Global Carbon Project

OptIC (OPTimisation InterComparison): An Intercomparison of Optimisation Techniques for Parameter Estimation in Terrestrial Biogeochemical Models

Overview and Protocol

Cathy Trudinger, Mike Raupach

Version 08 (10 Nov 05 CT)

1 Overview

1.1 Objective

The aim of the OptIC project is to comparatively evaluate several parameter estimation and data assimilation methods for the task of determining parameters in biogeochemical models from multiple sources of noisy data. Particular foci are (1) ability to handle data inadequacies such as noise, correlations and gaps; (2) ways to optimise the model-data synthesis process in the absence of good error specifications for the data; and (3) ways to handle multiple data sources with quite different properties.

The intention of OptIC is to focus specifically on the intercomparison of model-data synthesis methods, rather than biogeochemical models or data sources. This is the reason for the use of a transparently simple test model which embodies only the essential features of a real biogeochemical model, and artificially generated data for which "true" parameters are known.

The anticipated outcomes of OptIC are

- Better understanding of the strengths and limitations of a variety of model-data synthesis methods, especially with regard to handling noisy and incomplete data;
- Iterative improvement of methods by systematically tackling a sequence of progressively more challenging problems in terms of data inadequacies;
- Hence, design of parameter estimation and data assimilation methods for realistic assimilation of multiple data streams into biogeochemical models, to observe and predict the terrestrial balances (fluxes, storage changes) of carbon, water and related entities.

1.2 Approach

The method is to estimate parameters in a highly-simplified test model, using artificially generated data (from a forward run of the model) to which several forms of data degradation are added. The test model is a highly simplified two-equation representation of the carbon dynamics in a terrestrial biosphere model, with two state variables corresponding roughly to biomass carbon and soil carbon. The artificially generated data are surrogates for remote sensing data on

biomass (such as NDVI-based estimates of green leaf area index) and in-situ store measurements (such as soil carbon). Possible forms of data degradation added to the artificially generated test data include noise with Gaussian or skewed distributions, correlation among different data streams, temporal correlations, drifts, extreme outliers, and gaps in one or more data streams.

The test model (but not the parameter values) and degraded data sets are made available to investigators and teams wishing to test particular methods. The investigators return their results for estimated parameters and other metrics (including parameter uncertainties and model forecast capability with the estimated parameters) to a central archive where objective evaluations of the success of the estimations are made.

To date, methods for inclusion in the intercomparison are: (1) down-gradient search by the Levenberg-Marquardt method; (2) down-gradient search using a model adjoint to find the gradient; (3) the Extended Kalman Filter (EKF); (4) the Ensemble Kalman Filter (EnKF); (5) the Metropolis-Hastings method; (6) a Genetic Algorithm.

2 Background

2.1 Context and Related Work

The roles of terrestrial and oceanic biogeochemical processes in the climate system and the earth system are now well recognised (Steffen *et al.* 2004; Field and Raupach 2004). There is a corresponding recognition of the need to include these processes in climate and earth system models, especially for long-term (multiannual and greater) simulations.

Data assimilation into models has been under intensive development in meteorological and ocean forecasting since the 1980s, and has led to major improvements in forecast ability. Increasingly, data assimilation and parameter estimation methods (collectively termed model-data synthesis) are being used to constrain models of biogeochemical cycles (both in stand-alone form and as components of climate system models) with multiple sources of data (Raupach *et al.* 2005). The application of these approaches into biogeochemical modelling has accelerated very quickly since about 2000, with the active support of several international initiatives. These include:

- The Global Carbon Project (GCP), a joint project of IGBP, IHDP, WCRP and Diversitas (the Earth System Science Partnership, ESS-P). The Global Carbon Project has identified model-data synthesis as a primary tool for resolving patterns and trends in the carbon cycle at global and regional scales. The OPTimisation InterComparison (OptIC) is a contribution to Activity 1.2 on "Model Development and Model-Data Fusion" of the GCP Implementation Strategy (Global Carbon Project 2003, available at http://www.globalcarbonproject.org/science_plan_and_implementation.htm)

- The Integrated Global Carbon Observation (IGCO) Theme (http://ioc.unesco.org/igospartners/carbon.htm) of the Integrated Global Observation Strategy Partnership (IGOS-P) (http://ioc.unesco.org/igospartners). The overall objective of the IGCO Theme is to develop a flexible and robust strategy for international global carbon observations over the next decade, combining remote and in-situ observations and bringing together observational strategies for land, oceans and atmosphere.

