Published paper
Kennedy, M.C., O'Hagan, A., Anderson, C.W., Lomas, M., Woodward, F.I., Heinemeyer, A. and Gosling, J.P. (2008) *Quantifying uncertainty in the biospheric carbon flux for England and Wales.* Journal of the Royal Statistical Society: Series A (Statistics in Society), 171 (1). pp. 109-135.
http://dx.doi.org/10.1111/j.1467-985X.2007.00489.x

Quantifying Uncertainty in the Biospheric Carbon Flux for England and Wales

Marc Kennedy, Clive Anderson, Anthony O'Hagan*, Mark Lomas, Ian Woodward
University of Sheffield, UK
Andreas Heinemeyer
University of York, UK

3rd April, 2006

Abstract

A crucial issue in the current global warming debate is the effect of vegetation and soils on carbon dioxide (CO_2) concentrations in the atmosphere. Vegetation can extract CO_2 through photosynthesis, but respiration, decay of soil organic matter, and disturbance effects such as fire return it to the atmosphere. The balance of these processes is the net carbon flux. In order to estimate the biospheric carbon flux for England and Wales, we address the statistical problem of inference for the sum of multiple outputs from a complex deterministic computer code whose input parameters are uncertain. The code is a process model which simulates the carbon dynamics of vegetation and soils, including the amount of carbon stored as a result of photosynthesis and the amount returned to the atmosphere through respiration. The aggregation of outputs corresponding to multiple sites and vegetation types in a region gives an estimate of the total carbon flux for that region over a period of time. Expert prior opinions are elicited for marginal uncertainty about the relevant input parameters and for correlations of inputs between sites. A Gaussian process model is used to build emulators of the multiple code outputs and Bayesian uncertainty analysis is then used to propagate uncertainty in the input parameters through to uncertainty on the aggregated output. Numerical results are presented for England and Wales in the year 2000. It is estimated that vegetation and soils in England and Wales constituted a net sink of 7.61 MtC (1 MtC = 10^{12}g of carbon) in 2000, with standard deviation 0.61 MtC resulting from the sources of uncertainty considered.

1 Introduction

The continual and accelerating rise in atmospheric CO_2 concentrations, largely as a result of anthropogenic activities, underlies increasing concerns about global warming and climate change. The potentially serious long-term effects of uncontrolled greenhouse gas (GHG) emissions motivated the UN Framework Convention on Climate Change, under which parties to the Convention

[1]Address for correspondence: Tony O'Hagan, Department of Probability and Statistics, University of Sheffield, Hounsfield Road, Sheffield S3 7RH, UK
Email: a.ohagan@sheffield.ac.uk

must take steps to reduce emissions, and in addition must periodically report national inventories of anthropogenic emissions and removals by sinks of all GHGs. The principal terrestrial sink for CO_2 is through its take-up by vegetation or soils, although such biospheric systems can also act as sources of CO_2. Estimating the intensities of such carbon sinks or sources is difficult, and no methodology for it has been agreed by the Intergovernmental Panel on Climate Change. However, whatever estimation method is employed, it is imperative for rational discussion and use that it should be accompanied by a realistic assessment of its uncertainty. An additional consideration is that the Kyoto Protocol allows countries to use new carbon sinks (from afforestation and reforestation) to offset emissions. Estimating the effectiveness of such activities in sequestering carbon is therefore a scientific matter of some importance, and again estimates of uncertainty are needed for any well-founded assessment.

The UK Centre for Terrestrial Carbon Dynamics (CTCD) is a multi-disciplinary consortium of researchers in British academic and governmental institutions, established to address these issues. Its purpose is to advance scientific understanding of the role played by terrestrial ecosystems in the carbon cycle, with particular emphasis on forest ecosystems. The ultimate goals of the Centre are: to estimate carbon fluxes and their uncertainties at different space/time resolutions; to further scientific understanding and, through new methodology, instrument advances and data acquisition (particularly from satellites) to reduce these uncertainties; and to deliver relevant findings in accessible formats to the scientific community and to policy makers. In pursuit of these objectives, a variety of computer models have been created for simulating carbon dynamics over different geographical and climatic scenarios.

Carbon is exchanged in various chemical forms between vegetation, soils and the atmosphere, and provides the basic unit to quantify such exchanges. Existing estimates of UK biospheric carbon fluxes make use of a simple dynamic carbon-accounting model, as described in Milne and Cannell (2005). The primary inputs to this model are forest inventories and annual planting and felling information. Estimates of uncertainty from these approaches are not completely established, the methods used provide limited understanding of the processes operating, and they have no predictive power with respect to changes in climatic drivers and soil properties. These factors have motivated the development of more detailed models that attempt to estimate carbon fluxes by representing the processes that give rise to them. We present here an analysis of the total biospheric carbon flux for England and Wales in the year 2000, based on a dynamic vegetation model incorporating the latest scientific understanding of those processes that drive carbon exchange between vegetation, soils and the atmosphere. The model-based approach is embedded in a statistical framework that facilitates rigorous assessment of the uncertainty of flux estimates, and in particular makes it possible to address such questions as: what are the main sources of uncertainty?; what is the total uncertainty and can it be partitioned among its various components?; and how can uncertainty be reduced? The paper describes the statistical framework and the development of uncertainty estimates for aggregate quantities within it. The discussion of aggregation is a novel technical contribution to Bayesian uncertainty analysis.

This paper specifically addresses model-based estimation of carbon flux as represented by net biome productivity (NBP). NBP is the net uptake of CO_2 by the land surface (vegetation and soils) from the atmosphere, with positive NBP indicating a carbon sink. It can be expressed

as

$$NBP = GPP - \text{plant respiration} - \text{soil respiration} - \text{disturbances},$$

where Gross Primary Production (GPP) is photosynthetic fixation by vegetation, and the disturbances include carbon released by fire, harvesting, etc. Transport effects can also be absorbed into the disturbance term.

1.1 The Sheffield Dynamic Global Vegetation Model

The Sheffield Dynamic Global Vegetation Model (SDGVMd, where 'd' indicates the daily version) is a dynamic vegetation model designed to calculate the carbon fluxes and pools in the biosphere, and especially variations under changing climate (Woodward et al. (1995), Cramer et al. (2001), Woodward and Lomas (2004)). It incorporates modules describing photosynthesis, evapotranspiration, hydrology, nitrogen cycling, disturbances, and soil carbon respiration and storage. The model represents the growth and decay of generic plant functional types (PFTs), rather than individual species, in discrete grid-cells that together can cover large areas, up to the whole land surface.

For the study reported in this paper, which is restricted to England and Wales (henceforth E&W), only four PFTs are needed, namely grassland, which accounts for 42.0% of the land cover, crops (32.1%), deciduous broadleaf trees (DcBl) (10.9%), and evergreen needleleaf trees (EvNl) (3.1%). Common species of DcBl are oak, beech, birch, ash and sycamore, and of EvNl, Sitka spruce, Scots pine and larch. In the land cover map we use, shrubland is represented as DcBl. Each of these PFTs is characterised by a vector of 25 parameters. A fifth non-productive PFT is included to account for urban and barren areas. Every grid-cell in the model contains a proportion of each PFT, determined by the land cover, with the proportions summing to 1. The explicit region occupied by each PFT within a grid-cell is not exploited.

In addition to the PFT parameters, the model needs several types of input data, including:

1. Climate data: for model runs representing the past the climate inputs use interpolated climate records or climate models, while runs into the future rely on predictions from global circulation models. In this study, monthly temperature, precipitation, air humidity and cloudiness for the year 2000 from the CRU/UEA dataset (Mitchell et al. (2004)) were used. Monthly data are downscaled to a daily time-step with a weather generator developed by Forest Research, UK (Evans et al. (2002)), based on an approach similar to that used in other weather generators e.g., Sharpley and Williams (1990). Precipitation is simulated using a Markov chain and a gamma distribution, while humidity and temperature are taken from Gaussian distributions.

2. Soil texture and bulk density (see §3.2).

3. Land cover from the LCM2000 (see §3.4).

A single simulation consists of a spin-up over 600 years to stabilize the various carbon pools, followed by a proper run from 1901 to 2000. We primarily concentrate on NBP and GPP for the year 2000, which is the most recent year for which full climate and land cover data are available to us. NBP is highly dependent on climate and can vary significantly between years,

even changing from a sink to a source at any given pixel. The same map of NBP could therefore look very different for a different year. The aggregate NBP and maps of GPP are expected to show less variation between years. We concentrate on England and Wales rather than the whole of Great Britain because of restrictions on the use of Scottish soils data.

Kennedy et al. (2005) applied Bayesian sensitivity analysis to SDGVMd during development of the computer code. In the present paper we use the same sensitivity analysis tools to identify the parameters that are important in generating uncertainty in model outputs. We in fact restrict our attention to a subset of the input variables used by the model, namely the PFT and soil texture parameters, and do not discuss uncertainties from the land cover map or the climate data. The phrase 'input parameters' hereafter refers to this subset.

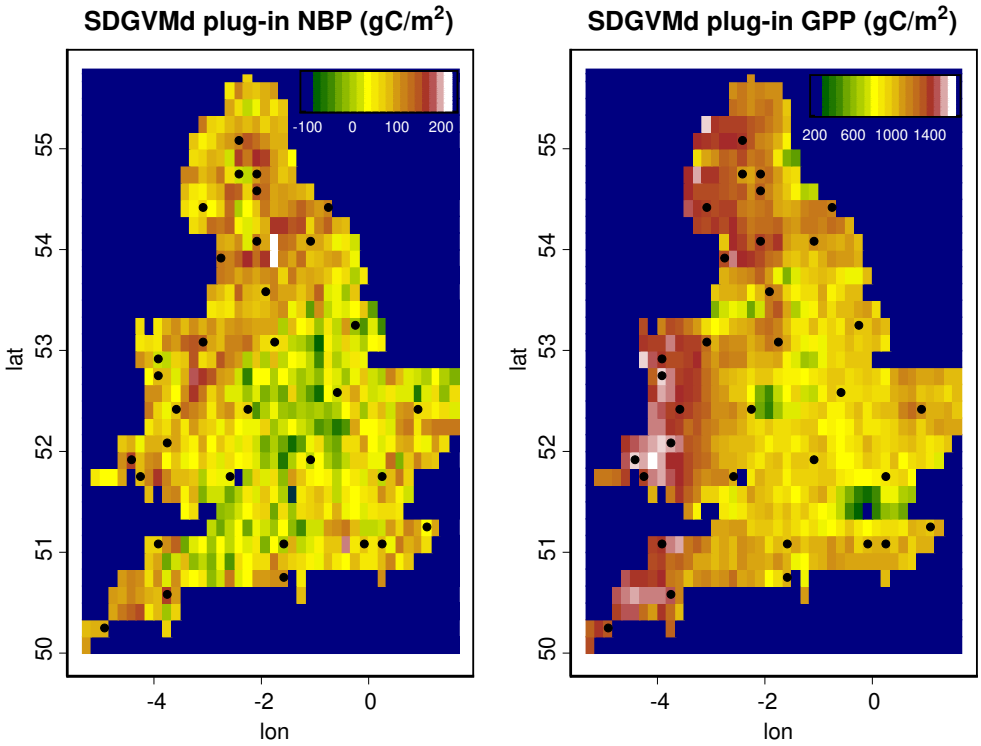

Figure 1: NBP and GPP plug-in estimates (gC/m^2) for England and Wales in the year 2000, using SDGVMd with fixed inputs. The dots indicate 33 sample sites at which uncertainty in these estimates will be assessed in detail.

The maps shown in Figure 1 are calculations by SDGVMd at 707 grid-cells, at 1/6th of a degree resolution. The maps show estimated NBP and GPP over E&W, in gCm^{-2}, for the year 2000. Each map is actually a weighted sum of four maps in which the weights are the proportions of land covered by the four relevant PFTs. In these calculations, the input parameters were fixed at their mean values (based on expert elicitation, described in §3.1), so that we refer to them as plug-in maps. In the following sections we use Bayesian uncertainty analysis to address the deficiencies in these maps that arise from ignoring uncertainty in the input parameters.

The analysis has two parts. First, uncertainty analysis of SDGVMd is used to derive the maps of mean NBP and GPP and to create maps of the associated uncertainties. This involves aggregation, at each site, over the 4 PFTs. Second, we calculate the uncertainty in the aggregate NBP for E&W as a whole. Both parts involve aggregation of model runs and make use of new uncertainty analysis methodology, which is described in §2. Note that we do not consider aggregation of GPP across E&W because, although the GPP map provides useful information about the main component of NBP at each site, the aggregate of GPP is of little scientific interest.

