A STUDY INTO THE EFFECT OF DIFFERENT FORMULATIONS OF THE METROPOLIS-HASTINGS ALGORITHM IN ESTIMATING MODEL PARAMETERS AND ERRORS

Edmund Ryan*, Mat Williams, Shaun Quegan***

*Department of Applied Mathematics, University of Sheffield

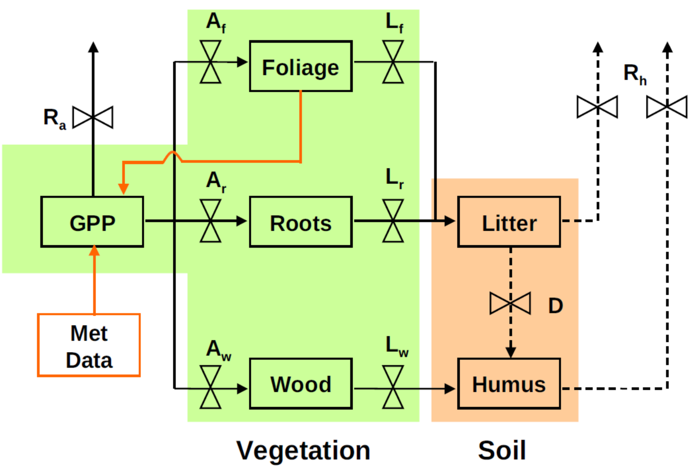
**Institute of Atmospheric and Environmental Sciences, University of Edinburgh

1. Background

REFLEX (REgional FLux Estimation eXperiment) is a Data Assimilation (DA) inter-comparison project which compared 8 different algorithms for estimating carbon (C) model parameters consistent with both measured carbon fluxes and states and a simple terrestrial C model. REFLEX was led by CTCD (M Williams, with many of the analyses undertaken by A Fox) and involved nine participating international groups.

Each participant used their own DA algorithm, but all used the same model, DALEC (Data Assimilation Linked Ecosystem Carbon model) and were provided with the same observational data (real and synthetic) for daily Net Ecosystem Exchange (NEE) of CO_2 and Leaf Area Index (LAI). Each participant was asked to estimate the parameters of DALEC optimised against the data, using their own DA algorithm. The synthetic data were generated by sampling a set of model outputs and adding noise and gaps to the data. The real data were obtained from FLUXNET sites in Europe.

The DALEC model
DALEC tracks the daily flows of carbon from photosynthesis (GPP) through to plant (Ra) and soil (Rh) respiration.
It has:
- 6 pools (boxes)
- 10 fluxes (arrows)
- 11 parameters.



2. Results of REFLEX

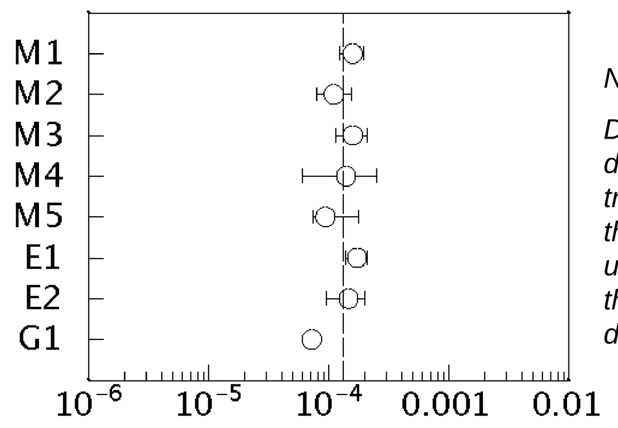
For about half of the parameters, the estimates from the 8 DA algorithms were unbiased compared to the synthetic truth, with narrow error bars, e.g. the soil respiration parameter (figure 1):

Figure 1

M1→ M5 are Metropolis algorithms

E1 & E2 use the Ensemble Kalman filter

G1 is a Genetic algorithm



Note:
Dashed line denotes the true value of the parameter used in creating the synthetic data.

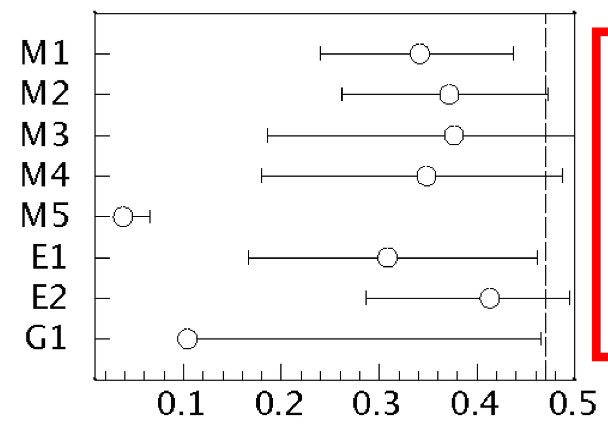
For the other parameters, most of the algorithms gave biased estimates, with a range of error intervals, e.g. the foliage allocation parameter (ie this parameter gives the fraction of NPP allocated to foliage, figure 2):

Figure 2

M1→ M5 are Metropolis algorithms

E1 & E2 use the Ensemble Kalman filter

G1 is a Genetic algorithm



Note that the true value is not even contained (or only just) within the error bars.