- The North American (NLDAS) and Global (GLDAS) Land Data Assimilation Systems (http://ldas.gsfc.nasa.gov), which are being developed mainly from a hydrological perspective to improve reanalysis and forecast simulations by numerical weather prediction (NWP) models, particularly for soil moisture stores and land-air energy fluxes.

2.2 Major Questions

A wide variety of methods is available for model-data synthesis in biogeochemical modelling, including many forms of down-gradient searching (sometimes using adjoint models for determining gradients); global search methods including genetic-algorithm, simulated-annealing and related approaches; and several forms of the Kalman filter. Raupach *et al.* (2005) survey these methods from the standpoint of terrestrial biogeochemical modelling, with particular emphasis on their requirements for uncertainty specification in the data. Despite this plethora of possible methods and activity, biogeochemical model-data synthesis remains a young science. Major questions include the following:

1. *What methods are appropriate for particular problems?* This must account for the complexity of the model (number of state variables); the characteristics of the data (number and uncertainty properties of the measurements; extent of prior knowledge on parameters); and whether data are available all at once (as in most parameter estimation problems) or sequentially (as in data assimilation problems for constraining forecast models).

2. *How do model-data synthesis methods respond to inadequacies in the data?* Inadequacies that characterise most real data include noise, correlations, gaps and outliers.

3. *How can particular methods be tuned for best performance in biogeochemical model-data synthesis?* This involves (typically) adjustment of information about model error and data error so that each exerts an appropriate influence on the final result. Tuning may be *a priori* or adaptive during the model-data synthesis process.

4. *How can model-data synthesis be carried out effectively using multiple data sources with very different spatial and temporal densities and error characteristics?* "Multiple constraints" refers to the simultaneous use of multiple kinds of observations (for example, atmospheric composition measurements, remote sensing, eddy-covariance fluxes, vegetation and soil stores, and hydrological data) in model-data synthesis. Issues requiring resolution include the need to handle data sources with quite different spatial and temporal scales of measurement, very different sample numbers and different error properties.

5. *How can parameter estimation be done in large spatial domains with high spatial heterogeneity and few on-ground measurement points?* This includes the issue of undersampling spatially variable parameters, and that of making spatially distributed parameter estimation computationally efficient over large domains.

Questions such as these suggest the need for methodological development in model-data synthesis for terrestrial biogeochemical applications. The Optimisation Intercomparison (OptIC) project is intended to contribute to this development.

3 Test Model

3.1 Formulation

The test model is a highly simplified representation of the carbon dynamics in a terrestrial biosphere model (TBM), with two state variables corresponding conceptually to stores of biomass carbon (x_1) and litter and soil carbon (x_2). These variables are governed by the equations

$$\frac{dx_1}{dt} = \underbrace{F(t)\left(\frac{x_1}{p_1+x_1}\right)\left(\frac{x_2}{p_2+x_2}\right)}_{\text{Net Primary Production (NPP)}} - \underbrace{k_1 x_1}_{\text{Litterfall}} + \underbrace{s_0}_{\substack{\text{Seed}\\ \text{production}}} \tag{1}$$

$$\frac{dx_2}{dt} = \underbrace{k_1 x_1}_{\text{Litterfall}} - \underbrace{k_2 x_2}_{\substack{\text{Heterotrophic}\\ \text{respiration}}} \tag{2}$$

where $F(t)$ is a forcing term describing input into the biomass store x_1 by net primary production (NPP), the flux of carbon into biomass by growth; p_1 and p_2 are scales for the limitation of production by lack of x_1 and x_2, respectively; k_1 and k_2 are rate constants for the decay of x_1 and x_2, respectively; and s_0 is a "seed production" term for x_1.