1.2 Uncertainty in the carbon flux

Some of the underlying physical and biological processes represented in SDGVMd are poorly understood and exact values for the parameters describing them are unknown. Measurement data are available from the literature for some vegetation parameters, but these are usually species-specific, and may not provide the most representative value for a more generalised PFT. Other parameters are either very difficult to measure or are tuning parameters which cannot be measured. Furthermore, any input value at a single grid-cell actually represents a large area so there is uncertainty due to upscaling from fine to coarse scale. The consequences of these uncertainties must be properly accounted for in any assessments based on the model outputs. Crucially, results from computer modelling exercises often neglect these uncertainties and treat the input estimates as if they were known.

A major deliverable of CTCD, to which this work contributes, is an estimate of the total UK carbon flux and its uncertainty. The wider issues involved in calculating this flux, and the various sources of uncertainty are described in detail in Quegan (2006). In the present paper, for the reasons given earlier, attention is first concentrated on producing maps of both the best estimate of the E&W NBP and GPP and their uncertainties, given the uncertainties arising from input parameters. From these maps we then derive inference about the total NBP. Spatial variations in the model output are due to differences in vegetation cover, soil and climatic conditions. Data for each of these three factors are used to drive the model. The input values for any set of grid-cells are clearly correlated, and these correlations induce correlations between the outputs for individual grid-cells. Similarly, model outputs of different PFTs at a site are correlated because they share the same climate and soil drivers. This dependence needs to be allowed for in any uncertainty analysis of the aggregate NBP.

Estimating aggregate NBP simply involves summing outputs for all relevant grid-cells. Each PFT is treated separately in the estimation (though over longer timescales they interact within the model itself), so that we can create maps of NBP and calculate the individual aggregate values for each of the PFTs, then combine the four aggregates to give the overall aggregate NBP, weighting according to land cover. This separate treatment of PFTs makes it straightforward to partition the sources of uncertainty and it also provides a route to account for uncertainty in the PFT proportions at each grid-cell (see Quaife (2006)).

Data for soil texture and bulk densities were obtained from the National Soil Research Institute via DEFRA (2003). The original data were measured on a 5 km grid, then a process involving interpolation and expert judgements was used to produce soil texture and bulk density information at 1 km resolution.

The first step in the analysis is to build a set of statistical emulators of the model outputs at individual sites as described in Appendix A. The emulators are used in a sensitivity analysis to identify the most influential inputs, and Bayesian uncertainty analysis is then applied to account for and quantify uncertainty in the output at each site, which derives from uncertainty in the influential model parameters defining vegetation and soil properties.

In the following section we introduce new methodology for inference about a sum of model outputs. Results for E&W carbon maps and total carbon flux are then given in §3. We conclude with discussion of these results and possible extensions in §4.

2 Bayesian uncertainty analysis for aggregates of model outputs

Recent work in the statistical analysis of computer models has concentrated on the case in which a model produces a single output (e.g. Sacks et al. (1989), Kennedy and O'Hagan (2001)). In practice most complex computer codes produce multiple outputs. Each output can of course be analysed individually using single output methods, and for many purposes this is sufficient. Often, however, inference is required about a function of multiple outputs: the sum of outputs is an important example. More generally consider the linear combination

$$F = \sum_{j=1}^{N} \gamma_j f_j(\boldsymbol{X}_j) \tag{1}$$

where each f_j is a single output of a code run with input vector $\boldsymbol{X}_j$. The $f_j(\cdot)$ may be distinct outputs from a multi-output code or even different codes, which could take distinct input vectors of varying dimensions. We formally recognise that the true values of the input vectors $\boldsymbol{X}_j$ are uncertain. In keeping with widespread convention we use upper case $\boldsymbol{X}_j$ to signify this uncertainty. The upper case F denotes that uncertainty in the $\boldsymbol{X}_j$ makes the aggregated output also uncertain. The question at issue for the uncertainty analysis of F is the distribution of F induced by uncertainty in the code input values.

For the E&W carbon flux, we need to sum contributions to NBP or GPP from different PFTs at a given site, and also to sum NBP over sites to obtain the aggregate for E&W. Therefore two variations of the sum (1) are considered.

Aggregation over plant functional types. A weighted sum (1) is required at each site, including a component of NBP for each of 4 vegetation types. The weight γ_j is the proportion of land covered by type j. Each PFT uses a different form of SDGVMd, so the f_j take different sets of inputs. This aggregation is described in more detail in §3.4.

Aggregation over sites. Each term represents NBP output at distinct sites or grid-cells for a single PFT, with weights γ_j representing the area (in m^2) of grid-cell j covered by that PFT. In this case we also assume that each $\boldsymbol{X}_j$ is $p-$dimensional, so that each code output has the same number of inputs. For a given PFT, SDGVMd can be run with a single input $\boldsymbol{x}$ at multiple sites. This multi-output code can therefore be represented as a vector function $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_N(\boldsymbol{x}))$ at, say, N sites. If exactly the same inputs were appropriate at each site, then $\boldsymbol{X}_1 = \boldsymbol{X}_2 = \ldots = \boldsymbol{X}_N = \boldsymbol{X}$ and F could be treated by single-output methods by simply reducing the output to a single aggregate value for each test run. In this paper we

consider the more structured case in which $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$ are distinct but correlated. Examples of correlations between SDGVMd inputs are considered in §3.5.

If the code is very fast then straightforward Monte Carlo estimation of the distribution of F may be feasible. This would require running the code for thousands of simulated input points, for all PFTs and sites, to create a sample of F values. Very often, however, (as, for instance, in the case of SDGVMd), the complexity of the code rules out such an approach, as a single run may take several minutes or even hours and it becomes too costly to generate enough runs for reliable estimation. A statistical representation of the code, based on a Gaussian process model and known as an *emulator*, can be used instead. This provides both a fast approximation for the code and a coherent framework for inference, in which input parameter uncertainty can be propagated through the emulator analytically. The emulator also accounts for the approximation uncertainty introduced by its own use as a surrogate to the code. A brief summary of emulation is given in Appendix A. The development builds upon the earlier work in Bayesian uncertainty analysis of Haylock and O'Hagan (1996) and O'Hagan and Haylock (1997).

We model individual outputs using separate emulators, and consider the covariance between outputs $f_j(\boldsymbol{X}_j)$, $f_{j'}(\boldsymbol{X}_{j'})$ that is induced by covariance between the inputs $\boldsymbol{X}_j$ and $\boldsymbol{X}_{j'}$. In principle, there is a further source of correlation arising from correlation in the emulators. This is addressed in Conti and O'Hagan (2006) by building a multivariate emulator of multiple outputs. However, the multivariate emulator makes stronger assumptions and the correlation induced in this way is small provided the emulation is sufficiently accurate. We therefore treat the emulators here as independent.

The joint distribution of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$, which formally characterises the uncertainty in model inputs, is denoted by G and the marginal distribution of $\boldsymbol{X}_j$ by G_j. Similarly the joint distribution of a pair $\boldsymbol{X}_j, \boldsymbol{X}_{j'}$ is denoted by $G_{jj'}$. Appropriate choices for G, and hence G_j and $G_{jj'}$, are clearly dependent on the application and require detailed discussions with a subject specialist – a vegetation modelling expert in our case. We use mixtures of normal distributions, which allows some important integrals required in the emulation theory to be computed analytically. This mixture class is flexible enough to accommodate a wide range of expert beliefs; indeed, a single component is adequate in practice for most inputs. In our application only one input, leaf lifespan for EvNl, requires a non-degenerate mixture distribution. The form of these distributions, with particular reference to EvNl leaf lifespan, is given in Appendix B.

We are interested in aspects of the distribution of the sum F. The distribution itself is complicated, so we focus on its mean and variance, which are expressed as

$$M = E(F \mid f_1, \ldots, f_N) = \sum_{j=1}^{N} \gamma_j \int f_j(\boldsymbol{x}_j)\, dG_j(\boldsymbol{x}_j), \tag{2}$$

$$V = \mathbb{V}\mathrm{ar}(F \mid f_1, \ldots, f_N) = M_2 - M^2, \tag{3}$$

where

$$\begin{aligned} M_2 = E(F^2 \mid f_1, \ldots, f_N) &= \sum_{j=1}^{N} \gamma_j^2 \int f_j(\boldsymbol{x}_j)^2\, dG_j(\boldsymbol{x}_j) \\ &\quad + 2 \sum_{j=1}^{N-1} \sum_{j'=j+1}^{N} \gamma_j \gamma_{j'} \int f_j(\boldsymbol{x}_j) f_{j'}(\boldsymbol{x}_{j'})\, dG_{jj'}(\boldsymbol{x}_j, \boldsymbol{x}_{j'}). \end{aligned} \tag{4}$$

M and V are features of the distribution of F that we want to compute to quantify the implications of parameter uncertainty. These quantities are themselves uncertain because we cannot compute integrals such as (2) without running the model at all possible input values. Instead, the emulation provides an estimate of each in the form of its posterior mean, and a posterior variance to describe uncertainty due to emulation; these are analogous to the estimates and standard errors that would be produced by the Monte Carlo approach. Computations required for these inferences about M and V, for aggregation over sites, are detailed in Appendix C.

3 Uncertainty analysis of E&W carbon flux

The following notational conventions are used throughout the remainder of the paper. A subscript k is used with variables or distributions associated with site k. A second site is indicated with subscript k' when required, for example if referring to a joint distribution or covariance between site-specific quantities. PFT specific quantities will use subscript t or t'. Input parameter components will use subscript $\_i$. The vector $\boldsymbol{x}$ is used as a generic code input, whereas $\boldsymbol{y}$ is used to represent a site coordinate in the maps.

There are too many grid-cells in the E&W map to emulate each of them as a separate output, since the process of building the emulator requires of the order of 200 model runs per cell and per PFT, so 33 sample sites were selected. As our estimates are based on outputs from a computer model, there were no restrictions on the choice of locations. Sample sites were selected to be representative of the varied E&W climatic conditions, to cover the whole region adequately and to include a wide range of land cover types. The selected sites can be seen in Figure 1. The spatial disposition of sites provided a wide range of inter-site distances that gave information about spatial correlations at different scales. A total of 132 emulators was created, one for each of the four PFTs at each sample site.

An initial sensitivity analysis was carried out for each PFT to determine the most important inputs in terms of contribution to the output variance. The ranges over which to test each input were provided by the plant modelling expert (Professor Ian Woodward, a member of CTCD and the originator of SDGVM). Different sets of important parameters emerged from the sensitivity analysis for each PFT. The significant plant inputs are listed in Table 1. In addition to these, 3 soil inputs were found to be important for all PFTs, as explained in §3.2. Inputs to which the output is not sensitive were given normal distributions in subsequent analysis, with parameters set so that the rough initial ranges represent mean ± 2 standard deviations. Each emulator for the grassland, crop, and DcBl PFTs was built as a function of the input parameters using 200 well chosen code runs. Emulators for EvNl required 300 runs to cover adequately the wider potential range of the leaf lifespan parameter, as explained in §3.1. Maximin Latin Hypercube designs (Morris and Mitchell (1995)) were used to select values of the input parameters for these training runs, using the initial ranges for each input selected by the expert.

3.1 Elicitation of expert opinions for PFT parameters

To elicit marginal prior probability distributions corresponding to the important vegetation input parameters the expert was questioned again, in more detail, on his beliefs about these parameters. The area covered by each 1/6th of a degree grid-cell ranges from about 190 km^2

to 220 km^2, so the parameter value used in SDGVMd should account for the fact that multiple species are likely to be represented in each cell. The expert was asked specifically to provide a possible range of values for individual grid-cell averages. The resulting minimum and maximum values were taken to be mean ± 2 standard deviations, assuming symmetric distributions. The implied mean value was taken as a suitable representative value of the input for use in SDGVMd. The implied variance reflects the fact that the types of species present in a single grid-cell varies across E&W, and it also includes the effect of uncertain knowledge about the overall mean for E&W. The expert agreed that normal distributions provided a fair representation of his uncertainty about these inputs.

The parameters and their elicited distributions are listed for each PFT in Table 1. For evergreen needleleaf species, the *leaf lifespan* parameter required special treatment because leaves can live for several years but are most likely to die during autumn. Suppose that a leaf dies d days into its yth year, so that its full lifespan is $365(y-1)+d$ days. The expert specified independent distributions for y and d, with d distributed as $N(200, 625)$, and with y having a discrete distribution on values $y = 1, 2, 3, 4, 5$.

Table 1: SDGVMd significant plant input parameters for each PFT and elicited probability distributions for their value within a single grid-cell.

