An improved analysis of forest carbon dynamics using data assimilation

MATHEW WILLIAMS*, PAUL A. SCHWARZ†, BEVERLY E. LAW†, JAMES IRVINE† and MEREDITH R. KURPIUS†

School of GeoSciences and NERC Centre for Terrestrial Carbon Dynamics, Darwin Building, University of Edinburgh, Edinburgh EH9 3JU, UK, †College of Forestry, Oregon State University, Corvallis, OR 97331, USA

Abstract

There are two broad approaches to quantifying landscape C dynamics – by measuring changes in C stocks over time, or by measuring fluxes of C directly. However, these data may be patchy, and have gaps or biases. An alternative approach to generating C budgets has been to use process-based models, constructed to simulate the key processes involved in C exchange. However, the process of model building is arguably subjective, and parameters may be poorly defined. This paper demonstrates why data assimilation (DA) techniques – which combine stock and flux observations with a dynamic model – improve estimates of, and provide insights into, ecosystem carbon (C) exchanges. We use an ensemble Kalman filter (EnKF) to link a series of measurements with a simple box model of C transformations. Measurements were collected at a young ponderosa pine stand in central Oregon over a 3-year period, and include eddy flux and soil CO_2 efflux data, litterfall collections, stem surveys, root and soil cores, and leaf area index data. The simple C model is a mass balance model with nine unknown parameters, tracking changes in C storage among five pools; foliar, wood and fine root pools in vegetation, and also fresh litter and soil organic matter (SOM) plus coarse woody debris pools. We nested the EnKF within an optimization routine to generate estimates from the data of the unknown parameters and the five initial conditions for the pools. The efficacy of the DA process can be judged by comparing the probability distributions of estimates produced with the EnKF analysis vs. those produced with reduced data or model alone. Using the model alone, estimated net ecosystem exchange of C (NEE) = $-251 \pm 197\,\mathrm{g\,C\,m^{-2}}$ over the 3 years, compared with an estimate of $-419 \pm 29\,\mathrm{g\,C\,m^{-2}}$ when all observations were assimilated into the model. The uncertainty on daily measurements of NEE via eddy fluxes was estimated at $0.5\,\mathrm{g\,C\,m^{-2}\,day^{-1}}$, but the uncertainty on assimilated estimates averaged $0.47\,\mathrm{g\,C\,m^{-2}\,day^{-1}}$, and only exceeded $0.5\,\mathrm{g\,C\,m^{-2}\,day^{-1}}$ on days where neither eddy flux nor soil efflux data were available. In generating C budgets, the assimilation process reduced the uncertainties associated with using data or model alone and the forecasts of NEE were statistically unbiased estimates. The results of the analysis emphasize the importance of time series as constraints. Occasional, rare measurements of stocks have limited use in constraining the estimates of other components of the C cycle. Long time series are particularly crucial for improving the analysis of pools with long time constants, such as SOM, woody biomass, and woody debris. Long-running forest stem surveys, and tree ring data, offer a rich resource that could be assimilated to provide an important constraint on C cycling of slow pools. For extending estimates of NEE across regions, DA can play a further important role, by assimilating remote-sensing data into the analysis of C cycles. We show, via sensitivity analysis, how assimilating an estimate of photosynthesis – which might be provided indirectly by remotely sensed data – improves the analysis of NEE.

Correspondence: Mathew Williams, fax + 44 131 662 0478,
e-mail: mat.williams@ed.ac.uk

Introduction

Quantifying the carbon (C) dynamics of the terrestrial biosphere is a current and major concern for earth system science. A major uncertainty concerns identifying whether regions or landscapes are sources or sinks for C. There are two broad approaches to quantifying landscape C dynamics – by measuring changes in C stocks over time, or by measuring fluxes of C directly. Stock inventories involve recording the mass of C in living biomass (leaves, stems, and roots), and in litter and soil pools, and quantifying the changes in pool sizes between sampling times (Turner *et al.*, 1995; Malhi *et al.*, 2002). The advantage of this approach is that measurements are generally cheap and simple, although labour intensive. The difficulty is that pools are spatially patchy (e.g., a large proportion of C can be in a few large tree stems), belowground pools (such as fine roots, root litter, soil C) are difficult to measure, and that monitoring is generally restricted to small plots by logistical constraints.