This test model is a severe but rational simplification of a full TBM, in the following sense. All biomass carbon (leaf, wood, root) is lumped into a single store x_1, and all litter and soil carbon into a single store x_2. These are respectively governed by equations of the form dx_1/dt = (NPP) − (litterfall) and dx_2/dt = (litterfall) − (heterotrophic respiration). The flux terms on each right hand side are identified in Equations (1) and (2) (except for seed production, which is discussed shortly). Litterfall is parameterised as a flux k_1x_1 which is an outflow from x_1 and an inflow to x_2, where k_1 is a rate constant. Likewise, heterotrophic respiration is parameterised as an outflow flux k_2x_2 from the x_2 pool. NPP depends on the availability of essential resources (light, water and nutrients) and also on the biomass invested in organs for resource acquisition (leaves for light, roots for water and nutrients). The light and water resources are together represented by a forcing term $F(t)$ equal to the NPP under given light and water inputs, without limitation by either lack of biomass investment in resource-gathering organs or nutrient availability. The time dependence in $F(t)$ accounts for fluctuating availability of light and water through variation in weather and climate. The actual NPP is less than $F(t)$ because of lack of biomass investment in resource-gathering organs and lack of nutrients, described respectively by the factors $x_1/(p_1 + x_1)$ and $x_2/(p_2 + x_2)$ (of Michaelis-Menten form). To account for nutrient limitation, x_2 (litter and soil carbon) is used rather than a soil nutrient store, since these two stores tend to vary together.

The parameter s_0 is a small "seed production" term for x_1. It represents growth of biomass from seed, assumed (unrealistically) to be a constant growth flux independent of x_1, x_2 and external conditions. It is included for the following reason: if s_0 is absent ($s_0 = 0$), then "extinction" of the test model biosphere is possible because $(x_1,x_2) = (0,0)$ is a stable equilibrium solution of Equations (1) and (2). As explained in the next section on the dynamical properties of the test model, this does not occur if $s_0 > 0$ (and s_0 is not too small).

This test model is not intended to be an actual TBM but rather a simplified version of a TBM with enough of the mathematical properties of a real TBM to answer the questions posed in the OptIC objectives. These properties include (a) nonlinearity, through the form of the NPP term; (b) multiple stores (here two); (c) data streams giving information about the stores or fluxes (such as NPP) but not directly about the parameters, which must be estimated.

Throughout, the forcing term $F(t)$ is assumed to be externally specified as a "log-Markovian" random process, that is, a process such that ln $F(t)$ is Markovian with specified mean and standard deviation. Specifically, we take $F(t) = p_0\exp(m(t))$, where p_0 is a measure of the mean magnitude of $F(t)$, and $m(t)$ is a dimensionless Markov process with zero mean, standard deviation σ_m, and time scale T_m. The process $m(t)$ obeys the Langevin equation, the stochastic differential equation $dm/dt = -(m/T_m) + \left(\sigma_m\sqrt{2/T_m}\right)\xi(t)$, where $\xi(t)$ is Gaussian white noise (Arnold 1974, Legg and Raupach 1982). In finite-difference form, $F(t)$ is given by:

$$F(t_i) = F_i = p_0 \exp(m_i)$$
$$m_i = a m_{i-1} + b\sigma_m w_i \qquad \left(\text{with } a = \exp(-\Delta t/T_m),\ b = \sqrt{1-a^2}\right) \tag{3}$$

where F_i and m_i are sequences of $F(t)$ and $m(t)$ at times t_i with increments Δt, and w_i is a Gaussian random number with zero mean and unit variance. This formulation ensures that $F(t)$ is always positive. The quantities determining $F(t)$ are p_0, σ_m and T_m. In most cases the values used are $p_0 = 1$, $\sigma_m = 0.5$ and $T_m = 10\Delta t$, with discretisation interval $\Delta t = 1$ time unit. These properties of $F(t)$ are given quantities, not model parameters to be estimated.

The test model has five parameters: p_1, p_2, k_1, k_2 and s_0. Figure 1 shows the behaviour of the model with "reference" parameter choices $p_1 = 1$, $p_2 = 1$, $k_1 = 0.2$, $k_2 = 0.1$, $s_0 = 0.01$, and a log-Markovian forcing function $F(t)$ specified as above.