PFT	Parameter	Distribution
EvNl	Leaf life span (days)	$0.18N(565, 625) + 0.52N(930, 625) + 0.23N(1295, 625)$ $+0.06N(1660, 625) + 0.01N(2025, 625)$
	Leaf mortality index (proportion)	$N(0.005, 4 \times 10^{-6})$
	Seeding density (individuals/m^2)	$N(0.15, 0.0025)$
DcBl	Leaf life span (days)	$N(190, 1600)$
crop	Budburst limit (degree-days)	$N(140, 56.25)$
	Fraction harvested	$N(0.505, 0.01051)$
	Maximum carbon store (gC/m^2)	$N(200, 2500)$
	Leaf mortality index (proportion)	$N(0.005, 2.25 \times 10^{-6})$
	Length of crop season (days)	$N(190, 1600)$
grassland	Leaf life span (days)	$N(180, 3600)$
	Budburst limit (degree days)	$N(135, 6.25)$
	Maximum carbon store (gC/m^2)	$N(200, 3600)$

3.2 Uncertainty in soil inputs

The important soil parameters were found to be *sand percentage*, *clay percentage* and *bulk density*. These parameters were generally found to be much more influential on the output than the PFT parameters for both DcBl and EvNl, but the opposite is true for grassland and crops, as seen in §3.3. (The reason for their importance is that the soil parameters play a critical role in the mechanisms governing the availability of water to plants, hence are central to growth and

biomass productivity.) The same soil parameters are inputs for each PFT, at a given site, and this implies a correlation between PFTs which must be accounted for in the total NBP at that site. This aggregation over PFTs is described in §3.4.

The following procedure was used to specify the input distributions for these soil parameters. For each of the sample sites, soil data were available at 1 km^2 resolution, giving a number of observations of the relevant parameters for each site. The number of available points varies with latitude and proximity to the coastline. The mean of these observations might be expected to give a good representative value for each parameter to use as input to SDGVMd. However, the soil input values which enable SDGVMd to best estimate vegetation growth and carbon dynamics in a grid-cell are not necessarily the mean values across the grid-cell, because of nonlinearity in the model's response to the soil inputs. The size of any discrepancy due to nonlinearity would be related to the variability of the parameters over the site. Furthermore, the data are not truly independent observations, having been derived by interpolation of raw data on a 5 km grid. Therefore, although we did assign the sample mean as the estimated value of each soil input, we represented uncertainty in each case, not by the sample variance divided by the number of 1 km^2 points but by half the sample variance. Single normal distributions were thought reasonable to describe uncertainty, and independence was assumed between different inputs. Although formally there should be some negative correlation between the percentages of sand and clay, this was judged to be negligible in the light of the small standard deviations on these inputs. The resulting input distributions are listed in Table 2 for each sample site.

3.3 Single Plant Functional Type maps

Bayesian uncertainty analysis produced estimates of the mean and variance of NBP and GPP for each PFT at each of the 33 sample sites. Between-site variations in these values were due to:

- different climatic conditions producing different SDGVMd outputs: these were captured by the differences between the emulators.
- different soil properties, characterised by the site specific probability distributions for soil parameters (Table 2): these variations were captured by the Bayesian uncertainty analyses.

SDGVMd is a highly non-linear function of its inputs. Therefore the expected value of an output with respect to the uncertain input distribution can be substantially different from the 'plug-in' output obtained simply by running the model with the input fixed at its mean value. At each sample site, both expected value of the outputs and the outputs with fixed inputs were calculated, and the differences were used to correct the fixed-input estimates elsewhere. Specifically, maps of the mean correction at the 707 E&W grid-cells were produced by interpolating the 33 differences. We used the kriging software GSLIB (Deutsch and Journel (1997)). The interpolated corrections for NBP are shown in Figure 2 for the four PFTs. Interpolation of these estimates to the remaining E&W grid-cells is reasonable, given that climatic variations over E&W are expected to be represented within the sample sites. A similar interpolation of the log-variances was carried out on the basis of the 33 sample site values of log-variance. We denote by $Em_t(\boldsymbol{y}_k)$ and $Ev_t(\boldsymbol{y}_k)$ the mean and variance of NBP for the tth PFT estimated at

Table 2: SDGVMd soil input parameters and their probability distributions at the sample sites.

Site	(lat, lon)	Soil sand %	Soil clay %	Bulk density (g/cm^2)
1	55.083 -2.417	N(34.12, 1.95)	N(22.82, 3.62)	N(0.814, 0.067)
2	54.75 -2.083	N(36.53, 14.25)	N(23.16, 4.60)	N(0.859, 0.1015)
3	54.75 -2.417	N(37.53, 28.20)	N(23.06, 15.99)	N(0.550, 0.123)
4	54.583 -2.083	N(36.93, 60.38)	N(25.33, 48.34)	N(0.815, 0.1215)
5	54.417 -0.75	N(43.27, 222.12)	N(22.36, 49.21)	N(1.214, 0.0325)
6	54.417 -3.083	N(40.17, 47.39)	N(17.24, 17.45)	N(0.689, 0.0625)
7	54.083 -1.083	N(43.08, 260.75)	N(28.75, 92.57)	N(1.379, 0.001)
8	54.083 -2.083	N(30.55, 63.32)	N(22.17, 3.19)	N(1.171, 0.0175)
9	53.917 -2.75	N(33.51, 33.02)	N(29.85, 27.75)	N(1.367, 0.007)
10	53.583 -1.917	N(50.99, 51.65)	N(18.53, 4.04)	N(0.647, 0.0915)
11	53.25 -0.25	N(29.77, 70.00)	N(36.45, 18.42)	N(1.327, 0.0145)
12	53.083 -1.75	N(22.63, 114.4)	N(25.14, 26.8)	N(1.257, 0.005)
13	53.083 -3.083	N(44.08, 58.18)	N(19.53, 16.00)	N(1.222, 0.021)
14	52.917 -3.917	N(36.18, 24.85)	N(20.20, 10.89)	N(0.684, 0.0705)
15	52.75 -3.917	N(39.97, 36.05)	N(18.2, 13.91)	N(0.819, 0.0400)
16	52.583 -0.583	N(28.88, 72.79)	N(38.21, 43.30)	N(1.275, 0.0030)
17	52.417 0.917	N(69.91, 163.47)	N(13.23, 52.70)	N(1.313, 0.0105)
18	52.417 -2.25	N(60.68, 203.54)	N(13.17, 23.66)	N(1.351, 0.0030)
19	52.417 -3.583	N(35.08, 3.55)	N(22.00, 4.76)	N(1.0172, 0.0145)
20	52.083 -3.75	N(34.81, 7.17)	N(21.59, 4.03)	N(1.001, 0.0105)
21	51.917 -1.083	N(24.47, 79.12)	N(40.77, 65.31)	N(1.273, 0.004)
22	51.917 -4.417	N(32.44, 6.83)	N(23.69, 3.96)	N(1.033, 0.0065)
23	51.75 0.25	N(19.75, 13.53)	N(43.24, 20.51)	N(1.350, 0.0005)
24	51.75 -2.583	N(36.85, 102.06)	N(24.31, 49.05)	N(1.218, 0.0030)
25	51.75 -4.25	N(37.86, 230.08)	N(21.53, 27.54)	N(1.286, 0.0030)
26	51.25 1.083	N(19.31, 26.37)	N(35.39, 46.22)	N(1.260, 0.0025)
27	51.083 0.25	N(16.63, 5.45)	N(24.77, 63.51)	N(1.329, 0.001)
28	51.083 -0.083	N(18.74, 20.03)	N(21.46, 45.45)	N(1.322, 0.001)
29	51.083 -1.583	N(19.41, 57.47)	N(34.90, 46.41)	N(1.226, 0.002)
30	51.083 -3.917	N(31.42, 5.98)	N(25.62, 3.6)	N(1.158, 0.00055)
31	50.75 -1.583	N(28.02, 159.45)	N(35.11, 92.82)	N(1.327, 0.0025)
32	50.583 -3.75	N(47.70, 101.35)	N(15.78, 34.03)	N(1.077, 0.003)
33	50.25 -4.917	N(32.92, 0.94)	N(25.68, 0.18)	N(1.198, 0.0015)

a given (latitude, longitude) coordinate $\boldsymbol{y}_k$. The emulator variance associated with $Em_t(\boldsymbol{y}_k)$ is denoted by $Vm_t(\boldsymbol{y}_k)$. Then, taking into account emulator uncertainty, the overall mean and variance for type t at $\boldsymbol{y}_k$ are $m_t(\boldsymbol{y}_k) = Em_t(\boldsymbol{y}_k)$ and $v_t(\boldsymbol{y}_k) = Ev_t(\boldsymbol{y}_k) + Vm_t(\boldsymbol{y}_k)$. Figure 3 shows the standard deviations (SD) derived from this interpolation for NBP. These would represent maps of mean and SD of NBP if E&W had 100% coverage of grassland, crop, DcBl or EvNl respectively. The SD maps also include a contribution from the uncertainty in the mean correction, which is due to emulation. The size of this contribution depends on the amount of spatial variation seen in the mean NBP or GPP values at the sample sites. It will also be largest in regions where there are few nearby sites. The mean percentage of the variance due to uncertainty in the interpolation (averaged over all 707 E&W grid-cells) is 5.9% for EvNl, 9.6% for grassland, 16.03% for crop and 23.21% for DcBl.

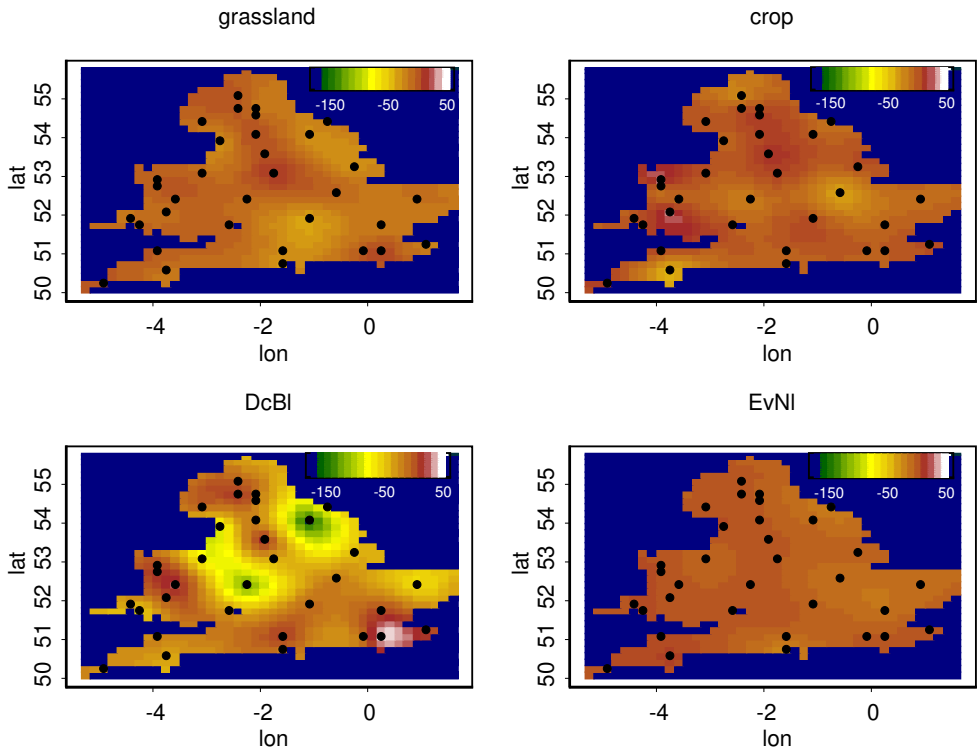

Figure 2: Mean NBP corrections (gC/m^2) with 33 sample sites shown.

Features and apparent anomalies in these graphs can be understood by examination of the emulation and interpolation analyses in more depth. For example, the map of NBP corrections for crops in Figure 2 shows significant North to South variations in a central band streching from Wales to East Anglia. The largest corrections 29.75 and 28.21 occur at latitudes 52.917 and 52.083. Between these latitudes are 5 sites which all have negative corrections, the most extreme being -39.29. Another feature seen in these plots is a large positive correction in NBP for DcBl in the South-East corner. This is due to the interpolation and caused by a large difference between values -16.35 at (51.083, -0.083) and 47.15 at (51.083, 0.25).

The most extreme corrections seen in Figure 2 occur for Deciduous Broadleaf at points (52.417, -2.25), (54.083, -1.083) and (54.417, -0.75). These are all negative corrections. An explanation for these corrections is suggested by examining the main effects of each input, as shown in Figure 4 for site (54.417, -0.75). Notice that the effect of soil bulk density is reasonably constant up to around 1.1, then has a sharp fall. The range over which bulk density for this site has significant probability (see Table 2) coincides with this sharp drop, resulting in the large negative correction. A positive correction is seen for Deciduous Broadleaf at site (51.083, 0.25). Here, the effect of increasing leaf lifespan is monotonic and positive, in contrast to the effect seen in Figure 4. The relatively low sand content is also in a region of the parameter space which leads to increasing NBP.