Direct determination of C fluxes has been revolutionized by the development of automated measurement systems of net carbon dioxide (CO_2) and water vapour exchange between land and atmosphere, via the eddy covariance technique (Baldocchi, 2003). Cuvettes and chamber measurements of soil effluxes, stem and foliage respiration, and leaf photosynthesis have enhanced the understanding of processes contributing to the eddy fluxes (Law *et al.*, 1999a, 2001a). Flux data can be generated somewhat continuously (e.g., half-hourly), and the instruments sample a 'footprint' of the surrounding landscape covering a few km^2. However, these data often have gaps, and filling these may introduce biases, and increase uncertainty. Also, night-time flux data can be biased when winds are light and intermittent (Goulden *et al.*, 1996), and complex terrain may jeopardize some of the assumptions of the approach (Finnigan *et al.*, 2003). Finally, the scale of measurement is often uncertain because the footprint varies depending on wind speed and direction (Schmid & Lloyd, 1999).

An alternative approach to generating C budgets has been to use process-based models, constructed to simulate the key processes involved in C exchange (Farquhar & von Caemmerer, 1982; Jarvis *et al.*, 1985; Parton *et al.*, 1988). The advantage of using models is that they can be extended across large spatial domains and into the future, given the relevant driving variables (Running *et al.*, 1999; White *et al.*, 2000; Rastetter *et al.*, 2003). Forecasts are possible because models incorporate a representation of the simulated system and its dynamics. Rules, such as the conservation of mass, can be enforced to guide system trajectories. The disadvantage of modelling is that the process of model construction is arguably subjective. Occam's razor – making models as simple as possible, but no simpler – is a useful guiding principle. But there is always a danger that the model's representation of the system is not accurate. Other problems include parametrization – setting the rate constants on fluxes in a compartment-flow model, for instance. Generally these parameters are unknown and have to be derived from data, somehow. There is always a danger that poorly defined parameters will be 'tuned' to give good output. When several parameters are tuned, the right answer may be generated for the wrong reason (Williams *et al.*, 2001a).

Generally, scientific papers on C budgets can be classified into stock (Phillips *et al.*, 1998), flux (Wofsy *et al.*, 1993), or model approaches (McGuire *et al.*, 1993), following the definitions given above. However, some papers do attempt to combine the methods. Ecological inventories are now being compared with micrometeorological data (Ehman *et al.*, 2002), and models are often corroborated against observations (Running, 1994; Williams *et al.*, 1996; Law *et al.*, 2003).

Models are generally parametrized with some subset of observational data, and tested against remaining data. Such tests are designed to show that the model can effectively describe the observed system by demonstrating a strong correlation, or a low mean error, between prediction and observation. This standard modelling approach assumes a primacy of the data over the model representation. But if this is the case, then the standard approach is inefficient. It would be much more worthwhile to use all the available data to improve the model and minimize confidence limits on predictions, rather than only a subset. Here we demonstrate the technique of data assimilation (DA), combining all available data with a model, to develop a full analysis of C cycling in a ponderosa pine forest.

Objectives

The objective of this paper is to demonstrate why the application of DA techniques results in improved

estimates of ecosystem C stocks and fluxes, with reduced uncertainty compared with the original observations, or the model alone. The argument of this paper is that combining measurements and modelling through DA generates more precise estimates of C dynamics, and simultaneously highlights areas where model improvement is required.

Methods

The premise of DA is that neither models nor observations can perfectly describe a system, but an analysis that combines model and data will provide a better estimate of system dynamics than model or observations alone. DA is a process for the optimal combination of information about a system, which evolved from the engineering approaches to filtering and control theory applied in missile guidance and interception (Maybeck, 1979). DA has been applied in meteorology for forecasting (Lorenc, 1986), and extended into stream ecology (Cosby, 1984), oceanography (Eknes & Evensen, 2002), and soil science (Heuvelink & Webster, 2001) for time series analysis.

DA is the process of finding the model representation that is most consistent with observations (Lorenc, 1995). DA recognizes that there are never sufficient observations to represent the state of a system at any one time. For a detailed, complete picture, further information is required, such as knowledge of the behaviour and probable structure of the system. In DA, knowledge of system evolution in time is usually embodied in a model. In sequential assimilation, the approach we demonstrate here, the model organizes and propagates forward information from previous observations (Lorenc, 1995). When new information becomes available, the prediction, or forecast, of the model can be compared with these observations and corrected. A poor model will drift and will be frequently and heavily corrected; an effective model will require little reinitialization by observations. However, it is not simply a question of fitting the model to the new data, as the assimilation process must also conserve the information provided by the model itself and by previous observations.

