Using data assimilation to understand the effect of disturbance on a managed woodland

Ewan Pinnington

October 16, 2016

Abstract

The response of forests and terrestrial ecosystems to disturbance is an important process in the global carbon cycle in context of a changing climate. Disturbance can take many forms, for example; felling, fire and insect outbreak. In current estimates of the global carbon budget disturbance is one of the least understood components. In this paper we investigate the effect of management practices on the carbon dynamics of a mature temperate woodland. In order to better understand ecosystem response we use the mathematical technique of data assimilation to combine a diverse set of observations with a mathematical model of ecosystem carbon balance. This allows us to combine the uncertainty from observations and prior model predictions to find the best possible estimate to the studied system. We develop new data assimilation techniques allowing for the assimilation of otherwise neglected information. The data assimilation techniques presented in this paper are applicable to other ecosystem models and data assimilation schemes. Previous statistical analyses of eddy covariance data at the study site had suggested that disturbance from thinning resulted in no change to net ecosystem carbon uptake. We show that this is likely due to reduced heterotrophic respiration post-disturbance. Our results support the theory of an upper limit for forest net carbon uptake due to the magnitude of ecosystem respiration scaling with gross primary productivity.

1 Introduction

1.1 Role of disturbance in the global C cycle

1.2 Current theories on terrestrial ecosystem response to disturbance

Particular focus on effect of management practices on woodland carbon dynamics.

It has been shown that tree roots provide a rhizosphere priming effect, greatly increasing the rate of soil organic carbon decomposition [Dijkstra and Cheng, 2007].

Moore et al. [2013] found little change in net CO_2 flux following disturbance from mountain pine beetle outbreaks in North America due to concurrent reductions in gross primary productivity and ecosystem respiration.

1.3 The role of data assimilation for improving estimates to a system

Combining the errors from model and observations to improve understanding of a systems response to change.

1.4 What does this paper do?

Analyse the effect of disturbance from management practices (thinning) for a deciduous managed woodland (Alice Holt) by the combination of a diverse set of observations with a mathematical model of forest carbon balance. Western side of the site thinned in 2014 and the Eastern side left unmanaged. The site has a flux tower positioned on the boundary between managed and unmanaged forest. Eddy covariance observation record was split between two sides using a flux footprint model found in REF. Previous statistical analysis of a management event in 2007 had suggested that there was no change in the Net Ecosystem Exchange (NEE) of CO_2 between the managed and unmanaged sides of the forest after thinning. From the years data after the 2014 management event and optimised model in this paper we find evidence to support this and seek an explanation for why the net uptake of carbon remains unchanged even after removing a large proportion of the trees from one side. The data assimilation techniques presented in this paper could be applied for similar analyses at other sites and provide a novel method to help elucidate the reasons behind ecosystem responses.

2 Observation and data assimilation methods

2.1 Alice Holt research forest

Alice Holt Forest is a research forest area managed by the UK Forestry Commission located in Hampshire, SE England. Forest Research has been operating a CO_2 flux measurement tower in a portion of the forest, the Straits Inclosure, since 1998 so it is one of the longer forest site CO_2 flux records, globally. The Straits Inclosure is a 90ha area of managed deciduous broadleaved plantation woodland, presently approximately 80 years old, on a surface water gley soil. The majority of the canopy trees are oak (*Quercus robur* L.), with an understory of hazel (*Corylus avellana* L.) and hawthorn (*Crataegus monogyna* Jacq.); but there is a small area of conifers (*Pinus nigra* J. F. Arnold) within the tower measurement footprint area in some weather conditions. Further details of the Straits Inclosure site and the measurement procedures are given in Wilkinson et al. [2012], together with analysis of stand-scale 30 minute average net CO_2 fluxes (NEE) measured by standard eddy covariance methods from 1998-2011.

As part of the management regime, the Straits Inclosure is subject to thinning; whereby a proportion of trees are removed from the canopy in order to reduce competition and improve the quality of the final tree crop. At the Straits an intermediate thinning method is used with a portion of both subdominant and dominant trees being removed from the stand [Kerr and Haufe, 2011]. The whole of the stand was thinned in 1995. Subsequently the Eastern side of the Straits was thinned in 2007 and then the Western side in 2014. The flux tower at the site is situated on the boundary between these two sides, allowing for the use of a footprint model to split the flux record and analyse the effect of this disturbance on carbon fluxes at the site. In Wilkinson et al. [2015] a statistically analysis of the eddy covariance flux record found that there was no significant effect on the net carbon uptake of the Eastern side after thinning in 2007. In this paper we focus on the effect of disturbance on the Western side after thinning in 2014.

