

Assimilation of global carbon cycle observations into an Earth system model to estimate uncertain terrestrial carbon cycle parameters

Daniel M. Ricciuto¹, Roman Tonkonojenkov², Nathan M. Urban², Richard D Wilkinson³, Damon Matthews⁴, Kenneth J. Davis⁵, and Klaus Keller²

- 1 Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA
- 2 Department of Geosciences, The Pennsylvania State University, University Park, PA, USA
- 3 Department of Probability and Statistics, University of Sheffield, Sheffield, UK
- 4 Department of Geography, Planning and Environment, Concordia University, Montreal, QC, CA
- 5 Department of Meteorology, The Pennsylvania State University, University Park, PA, USA

Abstract

Uncertainties about potential feedbacks in the terrestrial carbon cycle are a key driver of the uncertainty in carbon cycle and climate projections. Here we analyze how oceanic, atmospheric, and ice core carbon cycle observations improve key biogeochemical parameter estimates. We calibrate the University of Victoria Earth System Climate Model using atmospheric CO_2 concentration obtained from atmospheric and ice core measurements as well as estimates of decadal-scale oceanic carbon fluxes. We estimate the joint probability density function (PDF) of two key terrestrial feedback parameters quantifying carbon fertilization ($K_{\text{C}_{25}}$) and the respiration-temperature sensitivity (Q_{10}). To maximize computational efficiency, we use a relatively small ensemble of 48 model runs combined with a statistical emulation technique using principal component analysis. In the joint posterior PDF, there is a positive correlation between the two estimated parameters at low values of $K_{\text{C}_{25}}$ which reverses to a negative correlation at high values of $K_{\text{C}_{25}}$. This is a result of the nonlinearity of terrestrial carbon cycle feedbacks. CO_2 concentrations from Manua Loa are the most powerful constraint, and using a single CO_2 observation from the year 1999 can reproduce the general shape of the posterior PDF.

1. Introduction

Although terrestrial ecosystems currently absorb a considerable fraction of anthropogenic carbon emissions, the fate of this sink is highly uncertain. This uncertainty is caused by insufficient knowledge about key terrestrial carbon cycle/climate feedbacks, in particular the sensitivity of soil respiration to increasing global temperature. Modeling studies agree that the respiration-temperature sensitivity and other positive feedbacks are present in the climate system, but disagree on the mechanisms, timing and magnitude (Cox et al., 1999; Dufresne et al., 2002; Govindasamy et al., 2005; Zeng et al., 2004). These feedbacks will play a large role in regulating terrestrial carbon fluxes as climate change effects become more pronounced during this century. In the second half of the 21st century, models do not agree even on the sign of the net terrestrial carbon flux (Friedlingstein et al., 2006).

Although a compilation of results from different models is a useful gauge of cross-model uncertainty, there have been few studies of individual model uncertainty. An ensemble of runs from a single model would be an invaluable supplement; for example, such ensembles in the field of numerical weather prediction have led to improved forecasts and added value to information from deterministic forecasts (Toth & Kalnay, 1997; Molteni et al., 1996; Richardson, 2000). Fully coupled general circulation (GCM)/carbon cycle models generally employ state-of-the-art physics and have been tuned to reproduce past climate and carbon cycle observations accurately. However, in most terrestrial carbon cycle models, many key terrestrial processes are parameterized using empirical relationships because the underlying physical mechanisms are poorly

understood or too complex to model. Parameters derived in this way have uncertainties that are often not considered in modeling studies because of computational limitations. Ensembles of GCMs have demonstrated significant uncertainty in the degree of warming due to greenhouse gases (Stainforth et al., 2005; Piani et al., 2005; Murphy et al., 2004), but studies with fully coupled GCM/carbon cycle models have been limited to parameter sensitivity tests with limited numbers of model evaluations (Huntingford et al., 2008; Jones et al., 2003). While the cross-model uncertainty is well understood, most studies are silent on uncertainties within a single model.

To solve this problem, formal statistical model calibration techniques can help identify optimal parameter sets and parameter uncertainties that can be used make probabilistic predictions based on past observations. The computational cost of these complex models makes multidimensional parameter optimization extremely difficult, and this task is better suited for reduced complexity models. Model complexity, although crucial for progress in understanding underlying physical mechanisms, can be a liability for the tasks of validation against limited observations and future prediction; for example, it has been shown that simple, calibrated terrestrial carbon cycle models are capable of outperforming more complex models that have not been formally calibrated by regional to global scale data (Heimann et al., 1998). Earth System Models of Intermediate Complexity (EMICs) run much faster than fully coupled climate/carbon models, and therefore are ideal candidates for parameter optimization using large ensembles. An intercomparison of EMICs demonstrates a similar range of predictions to the fully coupled GCM/carbon cycle models (Plattner et al., 2008). EMICs have already been used in a number of parameter sensitivity studies to estimate key carbon cycle

uncertainties (Matthews et al., 2007; Eliseev & Mokhov, 2007). We build on these analyses by introducing a formal statistical method for calibrating an EMIC using historical observations of CO₂ and ocean fluxes. We calibrate two terrestrial carbon parameters that control two key feedbacks, respiration-temperature sensitivity (Q_{10}) and a parameter related to carbon fertilization (K_c) and test the following hypotheses:

1. The joint posterior PDF displays correlation because of the opposing effects of two key carbon terrestrial cycle feedbacks and the inability of observations to constrain them independently.
2. Including prior information about the distribution of Q_{10} significantly changes the shape of the posterior PDFs of both Q_{10} and K_c .

2. Methods

2.1 Model

We use the University of Victoria Earth System Climate Model version 2.8 (UVic ESCM), an intermediate complexity climate model comprised of general circulation ocean and dynamic/thermodynamic sea-ice components coupled to a vertically integrated energy/moisture balance atmosphere (Weaver et al., 2001). The land surface scheme is the Hadley Centre Met Office Surface Exchange Scheme (MOSES) (Cox et al., 1999). Also included in UVic is the Hadley Centre dynamic vegetation and terrestrial carbon cycle model Top-down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID) (Meissner et al., 2003; Matthews et al., 2005; Cox et al., 2000; Cox, 2001), and an inorganic ocean carbon cycle based on the OCMIP abiotic protocol (Ewen et al., 2004). Using an EMIC enables detailed studies of carbon cycle-climate

feedbacks with a much refined spatial resolution compared to simple box-models. The relative computational efficiency of EMICS combined with high-performance computing makes UViC an ideal candidate for data assimilation studies. This model has been used extensively in studies of terrestrial carbon cycle feedbacks (Matthews, 2007; Matthews & Keith, 2007; Matthews et al., 2005, 2007).

In this study, we calibrate two parameters in the TRIFFID submodel of UViC. TRIFFID explicitly models five plant functional types: broadleaf trees, needleleaf trees, C3 grasses, C4 grasses and shrubs. The five vegetation types are represented as a fractional coverage of each gridcell, and compete amongst each other for dominance as a function of the model simulated climate. In addition to simulating vegetation distributions, TRIFFID calculates terrestrial carbon stores and fluxes. The net terrestrial flux of carbon to the atmosphere can be calculated as the difference between soil respiration and net primary production (NPP). As nitrogen limitation on plant growth is fixed, terrestrial carbon uptake is determined primarily as a function of changing atmospheric carbon dioxide and climatic conditions. This global flux ($\text{kg m}^{-2} \text{ s}^{-1}$) is converted to or from ppmv using an atmospheric scale height of 8.5 km, which results in a conversion factor from GtC to ppmv of about 2.1. Carbon stores on land are represented by vegetation and soil carbon, and are updated by TRIFFID as a function of the flux of carbon to/from the atmosphere (a function of atmospheric CO₂ and climate) and changes in vegetation distributions (a function of climate). Atmospheric carbon dioxide is computed prognostically as a function of anthropogenic land use and fossil fuel emissions, ocean-atmosphere and land-atmosphere carbon fluxes. Anthropogenic emissions of carbon impose a perturbation to the equilibrium spin-up state, and the land

and ocean carbon pools and fluxes respond dynamically to the imposed increase in atmospheric carbon dioxide. The computed atmospheric carbon dioxide concentration is the resulting CO₂ in the atmosphere after global carbon sinks have responded to anthropogenic emissions.