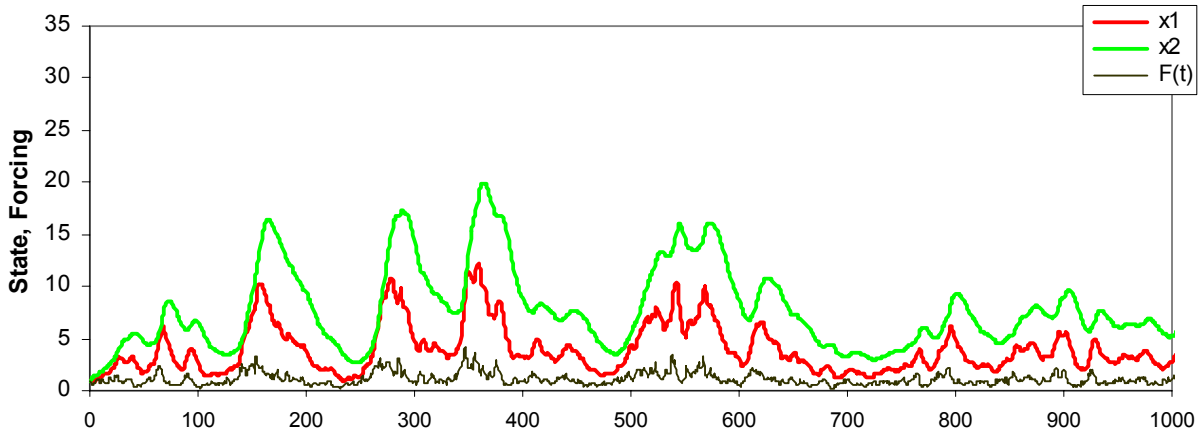

Figure 1: Test model behaviour with reference parameter choices $p_1 = 1$, $p_2 = 1$, $k_1 = 0.2$, $k_2 = 0.1$, $s_0 = 0.01$. Forcing function $F(t)$ is log-Markovian with $p_0 = 1$, $\sigma_m = 0.5$, $T_m = 10\Delta t$, and discretisation interval $\Delta t = 1$ (see text).

3.2 Dynamical Properties

Despite the test model being quite simple in form, it has some subtle properties as a dynamical system. Without providing details here, a brief summary of these properties is as follows.

Steady forcing F(t): Let us consider first the situation the forcing is steady and $F(t) = p_0$. In this case the model has equilibrium points or fixed points $\mathbf{x}_q = (x_{q1},x_{q2})$, at which $dx_1/dt = dx_2/dt = 0$. These points are solutions of

$$\begin{aligned} 0 &= \frac{dx_1}{dt} = p_0\left(\frac{x_1}{p_1+x_1}\right)\left(\frac{x_2}{p_2+x_2}\right) - k_1x_1 + s_0 \\ 0 &= \frac{dx_2}{dt} = k_1x_1 - k_2x_2 \end{aligned} \tag{4}$$

The second of Equation (4) implies that $x_2 = (k_1/k_2)x_1$ at equilibrium, and the first then implies that x_1 satisfies

$$\begin{aligned} g(x_1) &= c_0 + c_1x_1 + c_2x_1^2 + c_3x_1^3 = 0 \\ \text{with } \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} &= \begin{pmatrix} p_1p_2k_2s_0/k_1^2 \\ \left((p_1k_1+p_2k_2)s_0 - p_1p_2k_1k_2\right)/k_1^2 \\ \left(p_0 - p_1k_1 - p_2k_2 + s_0\right)/k_1 \\ -1 \end{pmatrix} \end{aligned} \tag{5}$$

Thus the equilibrium points are of the form $\mathbf{x}_q = x_1(1, k_1/k_2)$, where x_1 is a solution of the cubic equation $g(x_1) = 0$. This equation has either one or three real roots, yielding either one or three equilibrium points. The outcome is different depending on the value of the seed production s_0.

- *Seed production present:* In this case, $s_0 > 0$. With the reference parameter choices [$p_1 = 1$, $p_2 = 1$, $k_1 = 0.2$, $k_2 = 0.1$, $s_0 = 0.01$], there is just one real root of $g(x_1) = 0$, and hence just one equilibrium point. This is $(x_{q1},x_{q2}) = (3.428, 6.856)$. It is a stable (attracting) equilibrium point: solutions $((x_1(t), x_2(t))$ converge to this point as $t \to \infty$, from any starting point.

- *Seed production absent:* In this case, $s_0 = 0$. With all parameters set to their reference values except that $s_0 = 0$, there are three real roots of $g(x_1) = 0$, hence three equilibrium points, at (3.351, 6.702), at the origin (0,0), and at (0.149, 0.298). The first two of these are stable and the third is unstable. The reason for including a seed production term is to remove the (undesirable) stable point at the origin. With this equilibrium point present it is possible for solutions of the test model to be attracted to it if they stray into its basin of attraction close to the origin, and then be trapped there forever. Such an outcome would correspond to extinction of the terrestrial biosphere.