2.2 Model and data assimilation

2.2.1 DALEC2 ecosystem carbon model

The DALEC2 model is a simple process-based model describing the carbon balance of a forest ecosystem [Bloom and Williams, 2015] and is the new version of the original DALEC [Williams

et al., 2005]. The model is constructed of six carbon pools (labile (C_{lab}), foliage (C_f), fine roots (C_r), woody stems and coarse roots (C_w), fresh leaf and fine root litter (C_l) and soil organic matter and coarse woody debris (C_s)) linked via fluxes. The aggregated canopy model (ACM) [Williams et al., 1997] is used to calculate daily gross primary production (GPP) of the forest, taking meteorological driving data and the modelled leaf area index (a function of C_f) as arguments. Figure 1 shows a schematic of how the carbon pools are linked in DALEC2.

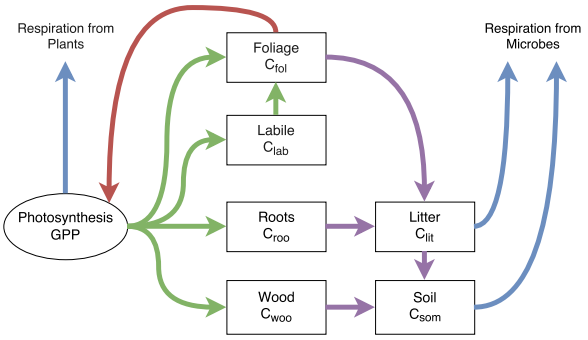

Figure 1: Representation of the fluxes in the DALEC2 carbon balance model. Green arrows represent C allocation, purple arrows represent litter fall and decomposition fluxes, blue arrows represent respiration fluxes and the red arrow represents the influence of leaf area index in the GPP function.

The model equations for the carbon pools at day i are as follows:

$$GPP^i = ACM(C_{fol}^{i-1}, c_{lma}, c_{eff}, \Psi) \tag{1}$$

$$C_{lab}^i = C_{lab}^{i-1} + (1 - f_{auto})(1 - f_{fol})f_{lab}GPP^i - \Phi_{on}C_{lab}^{i-1}, \tag{2}$$

$$C_{fol}^i = C_{fol}^{i-1} + \Phi_{on}C_{lab}^{i-1} + (1 - f_{auto})f_{fol}GPP^i - \Phi_{off}C_{fol}^{i-1}, \tag{3}$$

$$C_{roo}^i = C_{roo}^{i-1} + (1 - f_{auto})(1 - f_{fol})(1 - f_{lab})f_{roo}GPP^i - \theta_{roo}C_{roo}^{i-1}, \tag{4}$$

$$C_{woo}^i = C_{woo}^{i-1} + (1 - f_{auto})(1 - f_{fol})(1 - f_{lab})(1 - f_{roo})GPP^i - \theta_{woo}C_{woo}^{i-1}, \tag{5}$$

$$C_{lit}^i = C_{lit}^{i-1} + \theta_{roo}C_{roo}^{i-1} + \Phi_{off}C_{fol}^{i-1} - (\theta_{lit} + \theta_{min})e^{\Theta T^{i-1}}C_{lit}^{i-1}, \tag{6}$$

$$C_{som}^i = C_{som}^{i-1} + \theta_{woo}C_{woo}^{i-1} + \theta_{min}e^{\Theta T^{i-1}}C_{lit}^{i-1} - \theta_{som}e^{\Theta T^{i-1}}C_{som}^{i-1}, \tag{7}$$

where T^{i-1} is the daily mean temperature, Ψ represents the meteorological driving data used in the GPP function and Φ_{on}/Φ_{off} are functions controlling leaf-on and leaf-off. Descriptions for each model parameter used in equations (1) to (7) are included in the appendix in table 2. DALEC2 differs from the original DALEC in that it can be parameterised for both deciduous and evergreen sites with Φ_{on} and Φ_{off} being able to reproduce the phenology of either type of site. The full details of this version of DALEC can be found in Bloom and Williams [2015].