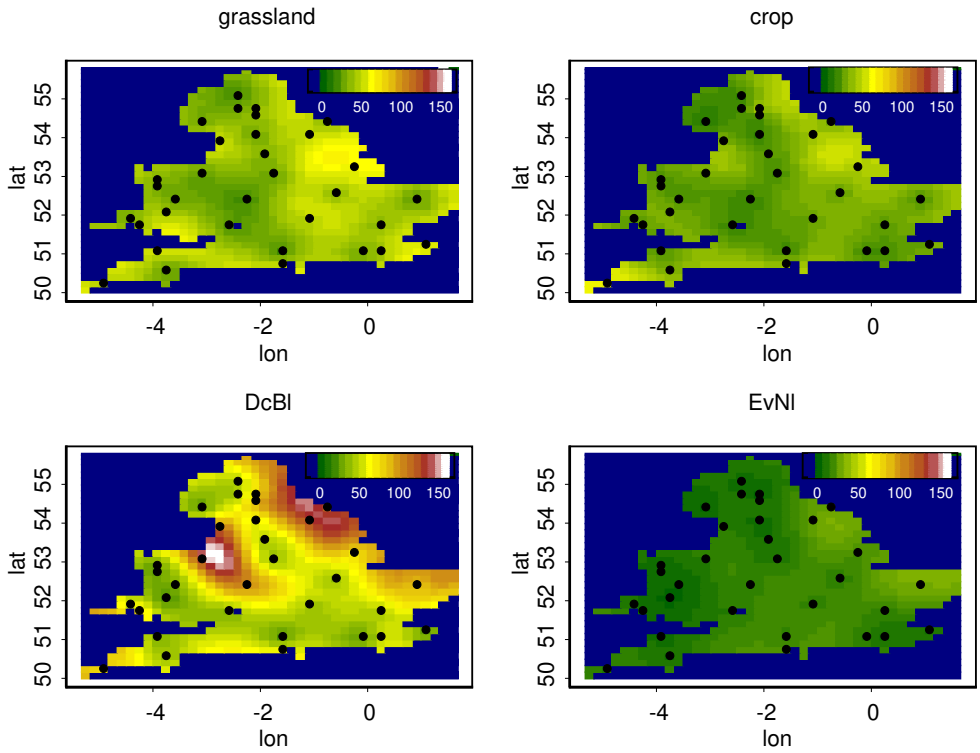

Figure 3: Standard deviation of NBP corrections (gC/m^2) with 33 sample sites shown.

The Deciduous Broadleaf PFT also gives rise to the largest standard deviations. The largest is at site (53.083, -3.083). To understand the source of this uncertainty, we can use the methods of Oakley and O'Hagan (2004) to partition the variance at the site into contributions from the main effects of each uncertain input. It is found that 81% of the total variance at this site is due to uncertainty in soil bulk density, while soil sand percentage and leaf lifespan contribute only approximately 4% and 2% respectively.

The combined contributions to total output variance from main effects of all soil parameters and all plant parameters are shown in Figure 5. The height of each bar, representing the sum of all main effects, is not 100% because there are also contributions from interaction effects, but it is clear that the overall influence of soil parameter uncertainty is much lower for grassland and crops than it is for the tree types. We return to this point in §4.

3.4 E&W combined map

The LCM2000 land cover map of Haines-Young et al. (2000) provides proportions of area covered by each PFT within each grid-cell (Figure 6). This can be used to combine the individual PFT maps $m_t(\cdot)$ and $v_t(\cdot)$, found in §3.3, into a single map for E&W. We will denote the proportions at grid-cell $\boldsymbol{y}_k$ by $\gamma_t(\boldsymbol{y}_k)$, where t ranges through the four plant types. For each E&W grid-cell

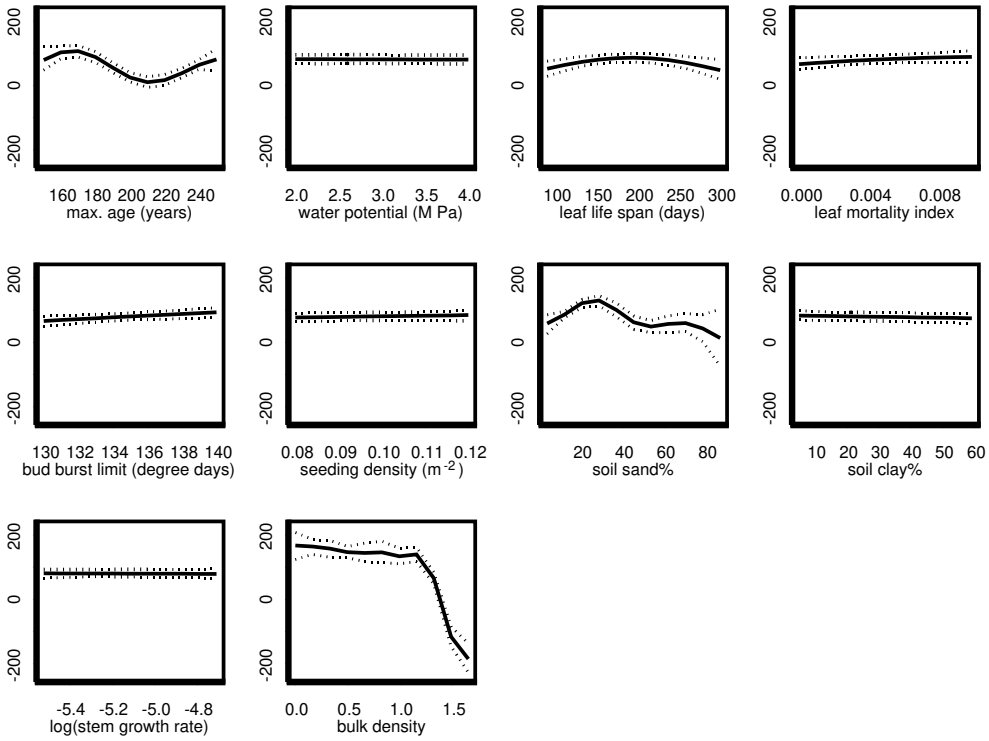

Figure 4: Main effects of input parameters on SDGVM NBP output for DcBl at site (54.417, -0.75) (gC/m^2). The solid lines are the mean main effects and the dotted lines show 95% probability bounds due to emulation.

$\boldsymbol{y}_k$, the mean NBP is

$$m(\boldsymbol{y}_k) = \sum_t \gamma_t(\boldsymbol{y}_k) m_t(\boldsymbol{y}_k),$$

and the variance of NBP is

$$v(\boldsymbol{y}_k) = \sum_t \gamma_t(\boldsymbol{y}_k)^2 v_t(\boldsymbol{y}_k) + 2 \sum_{t=1}^{3} \sum_{t'=t+1}^{4} \gamma_t(\boldsymbol{y}_k)\gamma_{t'}(\boldsymbol{y}_k) Ec_{t,t'}(\boldsymbol{y}_k). \quad (5)$$

The variance (5) includes covariances $Ec_{t,t'}(\boldsymbol{y}_k)$ between PFT outputs. These are non-zero because the 3 site specific soil input values are common to all PFTs. For each pair of types (t, t'), $Ec_{t,t'}(\boldsymbol{y}_k)$ is a result of interpolating values calculated at the sample sites, as with the means and variances. We now consider the calculation of $Ec_{t,t'}(\boldsymbol{y}_k)$ at a sample site k. Suppose the unknown input vector for PFT t at this site is denoted by $\boldsymbol{X}_{kt}$. Then

$$Ec_{t,t'}(\boldsymbol{y}_k) = \left(E[Var\{f_{kt}(\boldsymbol{X}_{kt}) + f_{kt'}(\boldsymbol{X}_{kt'})\}] - Ev_t(\boldsymbol{y}_t) - Ev_{t'}(\boldsymbol{y}_{t'})\right)/2,$$

where the term $E[Var\{f_{kt}(\boldsymbol{X}_{kt}) + f_{kt'}(\boldsymbol{X}_{kt'})\}]$ can be calculated using the methods of §2 with $N = 2$.

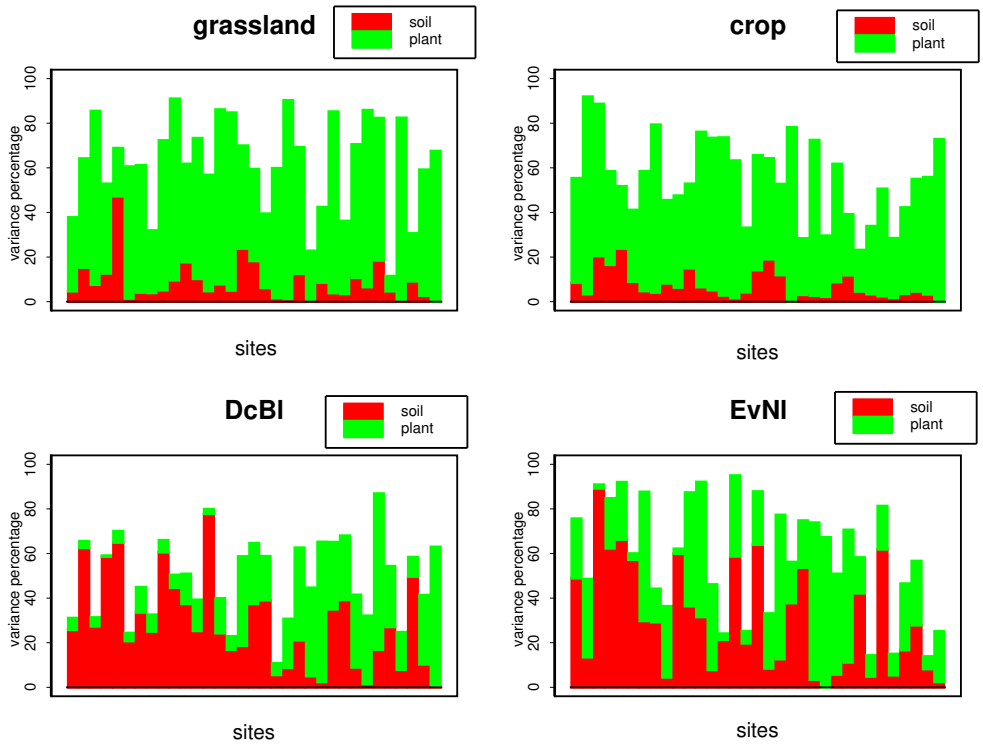

Figure 5: Proportion of the output variance explained by soil parameters and plant parameters for each sample site and each PFT. Sites are ordered as in Table 2.

The mean and standard deviation maps taking into account land cover are shown in Figures 7 and 8. The largest standard deviation for NBP occurs around site (54.417, -0.75) where there is a 36% proportion of deciduous broadleaf, according to LCM2000. The largest uncertainty in GPP occurs in West Wales. Inspection of the land cover map shows a 78.5% coverage of grassland and 10% DcBl, both of which produce high standard deviations in this region. Overall the differences between the corrected maps in Figures 7 and 8 and their plug-in counterparts (Figure 1) appear small. The most significant corrections of around -40 for NBP occur in small regions close to sites (54.417, -0.75), (53.75, -2.75) and (52.417, -2.25). For GPP there are negative corrections of approximately -100 to -150 in these regions.

3.5 Expert elicitation for between-site correlations

In order to quantify uncertainty about aggregates across E&W we need to identify correlations between sites. We have assumed independence between different inputs, but correlations arise between the values of a given input at different sites. For instance, if the expert were able to learn the value of the *leaf lifespan* for the DcBl PFT in one site, then it would influence his beliefs about the value of this parameter at another site. The principal reason for this is the common uncertainty about leaf lifespans for deciduous broadleaf species. There have been few

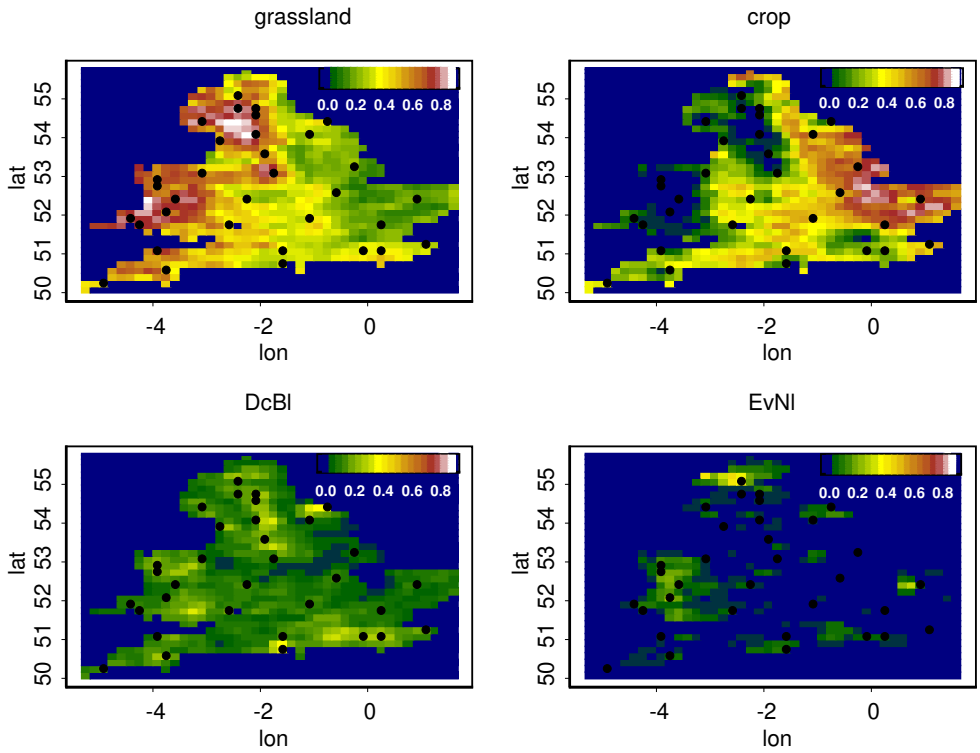

Figure 6: Percentage of land covered by each plant functional type, based on the LCM2000 land cover map.

recorded measurements of this parameter, covering very few species. The possibility of such correlations was considered for each input, and the following cases were identified.