Unsteady random forcing F(t): Now we turn to the case where the forcing $F(t)$ is random, as in Figure 1. The model must now be solved numerically. Figure 2 shows three model integrations over 10000 time steps, all with seed production present ($s_0 = 0.01$). The first is with reference parameters as in Figure 1 (this figure just extends the series from 1000 to 10000 steps). In the other two, k_1 is increased from 0.2 (the reference value) to 0.4 and then 0.5 (with all other

parameters left at reference values). At the larger k_1 values, the model exhibits a dual-mode behaviour, flipping between "active" and "dormant" states in response to vagaries in the forcing $F(t)$. It is not visually apparent what aspects of $F(t)$ cause the flip. This aspect of the model behaviour is reminiscent of the blooming of desert ecosystems in response to rain, interspersed with long periods of dormancy.

Figure 2: Illustration of dual-mode (active, dormant) behaviour of the test model. The three panels show numerical model integrations with k_1 = 0.2, 0.4 and 0.5, and other parameters set to reference values. The forcing $F(t)$ is identical in each case.

4 OptIC Process

4.1 General

The optimisation intercomparison involves the estimation of optimum or minimum-error values for four parameters (p_1, p_2, k_1, k_2) in the test model, using artificially generated data from a forward run of the model in which these parameters have been assigned "true" values unknown to participants. The time series (($x_1(t)$, $x_2(t)$) from this forward run are treated as observations for the parameter estimation, but like real data, they are subjected to degradation (noise, correlations, drifts, gaps) before being supplied to participants. Thus the "observations" are $z_1(t) = x_1(t)$ + noise, and $z_2(t) = x_2(t)$ + noise. The observation series may contain missing values, possibly representing a large fraction of the data.

Participants in the intercomparison know the form of the test model and receive datasets of the forcing, $F(t)$, and observations $z_1(t)$ and $z_2(t)$. The "aim of the game" is to use a parameter estimation method of choice to estimate (1) the "true" parameter values (p_1, p_2, k_1, k_2) used to generate the observations, and (2) characteristics of the degradation such as noise statistics.

The results will be compared using a number of different metrics, to assess the ability of different methods to recover the model parameters under various conditions. Uncertainty analysis will be an important component of the exercise. A workshop will be held in early 2006 to compare and discuss the results. We expect to produce a comparison paper from the project, and participants may also produce associated papers on specific methods. We also expect to produce a CD with the model code, supplied datasets, code for some of the optimisation calculations, results and analysis.

4.2 Who can participate?

The optimisation intercomparison is open to anyone who would like to participate. Participants can register their intention to submit results and their chosen method by email to Cathy Trudinger (cathy.trudinger@csiro.au).

5 Details of Experiments

5.1 General

There will be 16 experiments in total. The first 6 experiments will be ‘training’ experiments, each involving the same set of parameter values, initial values and forcing time series, but with different types of noise. These are denoted T1 through T6. There will then be 10 experiments (denoted A through J), each involving a different set of parameter values, a different forcing time series $F(t)$, and with different types of noise imposed on the data. It will be up to participants to choose values for any additional information required by their optimisation method, such as uncertainties in the measurements, initial parameter uncertainties or initial values for x_1 and x_2. Participants will submit their best estimates of the parameters, and, if possible, the parameter

covariance matrix, for each experiment (A through J as well as T1 through T6). Although participants know that experiments T1-T6 have the same parameter values, they should treat them as individual experiments, as this will give us the best indication of how the different types of error influence the solution. The forcing time series will be specified for 12000 timesteps, but the noisy observations available only for the first 10000 timesteps. We would like participants to calculate $x_1(t)$ and $x_2(t)$ for the full 12000 timesteps, using their best estimates of the parameters and the given forcing $F(t)$ for each experiment. This will test how differences in parameter estimates affect predictions of x_1 and x_2 when there is no data for assimilation.

The programs (in Fortran 90) used to generate the data and noise will be available to participants for testing of their optimisation calculation.