The key feedbacks examined in this study are heterotrophic respiration/temperature and carbon fertilization. Heterotrophic respiration in TRIFFID is controlled by a single carbon pool. The respiration rate depends on the size of this carbon pool, soil moisture and soil temperature. Respiration is calculated as:

$$R_H = \kappa_H C_s f(\theta) Q_{10}^{(T_s - 25^\circ C)/10} \quad (1)$$

Where C_s is the soil carbon (kg m⁻²), κ_s is the respiration rate per unit soil carbon at 25C, $f(\theta)$ is a dimensionless function modifying respiration based on the soil moisture θ , T_s is the soil temperature and Q_{10} is a parameter controlling the temperature dependence of respiration. The default value of Q_{10} in TRIFFID is 2.0, indicating a doubling of the respiration rate for every 10C increase in temperature. In this study, we vary the value of Q_{10} from a value of 1.0 to 4.0, which is comparable to a range of values observed across a wide array of ecosystem types (Fierer et al., 2006).

TRIFFID employs a Farquhar photosynthesis model in which the rate of photosynthesis when light is not limiting is controlled by the maximum carboxylation rate of Rubisco. This rate, W_c , is controlled by temperature, humidity, leaf nitrogen and atmospheric CO₂ concentration through a Michaelis-Menten relationship:

$$W_c = V_m(T_l, n_l) \left[\frac{c_i - \Gamma}{c_i + K_c \left(1 + \frac{O_a}{K_o}\right)} \right] \quad (2)$$

where V_m , the maximum carboxylation rate of Rubisco, is a function of T_l , the leaf temperature and n_l , the leaf nitrogen concentration. c_i is the internal leaf CO_2 concentration and K_c and K_o are the Michaelis-Menten constants for CO_2 and O_2 respectively. O_a is the partial pressure of atmospheric oxygen. c_i is related to the external CO_2 concentration c_c through the following closure scheme:

$$\left[\frac{c_i - \Gamma}{c_c - \Gamma} \right] = F_0 \left[1 - \frac{D_*}{D_c} \right] \quad (3)$$

Where D_* is the humidity deficit at the leaf surface and F_0 and D_c are PFT-specific parameters. The photorespiration compensation point Γ is the value of c_i at which photosynthesis balances photorespiration, and it is equal to 5.7 Pa at 25°C. Carbon fertilization is most strongly controlled by the K_c parameter, which is modified by air temperature using the following relationship:

$$K_c = K_{c_{25}} * 2.1^{(T-25)/10} \quad (4)$$

Where $K_{c_{25}}$ is the value of K_c at 25°C, which is 30 Pa in the default version UVic 2.8. The actual rate of gross photosynthesis is determined by taking a smoothed minimum of three limiting rates, including W_c , W_l (light limitation) and W_e (transport limitation of photosynthetic products for C_3 plants and PEP-carboxylase limitation for C_4 plants).

In this study, we calibrate the parameter $K_{c_{25}}$. Changing this parameter affects both the rate of gross photosynthesis and the dependence of this rate on atmospheric CO_2

concentration. At a given CO₂ concentration, a higher value of K_{c25} will result in lower photosynthesis rates; however, as CO₂ increases, this higher value of K_{c25} causes a larger carbon fertilization effect in which increased atmospheric CO₂ concentrations will result in larger relative uptake.

2.2 Observational Constraints

The model is constrained by observations of CO₂ concentrations and estimates of global ocean carbon fluxes. CO₂ concentration constraints include data from the Law Dome ice core between 1850-1959 (Etheridge et al., 1996; MacFarling Meure et al., 2006), and from the Mauna Loa Observatory from 1960-2004 (Keeling & Whorf, 2005). Both sites were chosen as proxies for the global average. The Law Dome dataset is an irregular time-series in which data points represent different samples of air in the ice core with different ages. We assume that the observational uncertainty of each estimate is independent and normally distributed with standard deviations of 8 ppm for the ice core data and 0.4 ppm for the Mauna Loa data. The mean air age of the sample was taken as the year of observation for the purpose of model evaluation. The annual Mauna Loa CO₂ used in the assimilation contains 40 annual data points (1960-1999).

Independent estimates of oceanic sink strength are also used to constrain our model. The chlorofluorocarbon (CFC) dataset, which was combined with measured DIC to estimate average annual uptake over the decades of the 1980s and the 1990s. McNeil *et al.* (2003) estimate an average annual ocean-wide uptake of 1.6 +/- 0.4 Gt C yr⁻¹ over the 1980s, and 2.0 +/- 0.4 Gt C yr⁻¹ over the 1990s. To compare the model output to

these estimates, modeled annual ocean fluxes were averaged over each decade. Because of the limited number of data points, we do not solve for process error or autocorrelation in the oceanic sink observations. The total error of these observations is considered to be normally distributed with a standard deviation equal to the reported observational error.

2.3 Experimental design

We use a 48-member ensemble of parameter sets of (Q_{10} , K_c) in the fully coupled UVic model with increased sampling density near the maximum likelihood solution (Fig. 1). For each ensemble member, we perform an equilibration run for 1,000 years using atmospheric CO_2 concentrations, solar forcing and land use data from the year 1800. At the end of the equilibration run, annual net primary productivity (NPP) and heterotrophic respiration (R_h) are approximately in balance, and annual global ocean carbon fluxes are near zero. Transient model runs are then performed from the year 1800 to 2000. Forcing for transient model runs included solar and orbital forcing (Berger, 1978), anthropogenic greenhouse gas emissions from fossil fuel burning, cement production and land use change. The land use change emissions are based on a maximum likelihood weighting of three land use change emissions estimates (Jain & Yang, 2005; Houghton, 2003) using a simple carbon cycle model constrained by CO_2 concentration and ocean flux observations (Miltich et al., in review).

2.4 Optimization technique

The aim of this study is to calibrate probabilistically the UVic model parameters K_c and Q_{10} , assigning higher probabilities to parameters for which the model output

better matches the observational constraints. We use a Bayesian data assimilation method to estimate the joint posterior PDF of these parameters. A Bayesian approach combines prior information about model parameters with the information contained in observations to derive this PDF. Markov Chain Monte Carlo (MCMC) is a powerful method for estimating PDFs that makes no structural assumptions about posterior PDFs and has been used with simple global-scale earth-system models to estimate parameters (Hargreaves & Annan, 2002; Ricciuto et al., 2008; Urban & Keller, 2008).

The usual MCMC approach to Bayesian inversion requires a dense sampling of the posterior density function. In our case, model output can be generated for only a small ensemble of runs at a limited number of parameter settings, making a direct application of the MCMC method infeasible. In order to apply MCMC assimilation techniques we construct a fast approximation, or emulator, of the UVic model. An emulator predicts what UVic output would be at a parameter setting at which the model was not run. Its prediction is an interpolation of the output from other ensemble members at related parameter values.