2.2.2 Data assimilation

We implement Four-Dimensional Variation data assimilation (4D-Var) with the DALEC2 model for joint parameter and state estimation. In 4D-Var we aim to find the parameter and initial state values such that the model trajectory best fits the data over some time window, given some prior information about the system. This prior information takes the form of an initial estimate to the

parameter and state variables of the model, $\mathbf{x}^b$, valid at the initial time. This prior is assumed to have unbiased, Gaussian errors with known covariance matrix $\mathbf{B}$. Adding the prior term ensures that our problem is well posed and that we can find a locally unique solution [Tremolet, 2006]. The prior used in this paper is derived from the assimilation of eddy covariance data from previous years. In 4D-Var we aim to find the parameter and initial state values that minimises the weighted least squares distance to the prior while minimising the weighted least squares distance of the model trajectory to the observations over the time window $t_0, \ldots, t_N$ [Lawless, 2013]. We do this by finding the state $\mathbf{x}_0^a$ at time t_0 that minimises the cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - \mathbf{h}_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{h}_i(\mathbf{x}_i)), \tag{8}$$

where $\mathbf{x}_0$ is the vector of parameter and initial state values to be optimised, $\mathbf{h}_i$ is the observation operator mapping the parameters and state to the observations $\mathbf{y}_i$ and $\mathbf{R}_i$ is the observation error covariance matrix. Further details of the implemented data assimilation scheme and specification of prior and observational errors can be found in Pinnington et al. [2016].

In this paper we assimilate day and nighttime NEE in order to increase the number of observations available to us and also better partition our modelled estimate of GPP and total ecosystem respiration, as discussed in section 2.3.1. As the DALEC2 model runs at a daily time step this requires us to relate the daily parameter and state values from the model to the twice-daily observations of NEE. We do this by writing two new observation operators, one relating the model state and parameters to daytime NEE and the other to nighttime NEE. The NEE of CO_2 at any given time is the difference between GPP and ecosystem respiration. For an observation of total daily NEE on day i we have,

$$NEE^i = -(1 - f_{auto})GPP^i(C^i_{fol}, \Psi) + \theta_{lit}C^i_{lit}e^{\Theta T^i} + \theta_{som}C^i_{som}e^{\Theta T^i}, \tag{9}$$

where all terms have the same meaning as describing in section 2.2.1. For total daytime NEE we have,

$$NEE^i_{day} = -GPP^i(C^i_{fol}, \Psi) + \delta_{day}f_{auto}GPP^i(C^i_{fol}, \Psi) + \delta_{day}\theta_{lit}C^i_{lit}e^{\Theta T^i_{day}} + \delta_{day}\theta_{som}C^i_{som}e^{\Theta T^i_{day}} \tag{10}$$

where δ_{day} is the day length, expressed as $\frac{\text{number of daylight hours}}{24}$, and T^i_{day} is the mean daytime temperature. For nighttime NEE we have,

$$NEE^i_{night} = \delta_{night}f_{auto}GPP^i(C^i_{fol}, \Psi) + \delta_{night}\theta_{lit}C^i_{lit}e^{\Theta T^i_{night}} + \delta_{night}\theta_{som}C^i_{som}e^{\Theta T^i_{night}} \tag{11}$$

where δ_{night} is the night length, expressed as $\frac{\text{number of night hours}}{24}$, and T^i_{night} is the mean nighttime temperature. Day length and night length are calculated using a solar model here, but could also be estimated using the record of incident solar radiation from the flux tower. These new observation operators allow for assimilation of day/nighttime NEE without the need for model development and can be applied to other ecosystem models to allow for the assimilation of finer temporal resolution eddy covariance data and possible improvements to the partitioning of photosynthesis and ecosystem respiration.

2.3 Observations

In order to assess the effect the 2014 thinning had on the Straits Inclosure an intensive field campaign was undertaken to measure leaf area index and also estimate woody biomass. From the site we also have a long record of flux data as discussed in section 2.1.

