a function of the difference $M(\theta)$-D

Implementation: (1) Select any starting point in parameter space, $\theta^{(0)}$

(2) Calculate the product of prior and likelihood: $\pi^{(0)}=P(\theta^{(0)}).P(D|\theta^{(0)})$

(3) Set i=0 (number of accepted candidates)

(4) Run the following loop until we have enough sample points:

(i) Generate a random new point, $\theta^{(candidate)}$

(ii) Calculate $\pi^{(candidate)} = P(\theta^{(candidate)}).P(D|\theta^{(candidate)})$

(iii) Calculate the Metropolis ratio: $\pi^{(candidate)} / \pi^{(i)}$

(iv) Accept the candidate with probability equal to *min*(Metropolis ratio,1)

(v) If the candidate is accepted: $\theta^{(i+1)} = \theta^{(candidate)}$, else: $\theta^{(i+1)} = \theta^{(i)}$

(vi) Set i=i+1

Result: (1) A representative sample $\{\theta^{(0)} \ldots \theta^{(n)}\}$ from the posterior pdf, $P(\theta|D)$

Implementation code in R

```
source("Preliminaries_rev2.R") # contain necessary data (lai, biomass..) and path for other data (tsoil, SWC, NEE..)

load("init.Rdata") # initial conditions

#*******************************************************************************************

# Define

# - the length of the MCMC chain (nChain) and burn-in (nBI : number of iterations to stabilize the model);

# - various matrices for storing the candidates (pChain);

# - the accepted states (pChAcc) with their index in the chain (jChAcc);

# - the mean vector (avePar) and covariance matrix (covPar).

#--npar : number of parameters

 nChain <- as.integer(100000)

 nBI    <- as.integer(2000)

 pChain <- matrix(0, nrow=nChain, ncol=npar)

 pChAcc <- matrix(0, nrow=nChain, ncol=npar)

 jChAcc <- matrix(0, nrow=nChain, ncol=npar)

covPar <- matrix(0, nrow=npar, ncol=npar)
```

```
#*******************************************************************************************

# Start the chain using uniform random values in [parmin, parmax].

pValues <- runif(npar, min=parmin, max=parmax)


 Param(1)         = pValues[1]

 Param(2)         = pValues[2]

...

    #modification of files before launching pasim

      det_param_vegetation()

# First call to pasim (to get outputs).

 pasim <- lance_pasim()


#-----

# Calculate corresponding prior probability and likelihood

 logPrior0  <- sum( dunif(pValues, min=parmin, max=parmax, log=TRUE) )

   logL0     <- calcul_logL(pasim)

#calcul_logL a function giving Likelihood:    <- sum( -0.5*((simulated - observed)/standard deviation (obs))^2 - 0.5*log(2.*pi) - log(standard deviation (obs)))


 psetMAP     <- pValues

 logMAP      <- logPrior0 + logL0

 pChain[1,] <- pValues


#*******************************************************************************************

# Define Variance-covariance matrix (vcovProp) for proposal generation an

# a scale factor (scalProp) according to Brooks and Gelman (1997).

#
```

```
# Note that by definition of vcovProp parameters are assumed to be at first independent.

#-----

 vcovProp <- diag( (0.01*(parmax - parmin)/parmod)^2 ) #construct a diagonal matrix

 scalProp <- 2.4^2/npar   #   This is the scaling factor proposed by

                          #   Brooks and Gelman (1997), it will be multiplied by the matrix


#*******************************************************************************************

# Build up the chain. Candidates for the parameter values (candidatepValues)

# are assumed to stem from a multivariate normal distribution (mvrnorm) with mean

# at the current state and covariance given by scalProp*covPar. If the candidates

# are outside [parmin, parmax], then candidates within this range are sampled using reflection at
parmin or parmax.

 covPar <- vcovProp

 kc     <- 0   # number of accepted parameters candidate


 for (j in 2:nChain)

   {

    candidatepValues  <- mvrnorm(n=1, pValues, scalProp*covPar, tol=1e-6, empirical=FALSE)

    reflectionFromMin <- pmin(0.,candidatepValues-parmin)

    reflectionFromMax <- pmax(0.,candidatepValues-parmax)

    candidatepValues  <- candidatepValues - 2.*reflectionFromMin - 2.*reflectionFromMax


#-----

#-- Call the model with the candidate values, and check acceptance criterion.


Param(1)         = candidatepValues [1]

Param(2)         = candidatepValues [2]

…
```

```
    # Modification of files before launching pasim
      det_param_vegetation()


    pasim <- lance_pasim()


#-----

#--  Update prior probabilities and likelihood


    logPrior1 <- sum( dunif(candidatepValues, min=parmin/parmod, max=parmax/parmod, log=T) )


#-----

#-- Check whether the candidates are accepted. If yes and if burn-in (2000 first iterations to stabilize the model) has been completed, then store corresponding values and positions within the chain in pChAcc and jChAcc.


    logalpha <- logPrior1 + logL1 - (logPrior0 + logL0)


    if (log(runif(1,0,1)) < logalpha)
     {
      pValues   <- candidatepValues
      logPrior0 <- logPrior1
      logL0     <- logL1


      if (j > nBI)
       {
        if ((logPrior0 + logL0) > logMAP)
         {
```

```
        logMAP  <- logPrior0 + logL0
        psetMAP <- pValues
       }
      kc <- kc + 1
      jChAcc[kc,] <- j
      pChAcc[kc,] <- pValues
     }
   }
pChain[j,] <- pValues
```

9. References

Aitchison, J., Dunsmore, I.R. (1975) Statistical prediction analysis. Cambridge University Press, Cambridge.