Our Gaussian process statistical emulator (Kennedy & O'Hagan, 2001; Rasmussen & C. Williams, 2006) is a generalization of the geostatistical interpolation method known as kriging (Cressie, 1993; Banerjee et al., 2003). Instead of the usual kriging practice of interpolating a measured quantity over physical space, a statistical emulator interpolates model output over parameter space. The main advantages of statistical emulators over other interpolation methods (e.g., linear interpolation, splines, parametric regression, neural networks) are that (i) emulators provide an estimate of the uncertainty in their predictions, and (ii) they make no strong assumptions about the

functional form of the model output. For data assimilation, statistical emulators which estimate their own errors are thus preferable to non-statistical emulators which do not. If emulator uncertainty is ignored, the resulting parameter estimates can be overconfident. Overconfident estimates can be produced if the assimilation erroneously excludes parameter settings for which the model fits the data but the emulator does not. Other statistical approaches such as parametric regression can also estimate their own prediction errors, but at the expense of making strong assumptions about what form the model output can take (e.g., smooth polynomial behavior). Gaussian process emulators do not make parametric assumptions and can fit model output with irregular local structure.

We construct independent emulators for atmospheric CO₂ concentrations and air-sea CO₂ fluxes, trained on the 48-member ensemble of UVic output in Q₁₀/Kc₂₅ space. To simplify the emulator, we first reduce the dimensionality of the CO₂ time series. By taking the leading four principal components of the time series, we reduce the emulated output from 57 data points (corresponding to 40 instrumental and 17 ice core measurements) to 4 data points. We use six emulators in total, each of which predicts a single scalar quantity as a function of Q₁₀ and Kc₂₅. Four of the emulators predict the loadings of the four leading principal components of the atmospheric CO₂ time series. The other two emulators predict the decadal averaged ocean CO₂ flux in the 1980s and 1990s. Details of the emulator construction and inversion may be found in the Appendix.

Bayesian inversion combines the emulator predictions with observations to estimate the model parameters. The observations are assumed to be a sum of the model output plus process and observation error, giving a likelihood function for the data

conditioned on the model parameters. The observation errors are described in Section 2.2. The process error for the CO₂ time series is assumed to be a normal first-order autoregressive AR(1) process of unknown variance and autocorrelation. The process error for the fluxes is ignored since the observation errors are large. The CO₂ concentration and flux errors are assumed independent of each other.

The inversion jointly estimates the posterior PDF of 10 parameters: the two model parameters (Q_{10} and K_{c25}), the variance and autocorrelation of the AR(1) process, the variance in emulator output for each of the four CO₂ principal component loadings, and the variance of the two flux emulators. To produce a posterior PDF for the two model parameters, we integrate the posterior distribution over the other eight "nuisance" parameters. This integration, known as marginalization, gives the reduced joint posterior for Q_{10} and K_{c25} . We also obtain one-dimensional marginal posteriors for Q_{10} and K_{c25} individually.

3. Results and discussion

The 48 UVic ensemble members produce a wide variation in predictions of CO₂ concentrations and ocean fluxes (Fig. 2). Most of the ensemble members fall within the range of uncertainty of the two separate ocean flux estimates (Figs 2b, 2c), but most of the members are out of the range of uncertainty of measured CO₂ concentrations, especially during the period of the Mauna Loa record when uncertainties are smaller (Fig 2a). The observations therefore add significant information about these carbon cycle parameters.

3.1 Posterior probability density functions

The joint probability surface (Fig. 3) and marginal PDFs for each parameter (Fig. 4) are heavily influenced by the choice of prior distribution for the Q_{10} parameter. When a uniform prior is used, the most likely values of K_{c25} and Q_{10} are their lowest allowable values of 7.5 Pa and 1.0 respectively. Observational studies of soil carbon efflux suggest that values of Q_{10} at study sites across many ecosystem types range from 2.2 to 4.6 (Fierer et al., 2006). We compile the 77 observed Q_{10} values into a PDF to be used as an informative prior distribution to compare to the base case of a uniform prior distribution (Fig. 4a). Using this informative Q_{10} prior shifts the maximum likelihood values for both K_{c25} and Q_{10} upwards to 37.5 Pa and 2.75 respectively. It is unknown whether these site-based studies can actually be useful to infer to a single value of Q_{10} or whether using a global value of Q_{10} can produce meaningful results. We use this example to stress the importance of prior information on the final result.

K_{c25} and Q_{10} display a strong correlation regardless of the Q_{10} prior distribution (Fig. 3), although this correlation is more apparent when the uniform Q_{10} prior is used. At values of K_{c25} less than 30 Pa and values of Q_{10} less than 2.5, we observe a linear correlation in the joint posterior PDF with a slope of roughly 15 Pa K_{c25} per unit of Q_{10} (Fig. 3a). The correlation represents the inability of the observational constraints to determine the relative strengths of terrestrial carbon cycle feedbacks. If estimates of anthropogenic emissions are reliable, the strength of the terrestrial carbon sink is relatively well constrained by the atmospheric CO_2 record and ocean flux estimates. Because carbon fertilization and the respiration – temperature feedbacks oppose each

other, their relative strengths are not well constrained; for example, $K_{C_{25}}$ of 10 Pa and a Q_{10} of 1.0 produces similar output to that from a $K_{C_{25}}$ of 25 Pa and a Q_{10} of 2.0.

At values of $K_{C_{25}}$ greater than 30 Pa, we observe a “hook” effect in which the maximum likelihood Q_{10} decreases (Fig. 3). This is because at high values of $K_{C_{25}}$, carbon fertilization saturates and then begins to decrease. This is evident when we examine effective values of the carbon fertilization parameter β as a function of $K_{C_{25}}$ and Q_{10} using the following relationship:

$$\beta = \frac{\frac{NPP}{NPP_0} - 1}{\log\left(\frac{[CO_2]}{[CO_2]_0}\right)} \quad (5)$$

Where NPP is the net primary productivity in the year 1999, NPP_0 is the net primary productivity in the year 1800, $[CO_2]$ is the atmospheric CO_2 concentration in 1999 and $[CO_2]_0$ is the atmospheric CO_2 concentration in 1800. Values of β range from 0.35 to 0.6 (Fig. 5), which is similar to values observed in other terrestrial carbon cycle models (Kicklighter et al., 1999). In UVic, NPP is also affected by autotrophic respiration, temperature, precipitation and vegetation distribution changes so that this effective β value is not purely driven by CO_2 fertilization. Although increasing $K_{C_{25}}$ increases the instantaneous carbon fertilization rate of gross photosynthesis as described by equation (2), other factors limit the increase in NPP over long timescales at high values of $K_{C_{25}}$. Because at high $K_{C_{25}}$, the carbon fertilization feedback decreases with increasing $K_{C_{25}}$,

lower values of Q_{10} are required to balance the carbon cycle in a way that matches observations.

This nonlinearity is compounded by differences in the equilibrium distributions of vegetation and soil carbon caused by differences Q_{10} and K_{c25} (Fig. 6). Equilibrium vegetation carbon decreases linearly with K_{c25} while equilibrium soil carbon increases with Q_{10} and decreases with K_{c25} . Large values of Q_{10} and low values of K_{c25} are therefore associated with high rates of carbon efflux, resulting in unrealistic solutions of a terrestrial carbon source in 1999 and atmospheric CO_2 concentrations above 420 ppm. Equilibrium soil carbon at high values of Q_{10} is higher because the turnover rates of soil carbon in northern high latitudes are much lower, leading to an increased respiration-temperature feedback because a) high latitudes experience earlier and stronger warming, b) more soil carbon is available to respire, and c) the sensitivity to increased temperature is higher. Increasing values of Q_{10} above 2.5 result in increased soil respiration that cannot be balanced by a corresponding increase in carbon fertilization (K_c) in a way that is consistent with the observational records. Conversely, low values of Q_{10} and moderate to high values of K_{c25} are associated with low equilibrium soil carbon and too much terrestrial uptake. Given the model structure of UVic and observational constraints, we conclude that high values of Q_{10} are unlikely because there is not enough carbon fertilization in the model, regardless of the K_{c25} value, to balance this level of respiration-temperature sensitivity.