2.3.1 Flux tower eddy covariance

From the flux tower we have half-hourly observations from January 1999 to December 2015, these consist of the NEE fluxes and meteorological driving data used in the DALEC2 model of temperatures, irradiance and atmospheric CO_2 concentration. The NEE data is subject to is subjected to u^* filtering and quality control procedures as described by Papale et al. [2006], but is not gap-filled. This quality controlled half-hourly NEE dataset is then split between observations corresponding to the Western and Eastern sides of the site using a flux-footprint model, see Wilkinson et al. [2015] for more details.

Usually NEE observations are averaged daily to combine with daily time-step ecosystem carbon models. To compute daily NEE observations we take the mean over the 48 measurements made each day, only selecting days where there is no missing data. As we have been strict on the quality control of the flux record and not allowed any gap filling this presents a problem in the number of NEE observations available to us. By splitting the flux record between two sides, we retrieve very few total daily observations of NEE for each side. In order to address this we instead compute day and nighttime NEE fluxes (NEE_{day} and NEE_{night} respectively) for use in data assimilation. To compute daytime or nighttime NEE observations we take the mean over the half-hourly day or nighttime (calculated using a solar model) measurements, again only taking periods where there are no gaps in the data so that we are only considering true observations. This provides us with many more observations for assimilation after data processing, as seen in table 1, as we are averaging over shorter time periods so have a smaller probability of gaps and erroneous data. In section 2.2.2 we give details of how we relate these twice daily observations of NEE to a daily time-step model.

Sector	NEE	NEE_{day}	NEE_{night}
East	3	57	2
West	9	50	18

Table 1: Number of observations of NEE, NEE_{day} and NEE_{night} for East and West sides of the Straits Inclosure for the year 2015.

Raupach et al. [2005] Comment daytime NEE errors between 20-50%, nighttime measurements much more uncertain. We have been stricter on u^* filtering at night and also increase the size of observation error on nighttime NEE obs.

Richardson et al. [2008] the measurement error in observations of daily NEE is between 0.2 to 0.8 $\text{gCm}^{-2}\text{day}^{-1}$. Richardson et al. [2008] also shows that flux errors are heteroscedastic. We assume a constant standard deviation of 0.5 $\text{gCm}^{-2}\text{day}^{-1}$ in the assimilated observations of daily NEE as we found this standard deviation gave the best weighting to the observations in the assimilation algorithm, producing the best results for the forecast of NEE after assimilation. Assuming this constant standard deviation also allows for correlations in time between observation errors to be included more easily. Ignoring the heteroscedastic nature of NEE errors may influence results by giving observations of larger magnitude a higher weight than would be realistic. Future work should try to incorporate the heteroscedastic nature of NEE errors.

2.3.2 Leaf area index

LAI measurements made with ceptometer, hemispherical photographs and litter traps also giving us estimates to the leaf mass area of forest.

2.3.3 Woody biomass

Method of point centred quarters used from REF and an allometric relationship between diameter at breast heigh and total above ground biomass and corse root biomass to find an estimate to total woody biomass for both sides of the forest.

Mensuration measurements for West of Forest from Ian,

2.4 Experimental setup

Assimilation of 2015 years data post disturbance for the optimisation of two parameter sets, one corresponding to the unmanaged East and the other set to the managed West.

In this paper we have split the flux tower eddy covariance record using a footprint model. This means we have observations of NEE for both the West and East sides of the forest stand. By combining these two distinct sets of observations with our prior model using 4D-Var data assimilation allows us to retrieve two unique sets of parameter and initial state values, .

3 Results

Show plots of East and West after assimilation and the change in optimised parameters. Confident in results as we know that even assimilating a single year of data we can accurately describe the carbon dynamics of the site for a long time period (15 years) into the future from Pinnington et al 2016.

4 Discussion

From the assimilation of multiple data streams with a model of ecosystem carbon balance we find evidence that reduced heterotrophic respiration following disturbance allows a managed woodland to exhibit an unchanged net carbon uptake when compared against an undisturbed section of the same woodland.

5 Conclusion

Wrap up the results and discussion.

This supports work suggesting that there could be an upper limit to net carbon uptake of the land surface due to increased GPP leading to increased ecosystem respiration REF Andreas.