Bayes, T. (1763) An essay towards solving a problem in the doctrine of chances. Phil. Trans. R. Soc. Lond., 53, 370-418.

Ben-Touhami, H., Lardy, R., Bahn, M., Bellocchi, G., Barra, V. (2012a) Bayesian calibration of the Pasture Simulation Model (PaSim) to simulate European grasslands under climate extremes: case study at Stubai (Austria). Proceedings of the 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012 (PSAM11 ESREL 2012), 25-29 June, Helsinki, Finland, 567-576

Ben-Touhami, H., Lardy, R., Gilgen, A.K., Buchman, N., Barra, V., Bellocchi, G. (2012a) Bayesian calibration of the Pasture Simulation Model (PaSim) to simulate Swiss grasslands under climate extremes. Proceedings of the 11th European Society for Agronomy Congress, 20-24 August, Helsinki, Finland, 441-5.

Ben-Touhami, H., Lardy, R., Klumpp, K., Barra, V., Bellocchi, G. (2013) Bayesian calibration of the Pasture Simulation model (PaSim) to simulate emissions from long-term European grassland sites: a case study at Laqueuille (France). Proceedings of Greenhouse Gases & Animal Agriculture Conference (GGAA2013), 23-26 Juin, Dublin, Ireland (in press).

Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., Acutis, M. (2010) Comparison of sensitivity analysis techniques: a case study with the rice model WARM. Ecol. Model., 16, 1897-1906.

Gabrielle, B., Laville, P., Duval, O., Nicoullaud, B., Germon, J.C., Hénault, C. (2006) Process-based modeling of nitrous oxide emissions from wheat-cropped soils at the subregional scale. Global Biogeochem. Cycles, 20, GB4018.

Gallagher, M., Doherty, J. (2007) Parameter estimation and uncertainty analysis for a watershed model. Environ. Modell. Softw., 22, 1000-1020.

Brooks, S.P., Gelman, A. (1997) General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat., 7, 434-455.

Haario, H., Saksman, E., Tamminen, J. (2001) An adaptive Metropolis algorithm. Bernoulli, 7, 223-242.

Helton, J.C., 1993. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliab. Engng. Syst. Safty., 42, 327-367.

Johnson, I.R., Chapman, D.F., Snow, V.O., Eckard, R.J., Parsons, A.J., Lambert, M.G., Cullen, B.R. (2008) DairyMod and EcoMod: biophysical pastoral simulation models for Australia and New Zealand. Aust. J. Exp. Agric., 48, 621-631.

Kennedy, M.C., O'Hagan, A. (2001) Bayesian calibration of computer models. J.R.Stat.Soc. B, 63, 425-464.

Klumpp, K., Tallec, T., Guix, N., Soussana, J.-F. (2011) Long-term impacts of agricultural practices and climatic variability on carbon storage in a permanent pasture. Global Change Biol., 17, 3534-3545.

Makowski, D., Hillier, J., Wallach, D., Andrieu, B., Jeuffroy, M.H. (2006) Parameter estimation for crop models. In: Wallach, D., Makowski, D., Jones, J.W. (Eds.), Working with dynamic crop models. Evaluation, analysis, parameterization and applications. Amsterdam: Elsevier, 101-149.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. (1953) Equation of state calculations by fast computing machines. J. Chem. Phys., 21, 1087-1092.

Morris, M.D. (1991) Factorial sampling plans for preliminary computational experiments. Technometrics, 33, 161-174.

Olesen, J.E., Carter, T.R., Díaz-Ambrona, C.H. Fronzek, S., Heidmann, T., Hickler, T., Holt, T., Minguez, M.I., Morales, P., Palutikof, J., Quemada, M., Ruiz-Ramos, M., Rubæk, G., Sau, F., Smith, B., Sykes, M. (2007) Uncertainties in projected impacts of climate change on European agriculture and terrestrial ecosystems based on scenarios from regional climate models. Climatic Change, 81, 123-143.

Patenaude, G., Milne, R., van Oijen, M. (2008) Integrating remote sensing datasets into ecological modelling: a Bayesian approach. Int. J. Remote Sens., 29, 1295-1315.

Romanowicz, R., Beven, K., Tawn, J.A. (1994) Evaluation of predictive uncertainty in nonlinear hydrological models using a Bayesian approach. In: Barnett, V., Turkman, K.F. (Eds.), Statistics for the environment 2: water related issues. New York: Wiley, 297-319.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. (2004) Sensitivity analysis in practice. Wiley, New York.

Sivia, D.S. (1996) Data analysis: a Bayesian tutorial. Oxford University Press, Oxford.

Soussana, J.-F., Graux, A.-I., Tubiello, F.N. (2010) Improving the use of modelling for projections of climate change impacts on crops and pastures. J. Exp. Bot., 61, 2217-2228.

Tubiello, F.N., Soussana, J.-F., Howden, S.M. (2007) Crop and pasture responses to climate change. P. Natl. Acad. Sci. U.S.A., 104, 19686-19690.

Van Oijen, M., Rougier, J., Smith, R. (2005) Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiol., 25, 915-927.

Van Oijen, M., Thomson, A. (2010) 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, 103, 55-67.

Wallach, D., Buis, S., Lecharpentier, P., Bourges, J., Clastre, P., Launay, M., Bergez, J.-E., Guérif, M., Soudais, J., Justes, E. (2011) A package of parameter estimation methods and implementation for the STICS crop-soil model. Environ. Modell. Softw., 26, 1-9.