3.2 Effect of individual constraints

Atmospheric CO₂ concentrations provide most of the constraint on the joint posterior PDF of Kc₂₅ and Q₁₀ (Fig 7). The last year of the Manua Loa record used in this study, 1999, provides enough information alone to reproduce the general shape of the joint posterior PDF. The ice core CO₂ concentrations (Fig. 7c) and decadal annual average air-sea fluxes (Fig. 7d) provide a much weaker constraint that only excludes combinations of low Q₁₀ and high Kc₂₅.

4. Conclusions

We present a statistically sound and computationally feasible methodology for the assimilation of observations into Earth system models of intermediate complexity. Using an ensemble of 48 model evaluations, we estimate the joint probability density function of two key carbon cycle feedback parameters constrained by observations of globally averaged atmospheric CO₂ concentrations and estimates of global ocean CO₂ fluxes. The respiration temperature sensitivity parameter Q₁₀ and carbon fertilization parameter Kc₂₅ are positively correlated at low values of Kc₂₅ and negatively correlated at high values of Kc₂₅. This behavior is a result of nonlinearities in the response of carbon fertilization to CO₂ and the response of respiration to temperature as well as differences in the equilibrium distributions of vegetation and soil carbon resulting from difference in Kc₂₅ and Q₁₀. Given the model structure, forcing data, constraints and uniform priors, high values of Q₁₀ are excluded but no value of Kc₂₅ can be ruled out. If an observationally based prior of Q₁₀ is used, low values of Q₁₀ and Kc₂₅ can also be excluded. CO₂ concentrations provide the best constraint on the joint posterior PDF. The acceptable

parameter combinations produce a similar response of the terrestrial carbon cycle during the year 1800-1999, but are expected to produce widely divergent future predictions.

Acknowledgements

We thank the UVic modeling community for providing and supporting such an excellent research tool. Comments by Murali Haran, Sham Bhat, and Josh Dorin helped to improve the paper. Any potential errors and omissions are, of course, ours. Financial support from the Penn State Earth and Environmental Systems Institute is gratefully acknowledged. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding entity.

References

- An, J, and A Owen (2001), Quasi-regression, *JOURNAL OF COMPLEXITY*, 17(4), 588-607.
- Banerjee, S, BP Carlin, et al. (2003), *Hierarchical modeling and analysis for spatial data*, Chapman & Hall, CRC.
- Berger, AL (1978), Long-term variations of caloric insolation resulting from earth's orbital elements, *Quaternary Research*, 9(2), 139-167.
- Cox, PM (2001), Description of the TRIFFID Dynamic Global Vegetation Model, *Hadley Centre technical note 24*.
- Cox, PM, RA Betts, et al. (1999), The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, *CLIMATE DYNAMICS*, 15(3), 183-203.
- Cox, PM, RA Betts, et al. (2000), Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model (vol 408, pg 184, 2000), *NATURE*, 408(6813), 750-750.
- Cressie, NAC (1993), *Statistics for spatial data*, Wiley, New York, NY.
- Dufresne, JL, P Friedlingstein, et al. (2002), On the magnitude of positive feedback between future climate change and the carbon cycle, *GEOPHYSICAL RESEARCH LETTERS*, 29(10).
- Eliseev, AV, and II Mokhov (2007), Carbon cycle-climate feedback sensitivity to parameter changes of a zero-dimensional terrestrial carbon cycle scheme in a climate model of intermediate complexity, *THEORETICAL AND APPLIED CLIMATOLOGY*, 89(1-2), 9-24.
- Etheridge, DM, LP Steele, et al. (1996), Natural and anthropogenic changes in atmospheric CO₂ over the last 1000 years from air in Antarctic ice and firn, *JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES*, 101(D2), 4115-4128.
- Ewen, TL, AJ Weaver, et al. (2004), Sensitivity of the inorganic ocean carbon cycle to future climate warming in the UVic coupled model, *ATMOSPHERE-OCEAN*, 42(1), 23-42.
- Fierer, N, BP Colman, et al. (2006), Predicting the temperature dependence of microbial respiration in soil: A continental-scale analysis, *GLOBAL BIOGEOCHEMICAL CYCLES*, 20(3).

- Friedlingstein, P, P Cox, et al. (2006), Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison, *JOURNAL OF CLIMATE*, 19(14), 3337-3353.
- Govindasamy, B, S Thompson, et al. (2005), Increase of carbon cycle feedback with climate sensitivity: results from a coupled climate and carbon cycle model, *TELLUS SERIES B-CHEMICAL AND PHYSICAL METEOROLOGY*, 57(2), 153-163.
- Hargreaves, JC, and JD Annan (2002), Assimilation of paleo-data in a simple Earth system model, *CLIMATE DYNAMICS*, 19(5-6), 371-381.
- Heimann, M, G Esser, et al. (1998), Evaluation of terrestrial Carbon Cycle models through simulations of the seasonal cycle of atmospheric CO₂: First results of a model intercomparison study, *GLOBAL BIOGEOCHEMICAL CYCLES*, 12(1), 1-24.
- Higdon, D, J Gattiker, et al. (2008), Computer model calibration using high-dimensional output, *JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION*, 103(482), 570-583.
- Houghton, RA (2003), Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850-2000, *TELLUS SERIES B-CHEMICAL AND PHYSICAL METEOROLOGY*, 55(2), 378-390.
- Huntingford, C, RA Fisher, et al. (2008), Towards quantifying uncertainty in predictions of Amazon 'dieback', *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, 363(1498), 1857-1864.
- Jain, AK, and XJ Yang (2005), Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with CO₂ and climate change, *GLOBAL BIOGEOCHEMICAL CYCLES*, 19(2).
- Jones, CD, P Cox, et al. (2003), Uncertainty in climate-carbon-cycle projections associated with the sensitivity of soil respiration to temperature, *TELLUS SERIES B-CHEMICAL AND PHYSICAL METEOROLOGY*, 55(2), 642-648.
- Keeling, CD, and TP Whorf (2005), Atmospheric CO₂ records from sites in the SIO air sampling network,, *Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.*
- Kennedy, MC, and A O'Hagan (2001), Bayesian calibration of computer models, *JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B-STATISTICAL*, 63, 425-450.