- We present novel techniques for the assimilation of day/night NEE using a modified observation operator. This facilitates the assimilation of a large amount of data that would have been previously neglected during data processing, without the need to alter the implemented model. Assimilation of NEE at this day/night time step could also help with the partitioning of NEE between GPP and ecosystem respiration.
- We demonstrate how data assimilation can be used to understand the effect of sudden changes to a studied system. Assimilating all available data streams after an event of disturbance allows us to assess changes to model parameter and state variables due to this disturbance.
- We find no significant change in net carbon uptake after disturbance for the managed side of the forest, despite a large proportion of trees being lost to felling. This was also found to be

true when a similar event occurred in 2007 REF Matt. It would be logical to assume that we would see decreased net carbon uptake due to this disturbance, resulting from fewer trees doing less GPP and heterotrophic respiration continuing at a similar rate as pre-disturbance. From our optimised model we find this unchanged net carbon uptake to be due to reduced heterotrophic respiration. So that even for a demonstrated decrease in GPP we have a relatively unchanged net carbon uptake. This supports work suggesting that there could be an upper limit to net carbon uptake of the land surface due to increased GPP leading to increased ecosystem respiration REF Andreas and others.

References

A. A. Bloom and M. Williams. Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological "common sense" in a modeldata fusion framework. *Biogeosciences*, 12(5):1299–1315, 2015. ISSN 1726-4189. doi: 10.5194/bg-12-1299-2015. URL `http://www.biogeosciences.net/12/1299/2015/`.

F. A. Dijkstra and W. Cheng. Interactions between soil and tree roots accelerate long-term soil carbon decomposition. *Ecology Letters*, 10(11):1046–1053, 2007. ISSN 1461-0248. doi: 10.1111/j.1461-0248.2007.01095.x. URL `http://dx.doi.org/10.1111/j.1461-0248.2007.01095.x`.

G. Kerr and J. Haufe. Thinning practice: A silvicultural guide. *Forestry Commission*, page 54, 2011.

A. S. Lawless. Variational data assimilation for very large environmental problems. In M. J. P. Cullen, M. A. Freitag, S. Kindermann, and R. Scheichl, editors, *Large scale Inverse Problems: Computational Methods and Applications in the Earth Sciences, Radon series on Computational and Applied Mathematics*, pages 55–90. De Gruyter, 2013.

D. J. P. Moore, N. A. Trahan, P. Wilkes, T. Quaife, B. B. Stephens, K. Elder, A. R. Desai, J. Negron, and R. K. Monson. Persistent reduced ecosystem respiration after insect disturbance in high elevation forests. *Ecology Letters*, 16(6):731–737, 2013. ISSN 1461-0248. doi: 10.1111/ele.12097. URL `http://dx.doi.org/10.1111/ele.12097`.

D. Papale, M. Reichstein, M. Aubinet, E. Canfora, C. Bernhofer, W. Kutsch, B. Longdoz, S. Rambal, R. Valentini, T. Vesala, et al. Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences*, 3(4):571–583, 2006.

E. M. Pinnington, E. Casella, S. L. Dance, A. S. Lawless, J. I. Morison, N. K. Nichols, M. Wilkinson, and T. L. Quaife. Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using four-dimensional variational data assimilation. *Agricultural and Forest Meteorology*, 228229:299 – 314, 2016. ISSN 0168-1923. doi: http://dx.doi.org/10.1016/j.agrformet.2016.07.006.

M. Raupach, P. Rayner, D. Barrett, R. DeFries, M. Heimann, D. Ojima, S. Quegan, and C. Schmullius. Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Global Change Biology*, 11(3):378–397, 2005.

A. D. Richardson, M. D. Mahecha, E. Falge, J. Kattge, A. M. Moffat, D. Papale, M. Reichstein, V. J. Stauch, B. H. Braswell, G. Churkina, B. Kruijt, and D. Y. Hollinger. Statistical properties of random {CO2} flux measurement uncertainty inferred from model residuals. *Agricultural and Forest Meteorology*, 148(1):38 – 50, 2008. ISSN 0168-1923. doi: http://dx.doi.org/10.1016/j.agrformet.2007.09.001.

Y. Tremolet. Accounting for an imperfect model in 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 132(621):2483–2504, 2006. ISSN 00359009. doi: 10.1256/qj.05.224. URL `http://doi.wiley.com/10.1256/qj.05.224`.