Soil inputs. Because detailed site-specific data (typically, samples of around 190 measurements) are available on soil inputs, it was deemed appropriate to treat these inputs as independent across sites.

Constants. At the other extreme are unknown constants that are common to all sites and have a single distribution. *Minimum stem growth rate* falls into this category (although it is not an influential parameter in any of the PFTs). It has no spatial variability, so we set the correlation to 1 between any pair of sites.

All other inputs except EvNl leaf lifespan. For all other PFT inputs $x_{\_i}$, the expert felt that there would be some spatial variability but also correlation. In the case of DcBl *leaf lifespan*, for instance, there is correlation due to uncertainty about the typical value that this parameter would take for deciduous broadleaf species, but there will also be spatial variability due to variations in the mix of species present at different sites. Although in principle sites close together might have higher correlations, it would have been impractical to elicit such a complex structure, and so a common correlation $\tau_{\_i}$ was assumed between all pairs of sites. For such an input, let $\bar{x}_{\_i}$ denote the average value of $x_{\_i}$ over E&W. In order to find a suitable value for $\tau_{\_i}$, the expert was asked to provide a possible range for $\bar{x}_{\_i}$, and from this the mean m_{exp} and variance v_{exp}

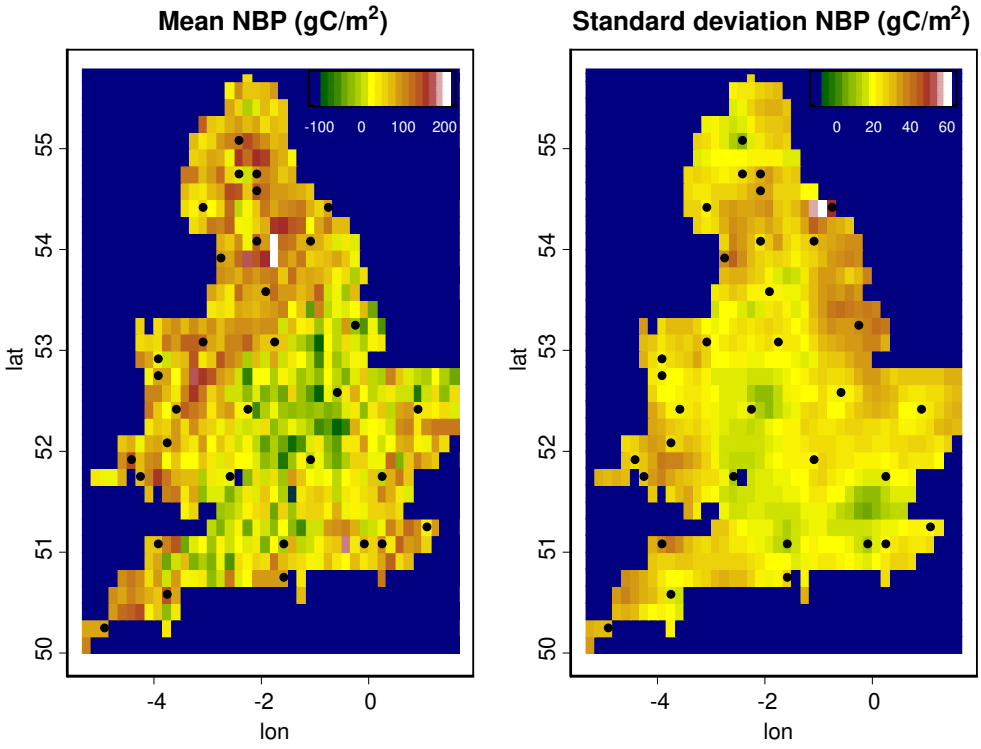

Figure 7: Mean and standard deviation of NBP using the LCM2000 land cover weighting. Dots indicate the 33 sample sites.

for his distribution for $\bar{x}_{\_i}$ were inferred as in §3.1. The individual grid-cell mean and variance of $x_{\_i}$, denoted by m_1 and v_1, are already known from the elicitation of §3.1. If we assume that $x_{\_i} \sim N(m_1, v_1)$ and $Cor(x_{\_i}, x'_{\_i}) = \tau_{\_i}$ for values at different grid-cells, and that the number of values averaged to obtain $\bar{x}_{\_i}$ is N, then $E(\bar{x}_{\_i}) = m_1$ and

$$Var(\bar{x}_{\_i}) = \frac{1}{N^2}\left(\sum_{k=1}^{N} v_1 + 2\sum_{k=1}^{N-1}\sum_{k'=k+1}^{N} \tau_{\_i} v_1\right). \tag{6}$$

By equating these to the expert's mean m_{exp} and variance v_{exp} and solving we find $m_1 = m_{exp}$ and

$$\tau_{\_i} = \frac{N(v_{exp}/v_1) - 1}{N - 1}. \tag{7}$$

This exercise was conducted for each influential vegetation input parameter. A default value $\tau_{\_i} = 0.5$ was used for inputs deemed non-influential.

EvNl leaf lifespan. In the case of inputs whose uncertainty marginally is represented by a single normal distribution, the above approach is adequate to identify joint normal distributions across different sites with correlation coefficient $\tau_{\_i}$. However, for EvNl *leaf lifespan* a more complex analysis is required to produce a joint distribution in the form of a mixture of normals. Details are given in Appendix B.

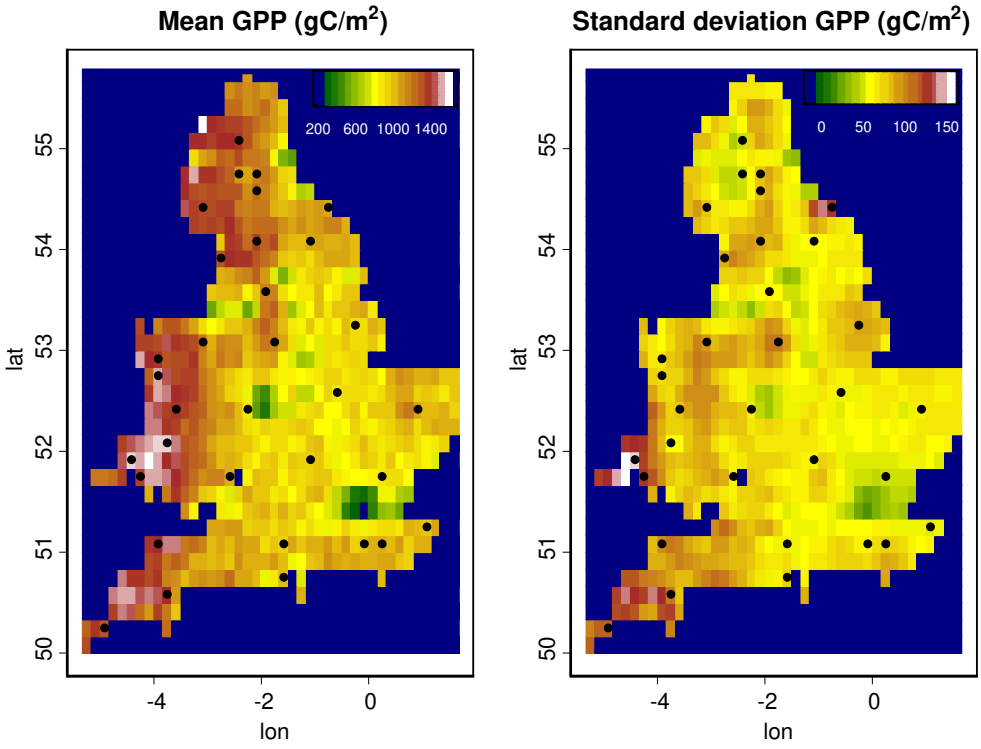

Figure 8: Mean and standard deviation of GPP using the LCM2000 land cover weighting. Dots indicate the 33 sample sites.

3.6 Inference for the E&W aggregates

The uncertainty analysis used to build the single PFT maps in §3.3 provided means and variances for NBP at the 33 sample sites. In order to estimate the variance of aggregate NBP for the whole of E&W, we need the correlations between any pair of grid-cells in E&W. To estimate these for PFT t, we first used the methodology of §2 with $q = 2$ to estimate the variances of all pairwise sums $Var\{f_{kt}(\boldsymbol{X}_{kt}) + f_{k't}(\boldsymbol{X}_{k't})\}$, for sample sites $k, k' = 1, \ldots, 33$ and $k < k'$. From these we derived the sample site covariance matrix $\boldsymbol{V}_{test}$ for the full set of 33 outputs, using the fact that $Cov(f_{kt}(\boldsymbol{X}_{kt}), f_{k't}(\boldsymbol{X}_{k't})) = [Var\{f_{kt}(\boldsymbol{X}_{kt}) + f_{k't}(\boldsymbol{X}_{k't})\} - Var\{f_{kt}(\boldsymbol{X}_{kt})\} - Var\{f_{k't}(\boldsymbol{X}_{k't})\}]/2$ and $Var\{f_{kt}(\boldsymbol{X}_{kt})\}$ is already calculated as the variance at sample site k.

The following approximation model was then used to extrapolate from our representative sites to the whole map for a single PFT, type t. Let $\boldsymbol{Y}_t$ represent the vector of NBP values over $N = 707$ E&W grid-cells of the tth PFT map. We assume that $\boldsymbol{Y}_t$ is distributed approximately as $\boldsymbol{Y}_t \sim N_N(\boldsymbol{m}_t, \boldsymbol{V}_t)$ where $\boldsymbol{m}_t = (m_t(\boldsymbol{y}_1), \ldots, m_t(\boldsymbol{y}_N))$ and

$$\boldsymbol{V}_t = \begin{pmatrix} v_t(\boldsymbol{y}_1) & \rho_t\sqrt{v_t(\boldsymbol{y}_1)v_t(\boldsymbol{y}_2)} & \ldots & \rho_t\sqrt{v_t(\boldsymbol{y}_1)v_t(\boldsymbol{y}_N)} \\ \rho_t\sqrt{v_t(\boldsymbol{y}_2)v_t(\boldsymbol{y}_1)} & \ddots & & \vdots \\ \vdots & & \ddots & \rho_t\sqrt{v_t(\boldsymbol{y}_{N-1})v_t(\boldsymbol{y}_N)} \\ \rho_t\sqrt{v_t(\boldsymbol{y}_N)v_t(\boldsymbol{y}_1)} & \ldots & & v_t(\boldsymbol{y}_N) \end{pmatrix}. \quad (8)$$

By using (8) we do not build spatial correlation into the model, since an earlier study revealed that the distance between grid-cells had little effect on the correlation.

Finally we selected values of ρ_t such that the sum of the terms in the sample site covariance matrix $\boldsymbol{V}_{test}$ for PFT t equalled the sum of elements in (8) with $N = 33$. This process was repeated for $t = 1, \ldots, 4$, based on the individual sample site emulators for type t, and produced estimates $\rho_1 = 0.012$ and $\rho_2 = 0.051$, $\rho_3 = 0.006$, $\rho_4 = 0.182$ for grassland, crop, DcBl and EvNl respectively. The correlations for DcBl, crop and grassland are extremely small, suggesting that correlations in the plant functional type parameters do not induce correlations between outputs.