5.2 Parameter ranges

Participants will estimate the values of four model parameters: p_1, p_2, k_1, k_2. The value of s_0 will be set throughout to 0.01. The "true" parameter values will lie within known "prior" ranges shown in Table 1. Participants may choose their estimates and uncertainty ranges (if required) based on these ranges.

Parameter	Minimum Value	Maximum Value
p_1	0.5	5
p_2	0.5	5
k_1	0.03	0.9
k_2	0.01	0.12
s_0	fixed at 0.01	

Table 1: Prior ranges for parameters

5.3 Types of observation noise

Below is a list of the types of noise that may be added to the data. The noise properties of z_1 and z_2 may be different. In the following, y_i is the uncorrupted observation (in our case $y_i = x_i$), z_i is the observation with noise added, and w_i is a Gaussian random number with zero mean and unit variance. In most cases the noise added to the observations has zero mean.

Not all of the following cases will be represented in the experiments.

1. Gaussian random noise with constant standard deviation σ:

$$z_i = y_i + \sigma w_i$$

2. Gaussian random noise with standard deviation σy_i (proportional to signal y_i):

$$z_i = y_i \left(1 + \sigma w_i\right)$$

3. Uniformly distributed noise added to signal:

$$z_i = y_i + \sigma u_i$$

where $u_i \in U[-1, 1]$, meaning that u_i is a random number drawn from a uniform distribution between −1 and 1.

4. Random noise with log-normal (skewed) distribution added to signal:

$$z_i = y_i + \exp(\sigma w_i) - \exp(0.5\sigma^2)$$

The last term is the population mean of the noise term $\exp(\sigma w_i)$, which is subtracted to give an unbiased (zero-mean) overall noise term.

5. Random noise with log-normal distribution multiplied by signal:

$$z_i = y_i \exp(\sigma w_i) - y_i\left[\exp(0.5\sigma^2) - 1\right]$$

The noise population mean, the last term, is subtracted to give an unbiased overall noise.

6. Random noise with Weibull distribution added to signal

$$z_i = y_i + b(-\ln v_i)^{1/c} - b\Gamma\left(1 + \frac{1}{c}\right)$$

where $v_i \in U[0, 1]$. The noise population mean, the last term, is subtracted to give an unbiased overall noise. The Weibull distribution has the probability density function $f(x) = cb^{-c}x^{c-1}\exp[-(x/b)^c]$, where b is the scale parameter and c the shape parameter.

7. Gaussian random noise with a time-invariant correlation between noise in z_1 and z_2 at each time instant:

$$\begin{pmatrix} z_{1,i} \\ z_{2,i} \end{pmatrix} = \begin{pmatrix} y_{1,i} \\ y_{2,i} \end{pmatrix} + \begin{pmatrix} w_{1,i} \\ w_{2,i} \end{pmatrix}\begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

where r is the correlation coefficient between noise in z_1 and z_2, and $w_{1,i}$ and $w_{2,i}$ are uncorrelated Gaussian random noise sequences with zero mean and unit variance.

8. Noise correlated in time (Markov sequence):

$$z_i = y_i + m_i$$
$$m_i = am_{i-1} + b\sigma_m w_i \qquad \left(\text{with } a = \exp(-\Delta t/T_m),\ b = \sqrt{1-a^2}\right)$$

where σ_m is the standard deviation of the Markov sequence and T_m is its integral time scale. This is the time scale over which the noise is temporally correlated. The above is similar to Equation (3).

9. Gaussian random noise plus the occasional Gaussian-distributed extreme outlier:

$$z_i = y_i + \sigma_n w_{n,i} + H(p_x - v_i)\sigma_x w_{x,i}$$

where σ_n is the standard deviation of the Gaussian random noise, σ_x ($>> \sigma_n$) is the standard deviation of the population of extreme outliers, $w_{n,i}$ and $w_{x,i}$ are uncorrelated Gaussian random noise sequences with zero mean and unit variance, v_i is a uniform random number between 0 and 1 ($v_i \in U[0, 1]$), $H(x)$ is the Heaviside step function ($H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x > 0$), and p_x is the (small) probability of extreme outliers. The last term is zero unless the random number v_i falls below p_x.