M. Wilkinson, E. Eaton, M. Broadmeadow, and J. Morison. Inter-annual variation of carbon uptake by a plantation oak woodland in south-eastern england. *Biogeosciences*, 9(12):5373–5389, 2012.

M. Wilkinson, P. Crow, E. Eaton, and J. Morison. Effects of management thinning on co_2 exchange by a plantation oak woodland in south-eastern england. *Biogeosciences Discussions*, 12(19), 2015.

M. Williams, E. B. Rastetter, D. N. Fernandes, M. L. Goulden, G. R. Shaver, and L. C. Johnson. Predicting gross primary productivity in terrestrial ecosystems. *Ecological Applications*, 7(3): 882–894, 1997.

M. Williams, P. A. Schwarz, B. E. Law, J. Irvine, and M. R. Kurpius. An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, 11(1):89–105, 2005.

Appendix

Parameter	Description	Background vector ($\mathbf{x}^b$)	Standard deviation	Range
θ_{min}	Litter mineralisation rate (day^{-1})	9.810×10^{-4}	2.030×10^{-3}	$10^{-5} - 10^{-2}$
f_{auto}	Autotrophic respiration fraction	5.190×10^{-1}	1.168×10^{-1}	$0.3 - 0.7$
f_{fol}	Fraction of GPP allocated to foliage	1.086×10^{-1}	1.116×10^{-1}	$0.01 - 0.5$
f_{roo}	Fraction of GPP allocated to fine roots	4.844×10^{-1}	2.989×10^{-1}	$0.01 - 0.5$
c_{lspan}	Determines annual leaf loss fraction	1.200×10^{0}	1.161×10^{-1}	$1.0001 - 10$
θ_{woo}	Woody carbon turnover rate (day^{-1})	1.013×10^{-4}	1.365×10^{-4}	2.5×10^{-5} -
θ_{roo}	Fine root carbon turnover rate (day^{-1})	3.225×10^{-3}	2.930×10^{-3}	$10^{-4} - 10^{-2}$
θ_{lit}	Litter carbon turnover rate (day^{-1})	3.442×10^{-3}	3.117×10^{-3}	$10^{-4} - 10^{-2}$
θ_{som}	Soil and organic carbon turnover rate (day^{-1})	1.113×10^{-4}	1.181×10^{-4}	$10^{-7} - 10^{-3}$
Θ	Temperature dependance exponent factor	4.147×10^{-2}	1.623×10^{-2}	$0.018 - 0.08$
c_{eff}	Canopy efficiency parameter	7.144×10^{1}	2.042×10^{1}	$10 - 100$
d_{onset}	Leaf onset day (day)	1.158×10^{2}	6.257×10^{0}	$1 - 365$
f_{lab}	Fraction of GPP allocated to labile carbon pool	3.204×10^{-1}	1.145×10^{-1}	$0.01 - 0.5$
c_{ronset}	Labile carbon release period (days)	4.134×10^{1}	1.405×10^{1}	$10 - 100$
d_{fall}	Leaf fall day (day)	2.205×10^{2}	3.724×10^{1}	$1 - 365$
c_{rfall}	Leaf-fall period (days)	1.168×10^{2}	2.259×10^{1}	$10 - 100$
c_{lma}	Leaf mass per area (gCm^{-2})	1.285×10^{2}	6.410×10^{1}	$10 - 400$
C_{lab}	Labile carbon pool (gCm^{-2})	1.365×10^{2}	6.626×10^{1}	$10 - 1000$
C_{fol}	Foliar carbon pool (gCm^{-2})	6.864×10^{1}	3.590×10^{1}	$10 - 1000$
C_{roo}	Fine root carbon pool (gCm^{-2})	2.838×10^{2}	2.193×10^{2}	$10 - 1000$
C_{woo}	Above and below ground woody carbon pool (gCm^{-2})	6.506×10^{3}	7.143×10^{3}	$100 - 10^{5}$
C_{lit}	Litter carbon pool (gCm^{-2})	5.988×10^{2}	5.450×10^{2}	$10 - 1000$
C_{som}	Soil and organic carbon pool (gCm^{-2})	1.936×10^{3}	1.276×10^{3}	$100 - 2 \times 10$

Table 2: Parameter values and standard deviations for background vector used in experiments.