- Kicklighter, DW, M Bruno, et al. (1999), A first-order analysis of the potential role of CO₂ fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models, *TELLUS SERIES B-CHEMICAL AND PHYSICAL METEOROLOGY*, 51(2), 343-366.
- MacFarling Meure, C, D Etheridge, et al. (2006), Law Dome CO₂, CH₄ and N₂O ice core records extended to 2000 years BP, *GEOPHYSICAL RESEARCH LETTERS*, 33(14), doi:10.1029/2006GL026152.
- Matthews, HD (2007), Implications of CO₂ fertilization for future climate change in a coupled climate-carbon model, *GLOBAL CHANGE BIOLOGY*, 13(5), 1068-1078.
- Matthews, HD, M Eby, et al. (2007), What determines the magnitude of carbon cycle-climate feedbacks?, *GLOBAL BIOGEOCHEMICAL CYCLES*, 21(2).
- Matthews, HD, and DW Keith (2007), Carbon-cycle feedbacks increase the likelihood of a warmer future, *GEOPHYSICAL RESEARCH LETTERS*, 34(9), doi:10.1029/2006GL028685.
- Matthews, HD, AJ Weaver, et al. (2005), Terrestrial carbon cycle dynamics under recent and future climate change, *JOURNAL OF CLIMATE*, 18(10), 1609-1628.
- McNeil, BI, RJ Matear, et al. (2003), Anthropogenic CO₂ uptake by the ocean based on the global chlorofluorocarbon data set, *SCIENCE*, 299(5604), 235-239.
- Meissner, KJ, AJ Weaver, et al. (2003), The role of land surface dynamics in glacial inception: a study with the UVic Earth System Model, *CLIMATE DYNAMICS*, 21(7-8), 515-537.
- Molteni, F, R Buizza, et al. (1996), The ECMWF ensemble prediction system: Methodology and validation, *QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY*, 122(529), 73-119.
- Murphy, JM, DMH Sexton, et al. (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations, *NATURE*, 430(7001), 768-772.
- O'Hagan, A (2006), Bayesian analysis of computer code outputs: A tutorial, *RELIABILITY ENGINEERING & SYSTEM SAFETY*, 91(10-11), 1290-1300.
- Piani, C, DJ Frame, et al. (2005), Constraints on climate change from a multi-thousand member ensemble of simulations, *GEOPHYSICAL RESEARCH LETTERS*, 32(23).
- Plattner, GK, R Knutti, et al. (2008), Long-term climate commitments projected with climate-carbon cycle models, *JOURNAL OF CLIMATE*, 21(12), 2721-2751.

- Rasmussen, CW, and CKI Williams (2006), *Gaussian Processes for Machine Learning*, MIT Press.
- Ricciuto, DM, MP Butler, et al. (2008), Causes of interannual variability in ecosystem-atmosphere CO₂ exchange in a northern Wisconsin forest using a Bayesian model calibration, *AGRICULTURAL AND FOREST METEOROLOGY*, 148(2), 309-327.
- Richardson, DS (2000), Skill and relative economic value of the ECMWF ensemble prediction system, *QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY*, 126(563), 649-667.
- Sabine, CL, RA Feely, et al. (2004), The oceanic sink for anthropogenic CO₂, *SCIENCE*, 305(5682), 367-371.
- Sacks, J, WJ Welch, et al. (1989), Design and analysis of computer experiments, *Statistical Science*, 4, 409-423.
- Santner, TJ, BJ Williams, et al. (2003), *The design and analysis of computer experiments*, Springer.
- Stainforth, DA, T Aina, et al. (2005), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *NATURE*, 433(7024), 403-406.
- Toth, Z, and E Kalnay (1997), Ensemble forecasting at NCEP and the breeding method, *MONTHLY WEATHER REVIEW*, 125(12), 3297-3319.
- Urban, NM, and K Keller (2008), Probabilistic hindcasts and projections of the coupled climate, carbon cycle, and Atlantic meridional overturning circulation systems: A Bayesian fusion of century-scale observations with a simple model, *Tellus A*, in review.
- Weaver, AJ, M Eby, et al. (2001), The UVic Earth System Climate Model: Model description, climatology, and applications to past, present and future climates, *ATMOSPHERE-OCEAN*, 39(4), 361-428.
- Zeng, N, HF Qian, et al. (2004), How strong is carbon cycle-climate feedback under global warming?, *GEOPHYSICAL RESEARCH LETTERS*, 31(20), doi:10.1029/2004GL020904.

Figure Captions

Figure 1: Design of the ensemble in Q10/Kcmult parameter space. Depicted are the 48 ensemble members.

Figure 2: UVic model predictions for all of (a) atmospheric CO₂ concentrations and (b) annual average ocean carbon uptake. Red circles with error bars indicate observations. Predictions for all ensemble members are shown.

Figure 3: Joint posterior probability distribution of Q10 and Kcmult. *Left:* Uniform prior for Q10. *Right:* Informative prior for Q10 based on laboratory experiments.

Figure 4: Marginal posterior probability distributions of Q10 (left) and Kcmult (right). Depicted are posteriors based on a uniform Q10 prior (black) and the informative Q10 prior (blue). The blue dashed curve is the informative prior.

Figure 5: Effective logarithmic beta fertilization factor relating changes in net primary productivity from the year 1800-1999 to the change in CO₂ concentrations from the years 1800-1999.

Figure 6: Equilibrium total terrestrial carbon pool including vegetation and soil carbon (a), change in terrestrial carbon pool size from 1800-1999 (b), equilibrium vegetation carbon (c), change in vegetation carbon from 1800-1999, (d), equilibrium soil carbon (e) and change in soil carbon from 1800-1999 (f); contours represent values in GtC.

Figure 7: Joint posterior probability distributions of Q_{10} and K_{cmult} from assimilations of subsets of the observational constraints. Uniform priors are assumed. *Upper left*: all constraints (same as Fig. 2, left panel). *Upper right*: instrumental atmospheric CO_2 concentrations. *Lower left*: ice core atmospheric CO_2 concentrations. *Lower right*: air-sea CO_2 fluxes.