The DA technique that we use here is the Kalman filter (KF) (Kalman, 1960), which has been widely used (Grewal, 1993), and, given various assumptions, has been shown to be an optimal, variance-minimizing analysis (Maybeck, 1979). The basic KF requires three assumptions: that a linear model can describe the system, and that the system and measurement noise are both white and Gaussian. Developments of the basic KF have provided means to deal with deficiencies in these assumptions (Grewal, 1993). The product of the KF is an estimate, or analysis, of the state variables that takes account of prior knowledge plus new observations to ensure that estimated errors are statistically minimized.

The KF generates a variance-minimizing analysis by combining the model forecast with the observations, weighted according to these prediction and measurement error covariances. If measurement noise is large, relatively less emphasis is placed on the current observation. If measurement noise is small, or the model estimation error is large, the corrected estimate approaches the observation. As inputs, the KF requires details on the covariances of both the model forecast and of the measurements. The error covariance of the measurements can be derived from knowledge of the accuracy of the techniques used to measure them. The error covariances for the model prediction are computed by solving an equation for the evolution in time of the error covariance matrix of the model state (Evensen, 2003).

The KF uses covariance matrices to store information on the uncertainty in the models, observations, and analysis. Storing, integrating, and inverting large covariance matrices is computationally expensive and the matrix operations required for the analysis are not always robust. Evensen (1994) suggested that, instead of storing a full covariance matrix, the same error statistics can be represented approximately using an appropriate ensemble of model states. The ensemble KF (EnKF, Evensen, 1994) uses a Markov Chain Monte Carlo method to solve the time evolution of the probability density of the model state. The probability density is represented by a large (ca. 100–1000) ensemble of model states, and these are integrated forward in time by a differential equation (i.e., the model) with a stochastic forcing term representing the model errors. Each ensemble member evolves in time according to

$$\psi_j^{k+1} = M(\psi_j^k) + dq_j^k, \qquad (1)$$

where ψ is the state vector, j counts from 1 to N, where N denotes the number of model state ensemble members, k denotes time step, M is the model operator or transition matrix, and dq is the stochastic forcing representing model errors from a distribution with mean zero and covariance Q (Burgers et al., 1998).

Similarly, observations are treated as random variables by generating an ensemble of observations from a distribution with the mean equal to the measured value and a covariance equal to the estimated measurement error (Burgers et al., 1998). Thus, we define the new observations

$$d_j = d + \varepsilon_j, \qquad (2)$$

where d are the observations, and ε are drawn from a distribution of zero mean and covariance equal to the

estimated measurement error. The analysis step of the EnKF updates each of the model state ensemble members using the following equation:

$$\psi_j^a = \psi_j^f + K_e(d_j - H\psi_j^f), \tag{3}$$

where ψ^f is the forecast state vector (i.e., the prediction) and ψ^a is the analysed estimate generated by the correction of the forecast. H is the observation operator, a matrix that relates the model state vector to the data, so that the true model state is related to the true observations by

$$d^t = H\psi^t. \tag{4}$$

K_e is the KF gain matrix, which determines the weighting applied to the correction (Burgers *et al.*, 1998).

We use the EnKF based on its ability to predict error statistics for strongly nonlinear systems and for its simplicity and numerical stability (Evensen, 2003). The analysis generated by the EnKF is a combination of the model forecast and a number of influence functions, one for each of the measurements (Allen *et al.*, 2002). These influence functions are computed from the ensemble statistics, and thus include cross-correlations between the different variables in the model. The influence functions summarize the correlations between model state variables that are determined from the ensemble of model runs. In effect, the ensemble serves as a large sensitivity analysis, quantifying how small changes in each state variable affect system dynamics. Changes to one variable resulting from a new observation are transmitted to other model variables according to the strength of these cross-correlations. This is a particularly powerful component of the EnKF – information on a single observation is transmitted to all state variables via the connections set out in the model. The FORTRAN code required to perform the EnKF is provided in Evensen (2003).