By combining the results from the four PFT maps and weighting according to the known land cover proportions $\gamma_t(\boldsymbol{y}_k)$ and areas a_k of each grid-cell $\boldsymbol{y}_k$, we obtain the following for the mean M^* and variance V^* of the total NBP carbon flux:

$$M^* = \sum_{t=1}^{4}\sum_{k=1}^{N} a_k \gamma_t(\boldsymbol{y}_k) m_t(\boldsymbol{y}_k)$$

and

$$V^* = \sum_{k=1}^{N} a_k^2 v(\boldsymbol{y}_k) + 2\sum_{t=1}^{4}\left\{\rho_t \sum_{k=1}^{N-1}\sum_{k'=k+1}^{N} a_k a_{k'} \gamma_t(\boldsymbol{y}_k)\gamma_t(\boldsymbol{y}_{k'})\sqrt{v_t(\boldsymbol{y}_k)v_t(\boldsymbol{y}_{k'})}\right\} \quad (9)$$

to give $M^* = 7.6118$ MtC (1 MtC $= 10^{12}$ gC) and $V^* = (0.6094 \text{ MtC})^2$. The mean is significantly lower than the aggregate based on the uncorrected plug-in model runs, which is 9.0613 MtC. The second term of (9) includes all between site covariances and the first term includes all the between PFT covariances through (5). As soil inputs are independent between sites, there are no between site covariances for distinct PFTs. Table 3 shows the contributions to these values from each of the different PFTs and the covariances between PFTs.

Table 3: Contribution to the mean and variance of total NBP from different plant functional types and covariances between these types.

PFT	Plug-in estimate (Mt C)	Mean (Mt C)	Variance (Mt C)^2
grassland	5.2793	4.6471	0.3229
crop	0.8525	0.4976	0.0380
DcBl	2.1319	1.6867	0.0087
EvNl	0.7976	0.7803	0.0007
covariances			0.0011
Total	9.0613	7.6118	0.3714

4 Discussion

We have illustrated a general methodology for creating uncertainty maps for carbon accounting based on outputs from a mechanistic computer model. We have also shown how the overall uncertainty can be partitioned into contributions from different plant functional types, each of

which can be further subdivided into individual parameter contributions at any given sample site. This allows modellers to identify and quantify the most significant sources of uncertainty, which should be helpful in the formulation of efficient strategies to reduce overall uncertainty.

The results presented here can be improved in a number of ways. The model itself is being continuously improved as the modellers gain greater understanding of the processes involved and also as a result of feedbacks provided by the statistical analysis presented here. It is important to remember that these results are purely model based. We have quantified uncertainty in the model output induced by uncertainties in soil and plant input parameters. The true physical value of the carbon flux includes extra uncertainty due to inadequacies in the computer model, which are not accounted for. The SDGVMd model does not account for grazing, for example, which could lead to overestimation of NBP from grassland. These inadequacies should be reduced as a result of the ongoing model improvements. If corresponding physical measurements became available these could be used to learn about the model inadequacy, as in Kennedy and O'Hagan (2001). Our results also assume that land cover is known for the year 2000 and that the climate inputs are known. A natural extension of the work presented here is to consider the impact of uncertainty in the land cover map. Given the mean and variance maps for individual PFTs, the proportions of each PFT within a grid-cell could be treated as uncertain rather than as fixed at the LCM2000 values. Daily climate data produced by the weather generator is assumed to be known, and this too could lead to underestimation in the total uncertainty.

The estimate M^* for the aggregate NBP is not directly comparable with other methods for carbon accounting, which focus specifically on removal of CO_2 by forests and include Scotland. The comparable estimate of NBP is the sum of DcBl and EvNl, which from Table 3 is 2.467 MtC. The value given in Milne and Cannell (2005) for the UK in 2000 is 4.173 MtC. From woodland inventories of Great Britain (Forestry Commision (2002)) and Northern Ireland (Northern Ireland Forest Service (2001)) woodland in E&W accounted for approximately 50.76% of the UK total in 2000. A simple scaling of the UK total therefore produces 2.118 MtC, which is smaller than our estimate of 2.467 MtC and in the lower tail of our distribution (the standard deviation is 0.0964 MtC). As already explained, incorporating uncertainty in land cover and other factors will further increase uncertainty. Scaling our estimate to account for differences in total forest area between the 2 methods gives 2.312 MtC. Discrepancies may also be partly due to model inadequacy. The SDGVMd plug-in estimate from Table 3 is 2.93 MtC, which appears to be too large.

Table 3 shows how the total flux and its uncertainty are clearly dominated by the contribution from grassland, even though both EvNl and DcBl generally have higher NBP per m^2. For example, the average mean NBP at the sample sites is 75.89 gC/m^2 for grassland, 104.67 gC/m^2 for DcBl and 172.63 gC/m^2 for EvNl. The dominant influence of grassland in the total is due to the fact that it has a much greater coverage than the forest types in the land cover map (see Figure 6). Further efforts to reduce uncertainty should therefore concentrate on reducing uncertainty in the input parameters for grassland. If uncertainty in the land cover map was included, we might expect the variance contributions from EvNl and DcBl to increase substantially.

Included in the variance calculation V^* is a contribution from the uncertainty due to the interpolation between sample sites. This accounts for 11.2% of the variance. Additional sample sites could be added in regions where the maps show high interpolation uncertainty to reduce

this contribution to the variance.

The parameters having the greatest impact on NBP output vary between PFTs. For DcBl, bulk density and leaf lifespan are the key parameters which explain most variation in the output. The main effects plotted in Figure 4 highlight something that needs further investigation, in that plants are dying within SDGVMd for large (but realistic) values of soil bulk density. This suggests a weakness in the soil process modelling within SDGVMd which must be addressed. The particular processes that operate on organic soils are also absent from SDGVMd, and this is a direction for future model development.

For EvNl, a range of parameters make substantial contributions to the output variance, although bulk density and soil clay are the dominant ones.

For grassland and crop PFTs, the soil parameters are less influential, although the effect of soil bulk density is still substantial. These are the PFTs contributing most to the variance of the carbon flux, so it is important to understand and verify the input-output relationships. Crops and grassland are better able to adapt to changes in water holding capacity of the soil, so we would expect these PFTs to be less dependent than trees on soil parameters.

Acknowledgements

The project is funded by the Natural Environment Research Council, contract number F14/G6/105. The authors would like to thank Iain Bradley from the National Soil Research Institute (NSRI) for help with the England and Wales soils dataset and to NSRI for providing soils data (licence 02/RGOB/171/OD7063V).

Appendix A: Emulation of computer outputs

Here we give a brief summary of the emulation process based on a Gaussian process model. A more extensive general discussion on emulation of computer models can be found in Kennedy and O'Hagan (2001). To simplify notation here we use subscript k for output k, but we could also use kt to refer to the output representing vegetation type t at site k. For a function with p-dimensional input vector $\boldsymbol{x}$, we assume the stationary Gaussian process prior model

$$\big[\boldsymbol{f}_k(\cdot) \mid \boldsymbol{\beta}_k, \boldsymbol{\sigma}_k, \boldsymbol{r}_k\big] \sim N\big(\boldsymbol{h}(\cdot)^T\boldsymbol{\beta}_k, \boldsymbol{\sigma}_k^2 c_k(\cdot,\cdot)\big),$$

where the correlation function has the product form

$$c_k(\boldsymbol{x}, \boldsymbol{x}') = \prod_{i=1}^{p} \exp\big\{-r_{k\_i}(x_{\_i} - x_{\_i}{}')^2\big\}.$$

The vector $\boldsymbol{r}_k = (r_{k\_1}, \ldots, r_{k\_p})$ determines the roughness in output k as a function of each input. Values of $\boldsymbol{r}_k$ are found by maximising the posterior density $p(\boldsymbol{r}_k \mid \boldsymbol{d}_k)$ given data $\boldsymbol{d}_k$. We choose regression functions $\boldsymbol{h}(\boldsymbol{x}) = (1, x_{\_1}, \ldots, x_{\_p})^T$, to account explicitly for the output having *a priori* an overall mean and a linear trend in each of the inputs.

We shall denote by $m_k^*(\cdot)$ and $v_k^*(\cdot,\cdot)$ the marginal posterior mean and covariance functions for the emulator of the kth output $f_k(\cdot)$. Conditional on the observed code data, comprising

input sets $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$ and corresponding outputs $\boldsymbol{d}_k = f_k(\boldsymbol{s}_1), \ldots, f_k(\boldsymbol{s}_n)$, these are given by

$$m_k^*(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x})^T \hat{\boldsymbol{\beta}}_k + \boldsymbol{t}_k(\boldsymbol{x})^T \boldsymbol{g}_k$$

and

$$\begin{aligned} v_k^*(\boldsymbol{x}, \boldsymbol{x}') &= \hat{\sigma_k}^2 \{c_k(\boldsymbol{x}, \boldsymbol{x}') - \boldsymbol{t}_k(\boldsymbol{x})^T \boldsymbol{A}_k^{-1} \boldsymbol{t}_k(\boldsymbol{x}') \\ &\quad + (\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{H}^T \boldsymbol{A}_k^{-1} \boldsymbol{t}_k(\boldsymbol{x}))^T (\boldsymbol{H}^T \boldsymbol{A}_k^{-1} \boldsymbol{H})^{-1} (\boldsymbol{h}(\boldsymbol{x}') - \boldsymbol{H}^T \boldsymbol{A}_k^{-1} \boldsymbol{t}_k(\boldsymbol{x}'))\}, \end{aligned}$$

where $\boldsymbol{t}_k^T(\cdot) = \big[c_k(\cdot, \boldsymbol{s}_1), \cdots, c_k(\cdot, \boldsymbol{s}_n)\big]$, $\boldsymbol{H}^T = \big[\boldsymbol{h}(\boldsymbol{s}_1), \cdots, \boldsymbol{h}(\boldsymbol{s}_n)\big]$, $\boldsymbol{A}_k = [c_k(\boldsymbol{s}_i, \boldsymbol{s}_j)] \in \mathbb{R}_{n,n}$, $\hat{\boldsymbol{\beta}}_k = \big(\boldsymbol{H}^T \boldsymbol{A}_k^{-1} \boldsymbol{H}\big)^{-1} \boldsymbol{H}^T \boldsymbol{A}_k^{-1} \boldsymbol{d}_k$ is the estimated vector of linear regression coefficients, $\boldsymbol{g}_k = \boldsymbol{A}_k^{-1}(\boldsymbol{d}_k - \boldsymbol{H}\hat{\boldsymbol{\beta}}_k)$ and $\hat{\sigma}_k^2 = \boldsymbol{d}_k^T \{\boldsymbol{A}_k^{-1} - \boldsymbol{A}_k^{-1} \boldsymbol{H} \big(\boldsymbol{H}^T \boldsymbol{A}_k^{-1} \boldsymbol{H}\big)^{-1} \boldsymbol{H}^T \boldsymbol{A}_k^{-1}\} \boldsymbol{d}_k$.

Appendix B: Normal mixtures for parameter uncertainty

The elicited distributions for individual uncertain model inputs are specified in Tables 1 and 2. The expert did not feel that acquiring more information about any of the inputs would materially change his beliefs about the others, so different inputs were assumed to be independent. However, we need to consider correlation between the different values that the same inputs might take at different sites, as discussed in §3.5. To simplify notation here we use subscript k to index output k, but we could also use kt to refer to the output representing vegetation type t at site k. We denote the marginal distribution of input i at site (i.e. output) k by

$$G_{k\_i} \sim \sum_{u=1}^{q_{\_i}} \alpha_{\_i}^u N(m_{k\_i}^u, v_{k\_i}^u), \tag{10}$$

where $\sum_{u=1}^{q_{\_i}} \alpha_{\_i}^u = 1$ for $i = 1, \ldots, p$. Superscripts u or u' will be used throughout to index components in a mixture distribution. The number of mixture components, $q_{\_i}$, is 1 for all inputs except EvNl *leaf lifespan*, for which $q_{\_i} = 5$. The values of $m_{k\_i}^u$ and $v_{k\_i}^u$ are given in Tables 1 (noting that for the PFT parameters these means and variances are the same for all sites) and 2. The independence between inputs means that for site k the distribution G_k is the product $\prod_{i=1}^p G_{k\_i}$.

As discussed in §3.5, independence was also assumed between the soil inputs in different sites, but there is correlation between the values of PFT inputs across sites. For a PFT input i whose uncertainty is represented by a single normal distribution, it is straightforward to introduce a correlation $\tau_{\_i}$, which was set to 1 for inputs that represent constants taking the same value in each site, to 0.5 for other non-influential PFT inputs, or to a suitable value obtained through eliciting the expert's beliefs about the E&W mean $\bar{x}_{\_i}$. However, some thought is needed to model the joint distribution of values of EvNl *leaf lifespan* across sites. The mixture distribution marginally reflects the fact that evergreen leaves may have a lifespan of several years. It would not have been appropriate to model the joint distribution $G_{kk'\_i}$ for outputs k and k' by a mixture of 5 bivariate normal components with margins the 5 components of the mixture given in Table 1, because this would imply lifetimes in both sites extending for the same number of years. Instead, the joint distribution was assumed to be a mixture of 25 terms. In general, we

write the joint distribution for any input i at sites k and k' by the mixture

$$G_{kk'\_i} \sim \sum_{u=1}^{q\_i} \sum_{u'=1}^{q\_i} \alpha_{\_i}^{(uu')} N\left(\begin{pmatrix} m_{k\_i}^{u} \\ m_{k'\_i}^{u'} \end{pmatrix}, \begin{pmatrix} v_{k\_i}^{u} & \tau_{\_i}\sqrt{(v_{k\_i}^{u} v_{k'\_i}^{u'})} \\ \tau_{\_i}\sqrt{(v_{k\_i}^{u} v_{k'\_i}^{u'})} & v_{k'\_i}^{u'} \end{pmatrix} \right), \qquad (11)$$

although for all inputs except for EvNl *leaf lifespan*, this mixture is degenerate. Independence between inputs then leads to the representation of the joint distribution $G_{kk'}$ as a product of bivariate Normal mixtures $G_{kk'}(\boldsymbol{x}, \boldsymbol{x'}) = \prod_{i=1}^{p} G_{kk'\_i}(x_{\_i}, x'_{\_i})$.