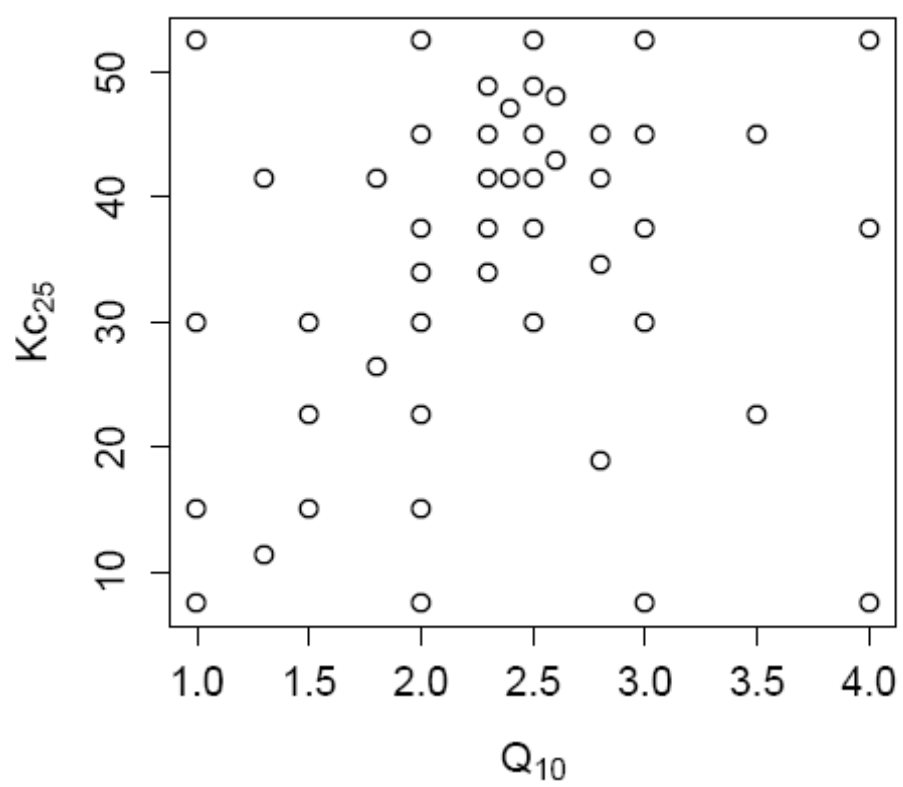


Figure 1

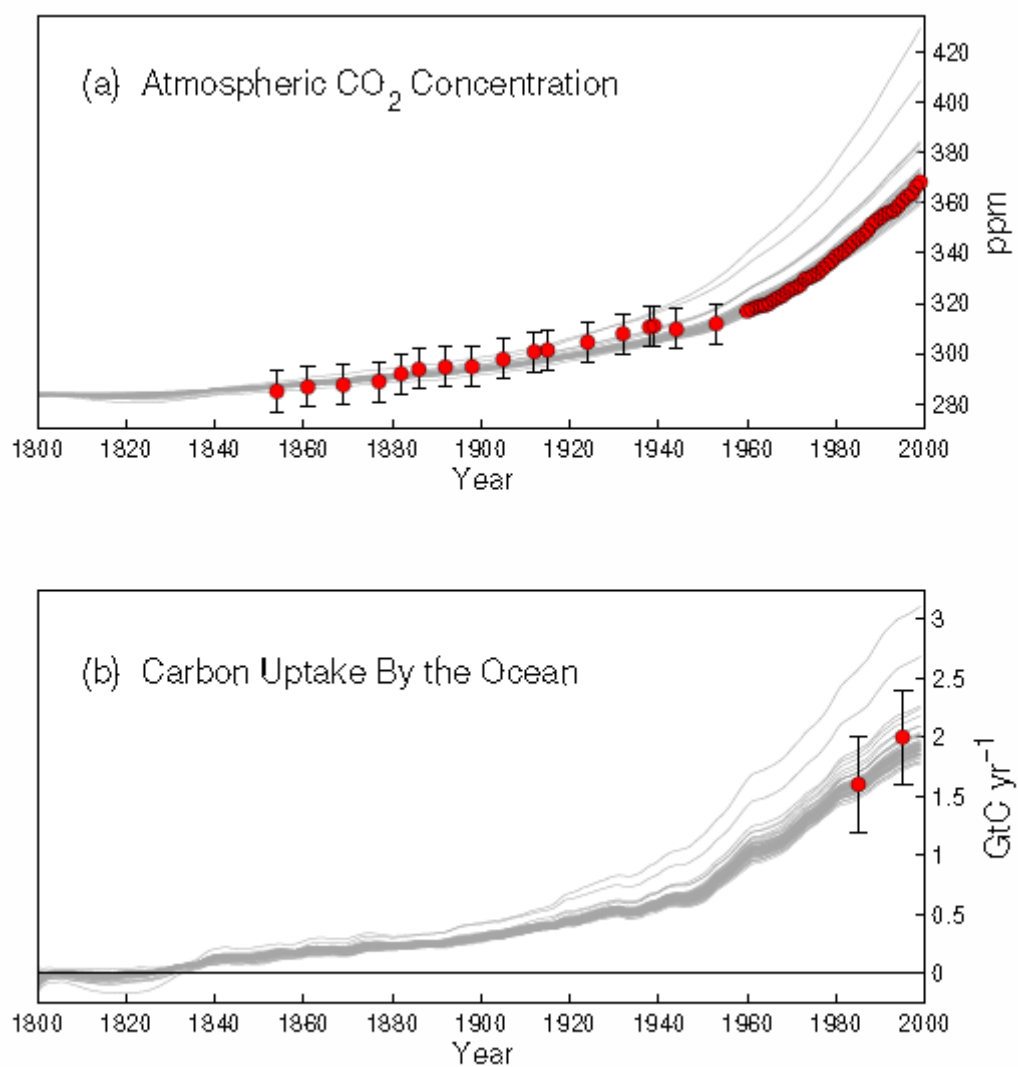


Figure 2

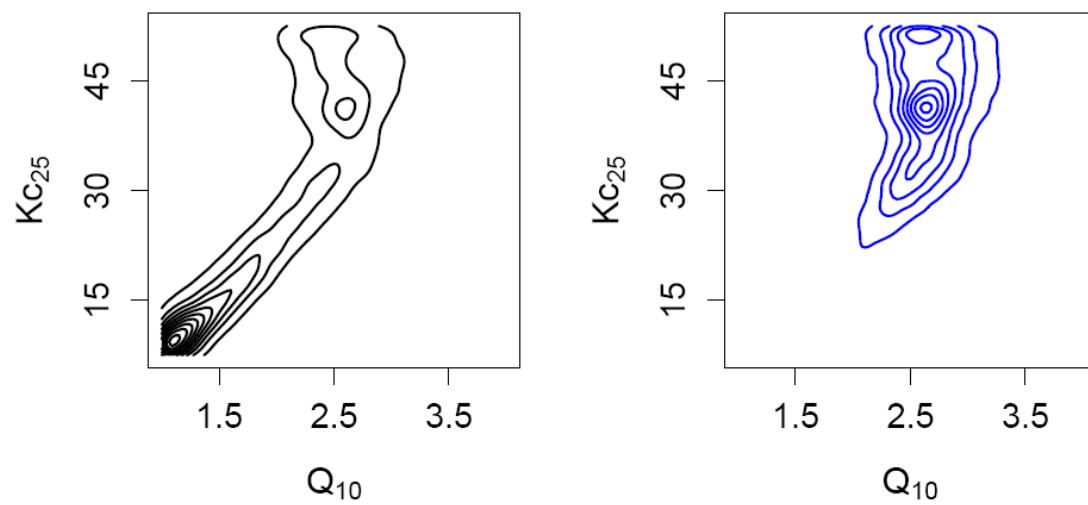


Figure 3

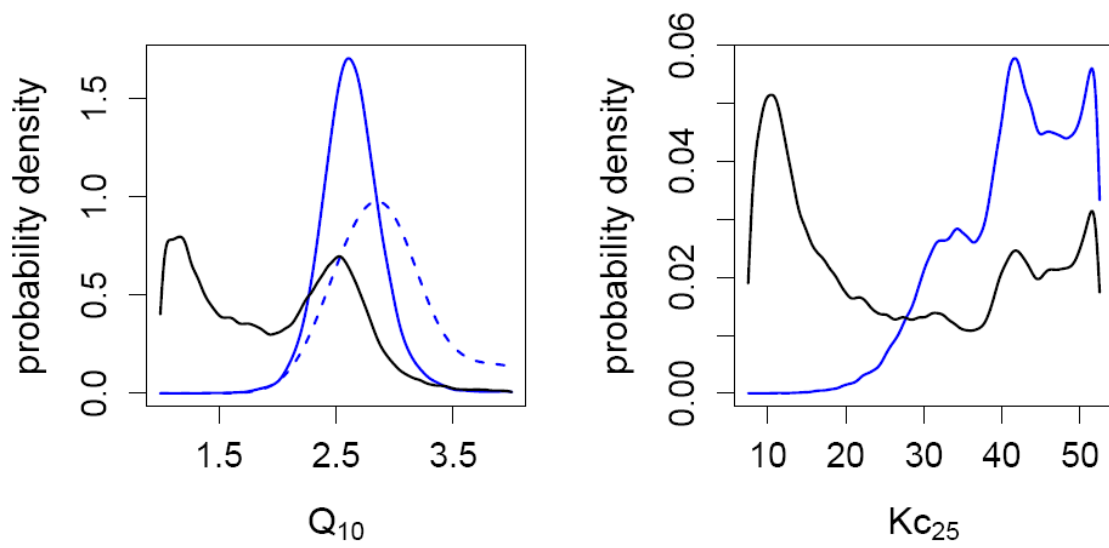


Figure 4

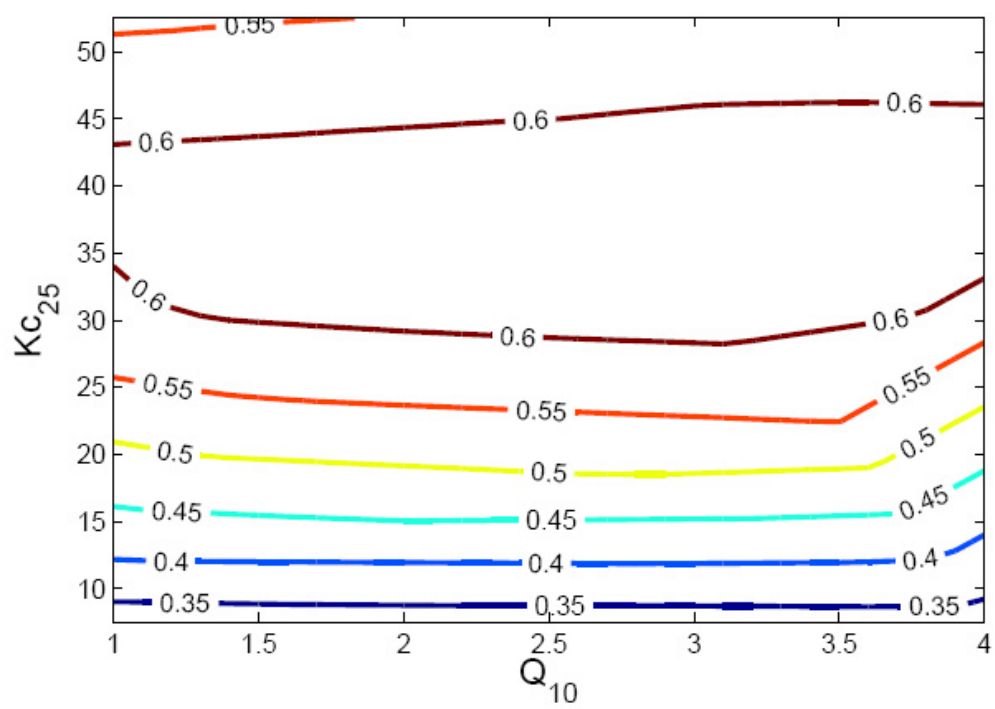


Figure 5

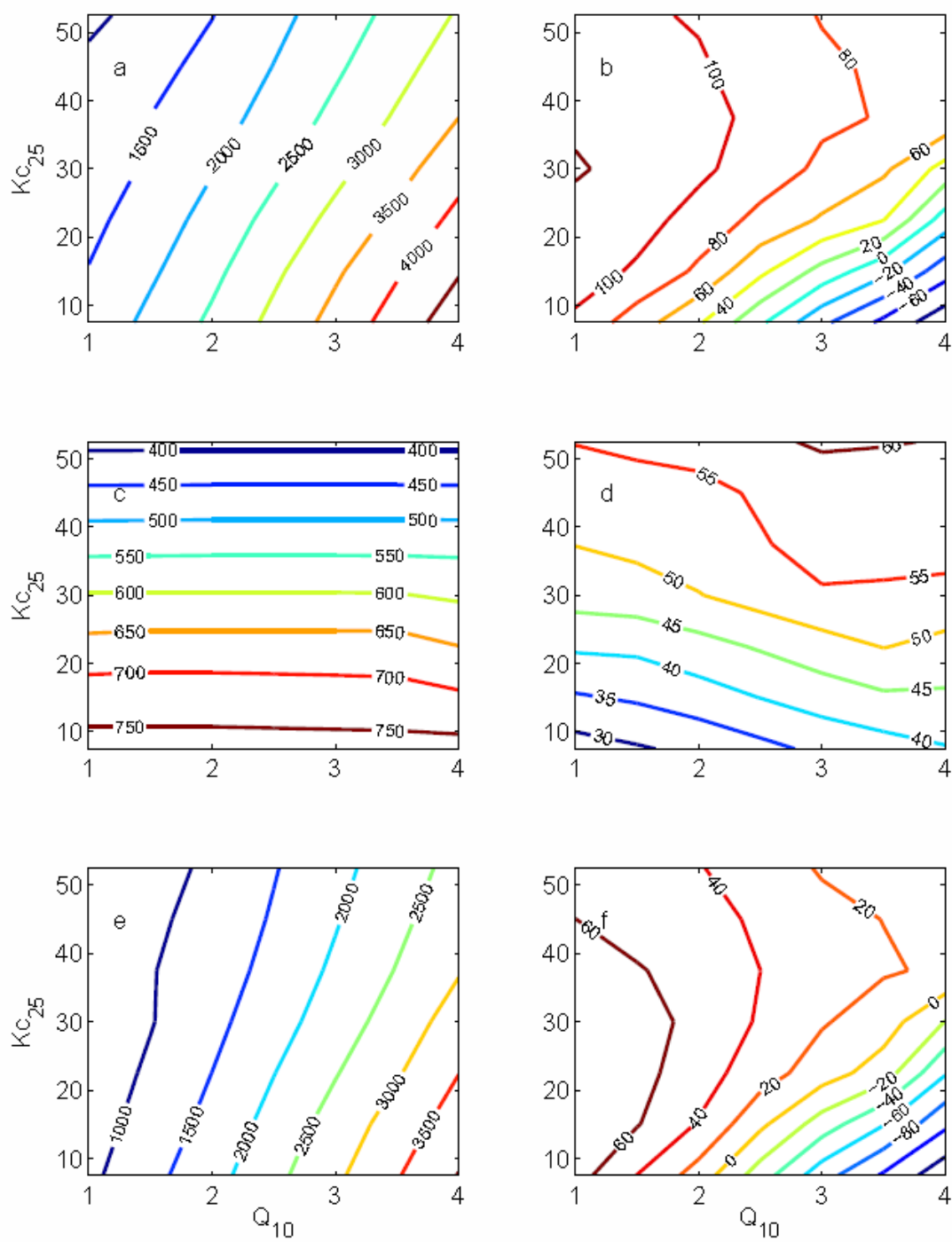


Figure 6

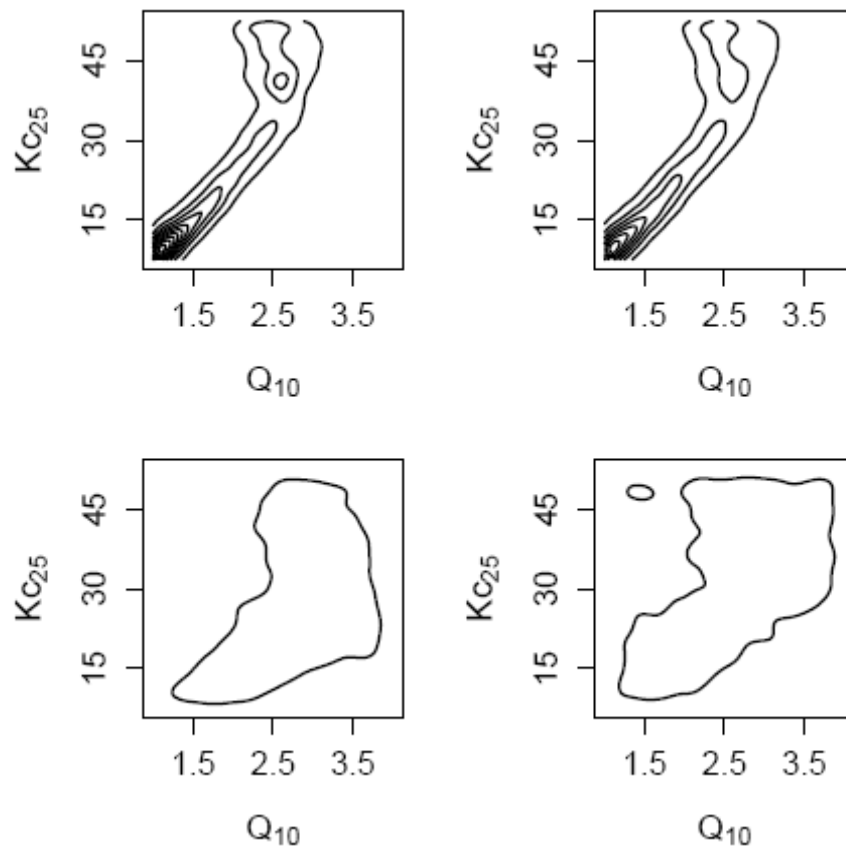


Figure 7

Appendix

To find the posterior distribution of the calibration parameters given the field measurements Z and the computer model $\eta(\cdot)$, we must posit a relationship between them. We use the framework described in Kennedy and O'Hagan (2001), namely that

$$Z = \eta(\theta) + \delta + \varepsilon.$$

Here Z represents the vector of observations, both CO₂ readings and flux measurements. We assume that the observations are equal to the model $\eta(\cdot)$ run at its 'best' input $\hat{\theta}$, plus an error term $\delta + \varepsilon$. We split the error term into two parts: (i) measurement error $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)$ and natural interannual variability δ