3. Aim of this work

Figures 1 and 2 arise from the use of synthetic data, generated from DALEC, so model error is not the issue. This raises two key questions:

(1) Why can't the DA methods get the right answer for the biased parameter (figure 2)?

(2) Why do the error estimates vary so much?

In my current work, I am trying to answer these questions.

4. Methodology

Since 6 of the DA algorithms were variations of the Metropolis-Hastings (M-H) algorithm, one possible reason for the differences in the parameter and error estimates may be differences in the formulation of the M-H algorithm.

Therefore, one of the M-H algorithms was chosen and sensitivity analysis conducted. Three factors (in the formulation) could be important:

(i) Step-size;

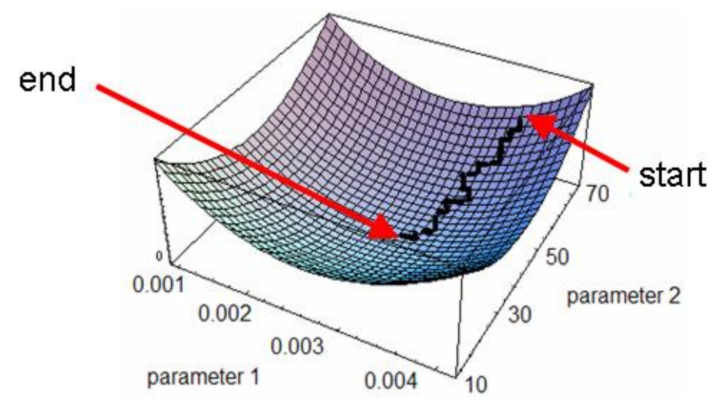
(ii) Cost function type (based on NEE & LAI or NEE alone);

(iii) The initial conditions of the small Carbon pools of foliage (Cf), roots (Cr) and litter (Clit).

In my study up until now, I have only conducted sensitivity analysis on (i) and (ii).

5. The M-H algorithm

This example shows how the M-H algorithm works with 2 parameters.



- The <u>step-size</u> tells us how big a step the algorithm should take when moving around the parameter space to find the optimal parameter set.
- The <u>cost function</u> is the sum of squares difference between the observations and the model estimates for a given set of parameters. The observations provided were NEE and LAI over a 3-year period. The algorithm works by stepping around the parameter space and at each step a parameter set is accepted if the sum of squares difference corresponding to the set of parameters where the algorithm has stepped to, is less than the sum of squares corresponding to the set of parameters where the algorithm has stepped from.

6. Design of the Experiment

After the M-H algorithm had converged we examined the parameter values for the 30,000 steps that followed. Using these converged values, we calculated the **posterior mean*** and the length of the **90% posterior error interval****, for each of the following two parameters:

(a) Ts, the parameter which controls soil respiration. This parameter showed small biases and small error bars in REFLEX.

(b) Fnf, the parameter that determines the fraction of NPP (NPP is GPP minus autotrophic respiration) which is allocated to foliage. This parameter gave significant biases.

In the runs of the M-H algorithm we examined the behaviour of the 2 parameters after:

(i) The step-size was varied over 10 values between 0.005 and 0.015 using a stratified random sample.

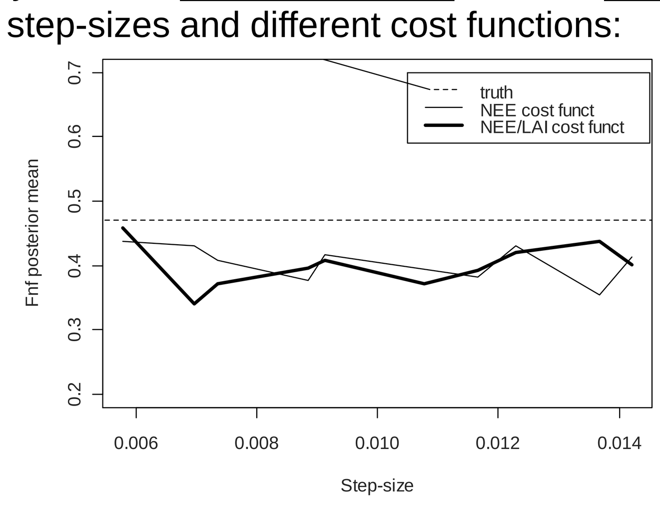
(ii) The cost function took two forms: using NEE observations only, or using equally weighted NEE and LAI observations.