Note that the $q_{\_i}^2$-component mixture (11) is constructed so that its marginal distributions agree with the $q_{\_i}$-component mixture (10), provided that the coefficients in (11) have the property that $\sum_{u'=1}^{q\_i} \alpha_{\_i}^{(uu')} = \alpha_{\_i}^{u}$ and by symmetry $\alpha_{\_i}^{(uu')} = \alpha_{\_i}^{(u'u)}$ for $i = 1, \ldots, p$. The parameter $\tau_{\_i}$ is the same in all of the mixture components. More generally, it could be allowed to vary, but this is unnecessary for the degenerate mixtures and in the case of EvNl *leaf lifespan* the common value reflects the expert's view (discussed in §3.1) that uncertainty about this input could be considered in terms of independent distributions for the number of years and the day number within the year.

We now consider how to specify the weights $\alpha_{\_i}^{(uu')}$ for the mixture components for EvNl *leaf lifespan*. It would not be appropriate to assume independence between the numbers of years at the two sites, so we cannot simply set $\alpha_{\_i}^{(uu')} = \alpha_{\_i}^{u}\alpha_{\_i}^{u'}$. Recall from Table 1 that the distribution of individual values $x_{\_i}$ of *leaf lifespan* at the grid-cell level is a five-component normal mixture. The five mixing proportions themselves constitute a discrete probability distribution, which, if there were any advantage in doing so, could be specified (in many ways) by discretization of a continuous distribution. For example, we could take a standard Normal distribution, partition the line into five intervals and choose the end-points of the intervals in such a way that the probabilities ascribed to the intervals by the Normal distribution were the five mixing proportions. In fact, this construction gives a simple basis for the generation of the bivariate mixing proportions $\alpha_{\_i}^{(uu')}$. Take a bivariate Normal distribution with standard margins and a correlation τ_α, and find the probability content of rectangles formed by the Cartesian product of the five linear intervals above. For each value of τ_α the resulting discrete bivariate distribution has the properties required for the $\alpha_{\_i}^{(uu')}$ in $G_{kk'\_i}$: symmetry and reduction to the correct marginal form. The means and variances of the mixture components in $G_{kk'\_i}$ are determined by the earlier marginal specifications (Table 1) and are the same for all sites, so $m_{k\_i}^{u} = m_{k'\_i}^{u}$ and $v_{k\_i}^{u} = v_{k'\_i}^{u}$ for $u = 1, \ldots, 5$ in (11). Thus by use of these values for the $\alpha_{\_i}^{(uu')}$ and the means and variances from Table 1, the joint distribution is specified up to the value of the correlation coefficient τ_α.

We have now formulated the joint distribution in terms of two correlation coefficients, $\tau_{\_i}$ representing correlation in the numbers of days within the year and τ_α representing correlation between the numbers of years. To identify values for these we have only the expert's expressed variance for the overall E&W average value $\bar{x}_{\_i}$ for EvNl *leaf lifespan*, which was $48,586$. This correlation is primarily driven by τ_α, so we simply fixed $\tau_{\_i} = 0.5$ and chose τ_α to match the

elicited variance of $\bar{x}_{\_i}$, which is given by

$$\begin{aligned} Var(\bar{x}_{\_i}) &= \frac{1}{N^2}\Bigg\{\sum_{k=1}^{N} Var_k(x_{\_i}) \\ &\quad +2\sum_{k=1}^{N-1}\sum_{k'=k+1}^{N}\left(\sum_{u=1}^{5}\sum_{u'=1}^{5}\alpha_{\_i}^{(uu')}(\tau_{\_i}\sqrt{(v_{k\_i}^{u}v_{k'\_i}^{u'})}+m_{k\_i}^{u}m_{k'\_i}^{u'})-\left(\sum_{u=1}^{5}\alpha_{\_i}^{u}m_{k'\_i}^{u}\right)^2\right)\Bigg\}, \end{aligned} \tag{12}$$

where

$$Var_k(x_{\_i}) = \sum_{u=1}^{5}\alpha_{\_i}^{u}\{v_{k'\_i}^{u}+(m_{k\_i}^{u})^2\}-\left(\sum_{u=1}^{5}\alpha_{\_i}^{u}m_{k\_i}^{u}\right)^2 \tag{13}$$

is the marginal variance of the single site *leaf lifespan*. The fitted value of τ_α is 0.27.

Appendix C: Closed forms for the integrals

The integrals developed in this appendix are for the case of aggregation over k for a single PFT, and we use subscript k for site k. Integrals required for aggregation over PFTs (at a fixed site k) are similar, but subscript k would be replaced by kt, for example, and summation is over t. In §2 the mean and variance of the aggregate output, for the special case of aggregating over q sites for a fixed PFT, are expressed as

$$\begin{aligned} M &= E(F \mid f_1,\ldots,f_q) = \sum_{k=1}^{q}\int f_k(\boldsymbol{x})\,dG_k(\boldsymbol{x}), \\ V &= \mathbb{V}\mathrm{ar}(F \mid f_1,\ldots,f_q) = M_2 - M^2, \end{aligned}$$

where

$$\begin{aligned} M_2 = E(F^2 \mid f_1,\ldots,f_q) &= \sum_{k=1}^{q}\int f_k(\boldsymbol{x})^2\,dG_k(\boldsymbol{x}) \\ &\quad +2\sum_{k=1}^{q-1}\sum_{k'=k+1}^{q}\int f_k(\boldsymbol{x})f_{k'}(\boldsymbol{x}')\,dG_{kk'}(\boldsymbol{x},\boldsymbol{x}'). \end{aligned}$$

The mean and variance of M with respect to the distribution of the unknown code can be expressed in terms of the emulator means and covariances. The mean is

$$E(M) = \sum_{k=1}^{q}\int m_k^*(\boldsymbol{x})\,dG_k(\boldsymbol{x}) = \sum_{k=1}^{q}(\hat{\boldsymbol{\beta}}_k^T\boldsymbol{R}_k + \boldsymbol{g}_k^T\boldsymbol{T}_k),$$

where

$$\boldsymbol{R}_k = \int \boldsymbol{h}(\boldsymbol{x})\,dG_k(\boldsymbol{x}), \qquad \boldsymbol{T}_k = \int \boldsymbol{t}_k(\boldsymbol{x})\,dG_k(\boldsymbol{x}).$$

The variance of M is

$$\begin{aligned}\mathbb{V}\mathrm{ar}(M) &= \sum_{k=1}^{q}\int v_k^*(\boldsymbol{x},\boldsymbol{x}')\,dG_k(\boldsymbol{x})dG_k(\boldsymbol{x}') \\ &= \sum_{k=1}^{q}\hat{\sigma}_k^2\big\{U_k-\boldsymbol{T}_k\boldsymbol{A}_k^{-1}\boldsymbol{T}_k+\big(\boldsymbol{R}_k-\boldsymbol{H}^T\boldsymbol{A}_k^{-1}\boldsymbol{T}_k\big)^T\big(\boldsymbol{H}^T\boldsymbol{A}_k^{-1}\boldsymbol{H}\big)^{-1}\cdot \\ &\quad\cdot\big(\boldsymbol{R}_k-\boldsymbol{H}^T\boldsymbol{A}_k^{-1}\boldsymbol{T}_k\big)\big\},\end{aligned}$$

where

$$U_k=\int\int c(\boldsymbol{x},\boldsymbol{x}')\,dG_k(\boldsymbol{x})\,dG_k(\boldsymbol{x}').$$

The emulator variance of V is complicated and is not given here. The mean of V is given by

$$E(V)=E(M_2)-\mathbb{V}\mathrm{ar}(M)-E(M)^2,$$

where

$$\begin{aligned}E(M_2) &= \sum_{k=1}^{q}\int\{v_k^*(\boldsymbol{x},\boldsymbol{x})+m_k^*(\boldsymbol{x})^2\}\,dG_k(\boldsymbol{x}) \\ &\qquad +2\sum_{k=1}^{q-1}\sum_{k'=k+1}^{q}\int\int m_k^*(\boldsymbol{x})m_{k'}^*(\boldsymbol{x}')\,dG_{kk'}(\boldsymbol{x},\boldsymbol{x}') \\ &= \sum_{k=1}^{q}\big\{\hat{\sigma}_k^2(1-tr(\boldsymbol{A}_k^{-1}\boldsymbol{P}_k)+tr\{(\boldsymbol{H}^T\boldsymbol{A}_k^{-1}\boldsymbol{H})^{-1}(\boldsymbol{Q}_k-\boldsymbol{S}_k\boldsymbol{A}_k^{-1}\boldsymbol{H} \\ &\qquad -\boldsymbol{H}^T\boldsymbol{A}_k^{-1}\boldsymbol{S}_k^T+\boldsymbol{H}^T\boldsymbol{A}_k^{-1}\boldsymbol{P}_k\boldsymbol{A}_k^{-1}\boldsymbol{H})\})+\hat{\boldsymbol{\beta}}_k^T\boldsymbol{Q}_k\hat{\boldsymbol{\beta}}_k+2\hat{\boldsymbol{\beta}}_k^T\boldsymbol{S}_k\boldsymbol{g}_k+\boldsymbol{g}_k^T\boldsymbol{P}_k\boldsymbol{g}_k\big\} \\ &\quad +2\sum_{k=1}^{q-1}\sum_{k'=k+1}^{q}\big\{\hat{\boldsymbol{\beta}}_k^T\boldsymbol{Q}_{kk'}\hat{\boldsymbol{\beta}}_{k'}+\hat{\boldsymbol{\beta}}_k^T\boldsymbol{S}_{kk'}\boldsymbol{g}_{k'}+\hat{\boldsymbol{\beta}}_{k'}^T\boldsymbol{S}_{k'k}\boldsymbol{g}_k+\boldsymbol{g}_k^T\boldsymbol{P}_{kk'}\boldsymbol{g}_{k'}\big\},\end{aligned}$$

in terms of the integrals

$$\begin{aligned}\boldsymbol{Q}_k &= \int\boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x})^T\,dG_k(\boldsymbol{x}), & \boldsymbol{Q}_{kk'} &= \int\int\boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}')^T\,dG_{kk'}(\boldsymbol{x},\boldsymbol{x}'), \\ \boldsymbol{S}_k &= \int\boldsymbol{h}(\boldsymbol{x})\boldsymbol{t}_k(\boldsymbol{x})^T\,dG_k(\boldsymbol{x}), & \boldsymbol{S}_{kk'} &= \int\int\boldsymbol{h}(\boldsymbol{x})\boldsymbol{t}_{k'}(\boldsymbol{x}')^T\,dG_{kk'}(\boldsymbol{x},\boldsymbol{x}'), \\ \boldsymbol{P}_k &= \int\boldsymbol{t}_k(\boldsymbol{x})\boldsymbol{t}_k(\boldsymbol{x})^T\,dG_k(\boldsymbol{x}), & \boldsymbol{P}_{kk'} &= \int\int\boldsymbol{t}_k(\boldsymbol{x})\boldsymbol{t}_{k'}(\boldsymbol{x}')^T\,dG_{kk'}(\boldsymbol{x},\boldsymbol{x}').\end{aligned}$$

The forms chosen for $\boldsymbol{h}(\cdot), G_k$ and $G_{kk'}$ allow us to express these integrals analytically. The general class of input distributions we use is described in Appendix B.

The closed forms for $\boldsymbol{R}_k$, $\boldsymbol{T}_k$, $\boldsymbol{S}_k$, $\boldsymbol{Q}_k$, $\boldsymbol{P}_k$ and U_k require only the marginal distribution G_k. To simplify the notation below we let $m_{k\_i}^* = \sum_{u=1}^{q\_i}\alpha_{\_i}^u m_{k\_i}^u$ denote the mean value of input i for output k.