10. Gaussian random noise plus shifts of random magnitudes lasting different intervals of time (mimicking calibration tank changes):

$$z_i = y_i + \sigma_n w_{n,i} + \zeta_i$$
$$\text{with } \zeta_i = \zeta_{i-1} \quad \text{if } H(p_s - v_i) = 0$$
$$\text{and } \zeta_i = \sigma_s w_{s,i} \text{ if } H(p_s - v_i) = 1$$

where ζ_i represents the random shift part of the noise, p_s is the (small) probability of a shift, and σ_x is the standard deviation of the zero-mean Gaussian distribution from which shifts are drawn.

11. Gaussian random noise plus drifts of different magnitudes, resetting to zero at random intervals (mimicking calibration tank drifts):

$$z_i = y_i + \sigma_n w_{n,i} + \zeta_i$$
$$\text{with } \zeta_i = \zeta_{i-1} + k \quad \text{if } H(p_d - v_i) = 0$$
$$\text{and } \zeta_i = 0,\ k = \sigma_d w_{d,i} \text{ if } H(p_d - v_i) = 1$$

where ζ_i is the drift part of the noise, p_d is the (small) probability of a resetting of the drift to zero, and σ_d is the standard deviation of the zero-mean Gaussian distribution from which drift rate is drawn. The drift rate k is held constant except for occasional resets.

5.4 Data sets

There will be 16 cases in total, with different types and magnitudes of observation noise, different forcing time series, different initial values for x_1 and x_2, and different parameters (p_1, p_2, k_1, k_2) to be estimated. For each case, the following will be supplied:

- Forcing time series $F(t)$ for 12000 timesteps
- Observations z_1 and z_2 (which include noise) for 10000 timesteps. Some cases may have some missing data for z_2, flagged as −999.

Initial values of x_1 and x_2 will not be given.

Some observations could have negative values, depending on the type and magnitude of the noise, and participants can choose how to deal with this.

5.5 Outputs required from parameter estimation

For each of the 16 cases, participants are requested to provide (where possible):

- Estimated parameters (p_1, p_2, k_1, k_2)
- Covariance matrix for parameters
- x_1 and x_2 calculated with the estimated parameters for the full 12000 timesteps

We recognise that some methods may not give estimates of the covariance matrix.

Please also provide up to one page (in ASCII text, Microsoft Word or LaTeX) about your method, how you applied it to parameter estimation, any assumptions that you made and whether you found any of the cases particularly difficult.

5.6 Submission of results

To Cathy Trudinger (cathy.trudinger@csiro.au), in ASCII format, one file per case, with:

- Parameter estimates (p_1, p_2, k_1, k_2) [formatted as in the Fortran statement write(*,'(4e16.6)')]
- Covariance matrix (4x4) for parameters p_1, p_2, k_1, k_2 [formatted as 4 lines of write(*,'(4e16.6)')]
- Time series of t, $x_1(t)$, $x_2(t)$ for the entire length of the forcing function record, $0 \le t \le 12000$ [formatted as 12001 lines of write(*,'(i6,2e16.6)')], with $t = 0$ referring to initial values.
- If any quantity is not available (such as covariance matrix), use the special value –999.
- The 10 files should be named like "SmithA.dat" using the name of the participant and the case letter (A to J or T1 to T6), and be emailed as attachments.

Results are due by **Friday 11 November 2005**.

References

Arnold L (1974) Stochastic Differential Equations: Theory and Applications. Wiley-Interscience, New York, 228 pp.

Field CB, Raupach MR, (Eds.) (2004) The Global Carbon Cycle: Integrating Humans, Climate, and the Natural World. Island Press, Washington, 526 pp.

Global Carbon Project (2003) Science Framework and Implementation. Earth System Science Partnership (IGBP, IHDP, WCRP, Diversitas) Report No. 1; GCP Report No. 1, Global Carbon Project, Canberra.

Legg BJ, Raupach MR (1982) Markov-chain simulation of particle dispersion in inhomogeneous flows: the mean drift velocity induced by a gradient in Eulerian velocity variance. *Boundary-Layer Meteorology* **24** 3-13.

Raupach MR, Rayner PJ, Barrett DJ, DeFries R, Heimann M, Ojima DS, Quegan S, Schmullius C (2005) Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Global Change Biology* .

Steffen WL, Sanderson A, Tyson PD, Jager J, Matson PA, Moore BI, Oldfield F, Richardson K, Schellnhuber HJ, Turner BLI, Wasson RJ (2004) Global Change and the Earth System: a Planet under Pressure. Springer, Berlin, 336 pp.