** The word 'posterior' is used here to denote the fact that we are calculating the mean 'after' convergence.*

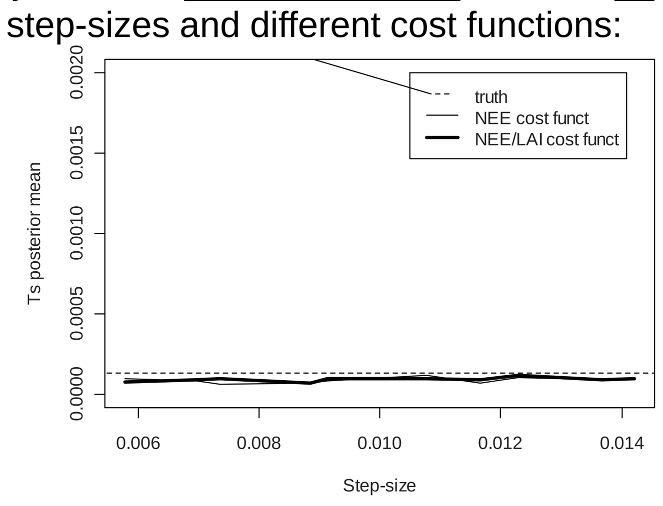
***A 90% posterior error interval is defined here as being the interval that has as its bounds the 5th and 95th percentiles.*

7. Results

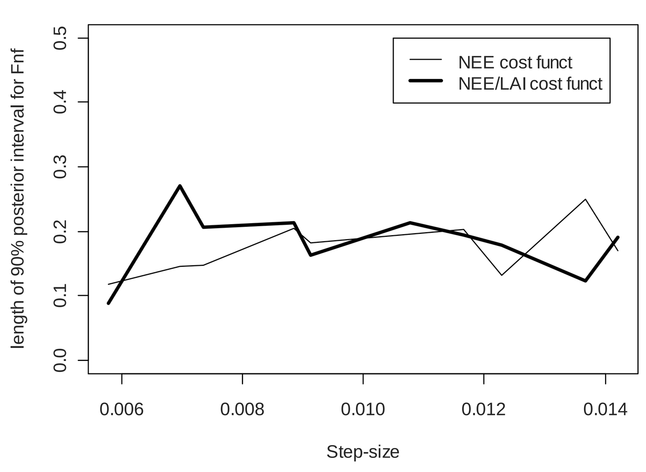
Sensitivity of the **<u>posterior mean</u>** of the **<u>Fnf</u>** parameter for different step-sizes and different cost functions:



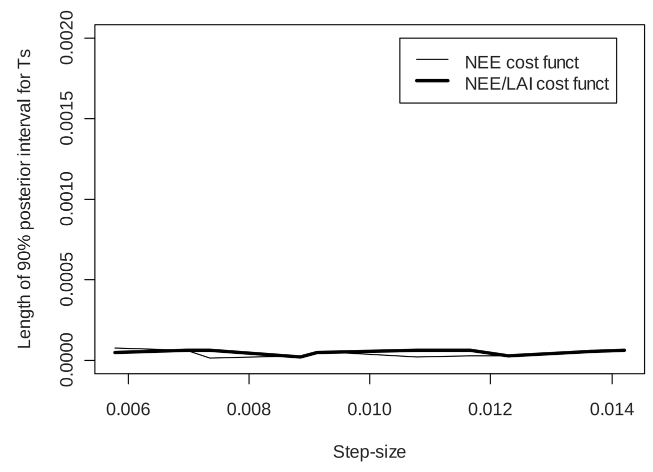
Sensitivity of the **<u>posterior mean</u>** of the **<u>Ts</u>** parameter for different step-sizes and different cost functions:



Sensitivity of the length of the **<u>90% posterior error interval</u>** of the **<u>Fnf</u>** parameter for different step-sizes and different cost functions:



Sensitivity of the length of the **<u>90% posterior error interval</u>** of the **<u>Ts</u>** parameter for different step-sizes and different cost functions:



8. Discussion of Results

It is clear from the 4 diagrams that the parameter estimates and length of the 90% error intervals are not sensitive to changes in the step-size (between 0.005 and 0.015). It also makes little difference as to whether the cost function should be composed solely of NEE, or NEE and LAI equally weighted; a possible reason for this is because there were very few LAI observations (less than 10), compared to around 250 NEE observations.

9. Conclusion

- The purpose of this work was to determine why the DA algorithms in REFLEX got it so wrong for the biased parameters and why the error bars varied in length.
- To find this out, we ran a Metropolis algorithm with different step-sizes and different cost-function types (NEE only and NEE & LAI equally weighted) for 2 parameters, and observed whether the parameter estimates and the length of the error bars changed.
- The results showed that the posterior mean (the algorithm's estimate of the parameter) and the length of 90% posterior error interval (the error bar) were not sensitive to changes in the step-size (between 0.005 and 0.015), nor to the type of cost function.
- This work is Stage 1 of the follow-on work to REFLEX. Stage 2 will investigate the effect of varying the initial conditions for the small C pools of Cf, Cr and Clit (Carbon stored in foliage, roots and litter respectively).

<u>References</u>

Fox, A, M Williams, AD Richardson, D Cameron, JH Gove, T Quaife, D Ricciuto, M Reichstein, E Tomelleri, C Trudinger, MT van Wijk (2009) <u>The REFLEX project: comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data</u>, ***Agricultural and Forest Meteorology*** *149, 1597–1615.*