The measurement error is assumed to be heteroskedastic zero-mean Gaussian noise with known variance (see Section 2.2). We model the natural variability of atmospheric CO₂ as a first-order autoregressive (AR(1)) process, with $\delta(t) = \rho\delta(t-1) + U$, where $U \sim N(0, \sigma_\delta^2)$. Our prior estimate of the correlation time $\tau = -1/\ln(\rho)$ is 15 ± 10 years, i.e. with a $N(15, 10^2)$ prior, to account for decadal-scale fluctuations in CO₂ concentration. The process variance σ_δ^2 has a Gamma(4, 0.6) prior distribution.

We use a Gaussian process emulator to describe our prior beliefs about the model $\eta(\cdot)$ (O'Hagan, 2006; Sacks et al., 1989; Santner et al., 2003). We update these beliefs in light of the ensemble of model runs $Y = \{(\theta_i; y_i = \eta(\theta_i))\}_{i=1, \dots, 48}$, so that our posterior belief about the model, $\eta(\theta)|Y$, interpolates the ensemble values.

The model output is multi-dimensional, and considered here to contain two parts. The first part is a time series of global atmospheric CO₂ concentration, and the second part contains two air-sea CO₂ flux predictions. We use independent emulators for both parts of the output. We apply principal component analysis to reduce the dimension of the output and build emulators from the input space to this reduced output space. The emulator assumes the UVic model is a map $\eta: \Theta \rightarrow Y$, where $\Theta = \{Q_{10}, K_{cmult}\}$ is the two-parameter input space and Y is the n -dimensional output space. (Here $n=57$ for the concentration emulator, comprising 17 ice core CO₂ measurements and 40 instrumental measurements (1960-1999); and $n=2$ for the flux emulator, with two decadal fluxes.) Our

emulator approach can then be described in three parts: (i) Dimension reduction, (ii) Emulation, and (iii) Reconstruction to the full space. These parts are detailed below.

Dimension reduction

We follow Higdon *et al.* (2008) and use principal component analysis (PCA) to project Y onto a lower dimensional space Y' of dimension m , where $m \ll n$. We emulate the m leading principal components (PCs). Relying on diagnostic plots, we found $m=4$ PCs (accounting for 98.2% of the variance) to be sufficient for the time series of CO₂ output. We rotate the flux data onto their principal components, but do not discard any PCs, so $m=n=2$. The PC rotation is performed to decrease the correlation between the two flux outputs, allowing us to use independent emulators for each output.

Emulation

Gaussian process emulators are used to model the output of UVic, i.e., the map from Θ to Y' . We build independent emulators to model each principal component score, with $\eta(\cdot) = (\eta_1^{pc}(\cdot), \dots, \eta_m^{pc}(\cdot))^T$. There are four concentration emulators and two flux emulators. The form of each emulator is a regression component plus a residual term

$$\eta_i^{pc}(\theta) = \beta_i h(\theta) + u_i(\theta)$$

where $u_i(\theta) \sim GP(0, \sigma_i^2, c_i(\cdot, \cdot))$ and $i=1 \dots m$. Here $h(\theta)$ is a vector of regression functions common to each emulator. Each β_i is a vector of regression coefficients specific to emulator i . We follow Rougier (2008) and use Legendre function regressors shifted onto the range of the two input parameters, as these are orthonormal on parameter space Θ . We model up to third order terms with the sum of powers restricted to be four or less (described by the triplet (4,4,3) in the notation of An and Owen (2001)), giving a collection of 13 regressors. We give the regression parameters $\{\beta_i\}$ flat improper priors and analytically integrate them out of the posterior.

The residual $u(\cdot)$ between the model output and the regression surface is characterized by a Gaussian process with a covariance function $c(\cdot, \cdot)$. We take the covariance function to be squared-exponential, $c_i(\theta, \theta') = \exp[-(\theta - \theta')^T B_i (\theta - \theta')]$, and separable, so

that B_i is a diagonal matrix of roughness parameters (inverse squared correlation length scales). There are two roughness parameters corresponding to the two UVic model parameters, so B_i is 2×2 . We take an empirical Bayes approach for the elements of matrix B_i , estimating their values and holding them constant in the subsequent analysis. In order to estimate the roughnesses B_i , we use Adler's theorem to estimate the length scale of $u(\cdot)$ from the number of its upcrossings (Rasmussen and Williams, 2006; Sec. 4.1). We give the emulator variance parameters improper Jeffreys prior ($1/\sigma^2$) distributions. Separate roughness and variance parameters are used for each emulator.

Details of how to update these beliefs in light of the ensemble Y , along with discussion of some of these modeling choices, can be found in Santner *et al.* (2003). We took the approach advocated in Rougier *et al.* (2008) and used extensive diagnostic plots, such as leave-one-out and one-step-ahead plots, to refine and test our choices for $h(\cdot)$, B , and $c(\cdot, \cdot)$.

Reconstruction

The reconstruction from the emulator prediction in Y' (principal component space) to a prediction in the original (time series) space Y is a two-part process. The time series is a linear combination of the PC eigenvectors with coefficients given by the emulated PC scores. This gives an overconfident prediction for the emulated time series, because some information is lost in the PCA projection when higher PCs are discarded, and we are thereby uncertain about what the reconstruction would be if all PCs had been included. To account for this, we add small random multiples of the discarded (non-emulated) PCs to the time series. We model the non-emulated PC scores as zero-mean Gaussian noise with variance equal to the sample variance of the scores for that component in the ensemble. Strictly speaking this second step is not necessary since the discarded PCs only account for 2% of the variance in model output. It was developed as a prototype for circumstances where the PC reconstruction may be less accurate.

Calibration

The framework of Kennedy and O'Hagan (2001) gives the likelihood function for the data, $\pi(Z|\hat{\theta}, \sigma_{\eta}^2, \rho, \sigma_{\delta}^2)$, given the regression, covariance, and prior parameter assumptions described above. We simulate the posterior distribution for the unknown model and statistical parameters using Markov chain Monte Carlo (MCMC). Ten parameters are estimated via MCMC (2 UVic model parameters, 2 AR(1) parameters, 4 PC emulator variances for the CO₂ time series, and 2 decadal flux emulator variances). In addition, for each of the six emulators, we analytically integrate out 13 regression coefficients and estimate two roughness parameters as described in the Emulation section. We use a Metropolis-within-Gibbs algorithm, with Metropolis updates for $\hat{\theta}$, ρ , σ_{δ}^2 , and a conditional Gibbs update for the emulator variances $\pi(\sigma_{\eta}^2|\hat{\theta}, \rho, \sigma_{\delta}^2, Z)$.