The integrals are computed as follows:

$$\boldsymbol{R}_k=(1,m_{k\_1}^*\ldots,m_{k\_p}^*)^T,$$

$$[\boldsymbol{T}_k]_j = t_{kj} = \prod_{i=1}^{p} \sum_{u=1}^{q\_i} \alpha_{\_i}^{u} (1 + 2r_{k\_i} v_{k\_i}^{u})^{-1/2} \exp\left\{ -\frac{r_{k\_i}(s_{j,i} - m_{k\_i}^{u})^2}{1 + 2r_{k\_i} v_{k\_i}^{u}} \right\},$$

where $s_{j,i}$ denotes the ith element of the jth data input point $\boldsymbol{s}_j$,

$$\boldsymbol{Q}_k = \begin{pmatrix} 1 & m_{k\_1}^{*} & \dots & m_{k\_p}^{*} \\ m_{k\_1}^{*} & \sum_{u=1}^{q\_1} \alpha_{\_1}^{u} \{ v_{k\_1}^{u} + (m_{k\_1}^{u})^2 \} & m_{k\_1}^{*} m_{k\_2}^{*} \quad \dots & m_{k\_1}^{*} m_{k\_p}^{*} \\ \vdots & m_{k\_2}^{*} m_{k\_1}^{*} & \ddots & \vdots \\ & \vdots & & \vdots \\ m_{k\_p}^{*} & m_{k\_p}^{*} m_{k\_1}^{*} & \dots & \sum_{u=1}^{q\_p} \alpha_{\_p}^{u} \{ v_{k\_p}^{u} + (m_{k\_p}^{u})^2 \} \end{pmatrix}.$$

For $i = 1$ we have

$$[\boldsymbol{S}_k]_{ij} = t_{kj}$$

and for $h = 1, \dots, p$

$$\begin{aligned} [\boldsymbol{S}_k]_{h+1,j} \;=\; & \prod_{\substack{i=1 \\ i \neq h}}^{p} \sum_{u=1}^{q\_i} \alpha_{\_i}^{u} (1 + 2r_{k\_i} v_{k\_i}^{u})^{-1/2} \exp\left\{ -\frac{r_{k\_i}(m_{k\_i}^{u} - s_{j,i})^2}{1 + 2r_{k\_i} v_{k\_i}^{u}} \right\} \\ & \times \sum_{u=1}^{q\_h} \alpha_{\_h}^{u} (1 + 2r_{k\_h} v_{k\_h}^{u})^{-1/2} \exp\left\{ -\frac{r_{k\_h}(m_{k\_h}^{u} - s_{j,h})^2}{1 + 2r_{k\_h} v_{k\_h}^{u}} \right\} \frac{m_{k\_h}^{u} + 2r_{k\_h} v_{k\_h}^{u} s_{j,h}}{1 + 2r_{k\_h} v_{k\_h}^{u}}. \end{aligned}$$

The (h, j)th element of $\boldsymbol{P}_k$ is

$$\begin{aligned} [\boldsymbol{P}_k]_{hj} = & \prod_{i=1}^{p} \sum_{u=1}^{q\_i} \alpha_{\_i}^{u} (1 + 4r_{k\_i} v_{k\_i}^{u})^{-1/2} \\ & \times \exp\left[-\frac{1}{2} \left\{ r_{k\_i}(s_{h,i} - s_{j,i})^2 + \frac{4r_{k\_i}}{1 + 4r_{k\_i} v_{k\_i}^{u}} \left(m_{k\_i}^{u} - \frac{s_{h,i} + s_{j,i}}{2} \right)^2 \right\} \right]. \end{aligned}$$

The remaining integrals $\boldsymbol{Q}_{kk'}$, $\boldsymbol{S}_{kk'}$ and $\boldsymbol{P}_{kk'}$ depend on the joint distribution $G_{kk'}$ of inputs for two sites k, k', which is formulated in Appendix B.

$$\boldsymbol{Q}_{kk'} = \begin{pmatrix} 1 & m_{k'\_1}^{*} & \dots & m_{k'\_p-1}^{*} & m_{k'\_p}^{*} \\ m_{k\_1}^{*} & \Sigma_{\_1}^{*}(\tau_{\_1}) & m_{k\_1}^{*} m_{k'\_2}^{*} \quad \dots & & m_{k\_1}^{*} m_{k'\_p}^{*} \\ \vdots & m_{k\_2}^{*} m_{k'\_1}^{*} & \ddots & & \vdots \\ & \vdots & & & \\ & & & & \vdots \\ m_{k\_p-1}^{*} & & & \ddots & m_{k\_p-1}^{*} m_{k'\_p}^{*} \\ m_{k\_p}^{*} & & \dots & m_{k\_p}^{*} m_{k'\_p-1}^{*} & \Sigma_{\_p}^{*}(\tau_{\_p}) \end{pmatrix}$$

where $\Sigma^*_{\_i}(\tau_{\_i}) = \sum_{u=1}^{q_{\_i}} \sum_{u'=1}^{q_{\_i}} \alpha_{\_i}^{(uu')} (\tau_{\_i} \sqrt{(} v^u_{k\_i} v^{u'}_{k'\_i}) + m^u_{k\_i} m^{u'}_{k'\_i})$,

$$
\begin{aligned}
[\boldsymbol{S}_{kk'}]_{h+1,j} &= \prod_{\substack{i=1 \\ i \neq h}}^{p} \sum_{u=1}^{q_{\_i}} \alpha^u_{\_i} (1 + 2 r_{k'\_i} v^u_{k'\_i})^{-1/2} \exp\left\{ - \frac{r_{k'\_i} (m^u_{k'\_i} - s_{j,i})^2}{1 + 2 r_{k'\_i} v^u_{k'\_i}} \right\} \\
&\quad \times \sum_{u=1}^{q_{\_h}} \sum_{u'=1}^{q_{\_h}} \left[\alpha_{\_h}^{(uu')} (1 + 2 r_{k'\_h} v^{u'}_{k'\_h})^{-1/2} \exp\left\{ - \frac{r_{k'\_h} (m^{u'}_{k'\_h} - s_{j,h})^2}{1 + 2 r_{k'\_h} v^{u'}_{k'\_h}} \right\} \right. \\
&\quad \left. \times \left\{ m^u_{k\_h} + \frac{2 r_{k'\_h} \tau_{\_h} \sqrt{(} v^u_{k\_h} v^{u'}_{k'\_h}) (s_{j,h} - m^{u'}_{k'\_h})}{1 + 2 r_{k'\_h} v^{u'}_{k'\_h}} \right\} \right]
\end{aligned}
$$

for $h = 1, \ldots, p$. For $h = 1$ we have

$$
[\boldsymbol{S}_{kk'}]_{hj} = t_{k'j}. \tag{14}
$$

Finally, the (h, j)th element of $\boldsymbol{P}_{kk'}$ is

$$
\begin{aligned}
[\boldsymbol{P}_{kk'}]_{hj} &= \prod_{i=1}^{p} \sum_{u=1}^{q_{\_i}} \sum_{u'=1}^{q_{\_i}} \alpha_{\_i}^{(uu')} \left((1 + 2 r_{k\_i} v^u_{k\_i} (1 - \tau_{\_i})) (1 + 2 r_{k'\_i} v^{u'}_{k'\_i} (1 + \tau_{\_i})) - 2 \tau_{\_i} (r_{k'\_i} v^{u'}_{k'\_i} - r_{k\_i} v^u_{k\_i}) \right)^{-1/2} \\
&\quad \times \exp\left[- r_{k\_i} r_{k'\_i} / \{ (1 + 2 r_{k\_i} v^u_{k\_i} (1 - \tau_{\_i})) (1 + 2 r_{k'\_i} v^{u'}_{k'\_i} (1 + \tau_{\_i})) - 2 \tau_{\_i} (r_{k'\_i} v^{u'}_{k'\_i} - r_{k\_i} v^u_{k\_i}) \} \right. \\
&\quad \times [(1 + 2 r_{k'\_i} v^{u'}_{k'\_i}) (m^u_{k\_i} - s_{h,i})^2 / r_{k'\_i} + (1 + 2 r_{k\_i} v^u_{k\_i}) (m^{u'}_{k'\_i} - s_{j,i})^2 / r_{k\_i} \\
&\quad \left. - 4 \tau_{\_i} \sqrt{v^u_{k\_i} v^{u'}_{k'\_i}} (m^u_{k\_i} - s_{h,i}) (m^{u'}_{k'\_i} - s_{j,i})] \right].
\end{aligned}
$$

References

S. Conti and A. O'Hagan. Bayesian emulation of complex dynamic computer models. Technical report, Department of Probability and Statistics, University of Sheffield, Sheffield UK, 2006.

W. Cramer, A. Bondeau, F. I. Woodward, I. C. Prentice, R. A. Betts, V. Brovkin, P. M. Cox, V. Fisher, J. A. Foley, A. D. Friend, C. Kucharik, M. R. Lomas, N. Ramankutty, S. Sitch, B. Smith, and A. Young-Molling C. White. Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models. *Global Change Biology*, 7:357–373, 2001.

DEFRA. UK soil database for modelling soil carbon fluxes and land use for the national carbon dioxide inventory. http://www2.defra.gov.uk/research/project_data/More.asp?I=SP0511, 2003. Project SP0511.

C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 1997.

S.P. Evans, T. Randle, and P. Henshall. A simplified weather generator model for use in process modelling of forest system dynamics in the UK. Technical report, Forest Research, 2002.

Forestry Commision. National inventory of woodland and trees. http://www.forestry.gov.uk/forestry/hcou-54pg4d, 2002.

R.H. Haines-Young, C.J. Barr, H.I.J. Black, D.J. Briggs, R.G.H. Bunce, R.T. Clarke, A. Cooper, F.H. Dawson, L.G. Firbank, R.M. Fuller, M.T. Furse, M.K. Gillespie, R. Hill, M. Hornung, D.C. Howard, T. McCann, M.D. Morecroft, S. Petit, A.R.J. Sier, S.M. Smart, G.M. Smith, A.P. Stott, R.C. Stuart, and J.W. Watkins. Accounting for nature: assessing habitats in the UK countryside. *DETR, London ISBN 1 85112 460 8*, 2000.

R. G. Haylock and A. O'Hagan. On inference for outputs of computationally expensive algorithms with uncertainty on the inputs. In *Bayesian statistics, 5 (Alicante, 1994)*, pages 629–637. Oxford Univ. Press, New York, 1996.

M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B*, 63(3):425–464, 2001.

M. C. Kennedy, C. W. Anderson, S. Conti, and A. O'Hagan. Case studies in Gaussian process modelling of computer codes. *Reliability Engineering and System Safety.*, 2005. Forthcoming.

R. Milne and M. G. R. Cannell. Estimating forest and other terrestrial carbon fluxes at a national scale: The UK experience. In H. Griffiths and P. G. Jarvis, editors, *The Carbon Balance of Forest Biomes*, pages 57–75. Taylor & Francis, Abingdon, 2005.

T. D. Mitchell, T. R. Carter, P. D. Jones, M. Hulme, and M. New. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). Tyndall Centre Working Paper 55. http://www.tyndall.ac.uk/publications/working_papers/wp55_summary.shtml, 2004.

M. D. Morris and T. J. Mitchell. Exploratory designs for computational experiments. *J. Statist. Planng Inf.*, 43:381–402, 1995.

Northern Ireland Forest Service. Annual report 2000/2001. http://www.forestserviceni.gov.uk/publications/annual_reports/report00_01/00_01_link.htm, 2001.

J. E. Oakley and A. O'Hagan. Probabilistic sensitivity analysis of complex models: a Bayesian approach. *J. Roy. Statist. Soc. Ser. B*, 66:751–769, 2004.

A. O'Hagan and R. G. Haylock. Bayesian uncertainty analysis and radiological protection. In V. Barnett and K. F. Turkman, editors, *Statistics for the Environment 3, Pollution Assessment and Control*, pages 109–128. Wiley, Chichester, 1997.

T. Quaife. The impact of error in satellite derived land cover data on estimates of biospheric carbon fluxes predicted by a dynamic vegetation model. Technical report, Department of Geography, University College London, London, 2006.

S. Quegan. Calculating the uncertainty in calculations of carbon flux by dynamic vegetation models: general principles and a UK case study. Technical report, University of Sheffield, Sheffield UK, 2006.

J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statist. Sci.*, 4(4):409–435, 1989. With comments and a rejoinder by the authors.

A. N. Sharpley and J. R. Williams. EPIC-Erosion Productivity Impact Calculator, 1. model documentation. Technical report, U.S. Department of Agriculture, Agricultural Research Service, 1990. Tech. Bull. 1768.

F. I. Woodward and M. R. Lomas. Vegetation dynamics - simulating responses to climatic change. *Biol. Rev.*, 79:643–670, 2004.

F. I. Woodward, T. M. Smith, and W. R. Emanuel. A global primary productivity and phytogeography model. *Global Biogeochemical Cycles*, 9:471–490, 1995.