The study area

The Metolius young ponderosa pine site is located in a Research Natural Area (44°26′N, 121°34′W, elevation 1165 m) in the eastern Cascades, near Sisters, Oregon. The site was clear cut in 1978, and has regenerated naturally since then, with some recent thinning. In 2002, there were 431 trees ha^{-1}, with a mean diameter at breast height (DBH) of 11.3 cm. The understorey vegetation is sparse with patches of bitterbrush (*Purshia tridentata*) and bracken fern (*Pteridium aquilinum*), and a groundcover of strawberry (*Fragaria vesca*). The site is in a semiarid region that experiences warm, dry summers and wet, cool winters. While the flux tower samples a variable footprint, most of the remaining data were

collected within a 100 m × 100 m plot located ~50 m upwind of the tower.

Flux measurements

A variety of flux data were used to generate daily estimates of specific C fluxes for use in the analysis. We made continuous eddy covariance measurements to determine half-hourly fluxes of CO_2 throughout 2000–2002 at 12 m height, ~9 m above the canopy (Law *et al.*, 1999a,b; Anthoni *et al.*, 2002). Data were screened to remove possible eddy covariance instrumentation and sampling problems (e.g., friction velocity, $u^* < 0.2\,m\,s^{-1}$), and fluxes were also rejected when unreasonably large CO_2 fluxes ($|F_c| > 25\,\mu mol\,m^{-2}\,s^{-1}$) were observed. CO_2 concentration was measured every 30 min at 1, 3, and 12 m above the ground. The buildup or release of CO_2 from within the canopy was quantified by determining the rate of change of CO_2 along the vertical profile. The total CO_2 flux was then calculated as the sum of CO_2 flux and this change in storage. Data gaps were filled based on seasonal empirical relationship with environmental variables (PAR, VPD) derived from valid flux data (Anthoni *et al.*, 1999). We generated daily net ecosystem exchange of CO_2 (NEE) data for days in which less than 25% of the 48 possible half-hour measurements were gap-filled; for the 3-year period of this study, this amounted to 684 daily NEE values.

Other periodic gas exchange measurements included foliage and stem respiration, and soil surface CO_2 fluxes (Law *et al.*, 1999b). Soil respiration was measured using six automated chambers installed in 1999 (Irvine & Law, 2002); total daily effluxes were recorded on 401 days during 2000–2002. Root contributions to soil C effluxes were measured by recording CO_2 efflux with a manual system, removing a 30 cm soil core, extracting the roots, and then measuring root CO_2 effluxes (methods in Law *et al.*, 2001a). These soil/root data were collected on 3 days in 2000 (days 153, 201, and 278), and 2 days in 2001 (115 and 227). Pine foliage respiration was determined using a temperature function fitted to 12 nights of cuvette data collected during 1999 and 2001 (Law *et al.*, 1999b); shrub foliage respiration rates were determined similarly using five nights of data (Law *et al.*, 2001a). Site-level estimates of total foliar respiration were determined from air temperature data, temperature response functions for trees and shrubs, and partitioning of estimated site leaf area index (LAI) between trees and shrubs. We estimated sapwood respiration at the *Y* site using a temperature response function developed from data at a nearby old-growth site, using sapwood volume estimates derived from tree rings at the young site

(Law et al., 1999a, 2001a). Because sapwood of young trees has higher respiration rates than old trees, we likely underestimated the respiration loss from the system, but in the old forest, sapwood respiration accounted for only 10% of total ecosystem respiration (Law et al., 1999a).

We limited estimates of total ecosystem respiration to those 401 days where soil measurements were available. Estimating foliar and stem respiration on these days required temperature and biomass data for each of these days. Errors will be dominated by uncertainty in the foliar biomass and sapwood volume estimates, although the latter should be relatively small. To estimate total autotrophic respiration, we combined estimates of the root fraction of soil effluxes with foliar and stem respiration measurements. Because there were no automated soil chamber data within 55 days of the root respiration measurement on day 115 in 2001, we discarded this data point. Thus, over the 3-year period there were 4 days when we were able to produce estimates of autotrophic respiration.

To provide independent estimates of gross primary production (GPP) (direct observations were lacking), we used the soil-plant-atmosphere (SPA) model (Williams et al., 1996), which is based on the underlying biochemistry of carboxylation, of light interception, gaseous CO_2 exchange, and the impacts of soil moisture on stomatal conductance. The SPA model has multiple canopy layers and a 30 min time step, and has been previously parametrized and applied in ponderosa pine ecosystems (Williams et al., 2001b), generating reasonable estimates of C fixation throughout the annual cycle. The SPA model requires as driving variables a daily phenology of LAI, foliar N, and root biomass, which we generated from interpolations of the limited data available over time. We used measured estimates of maximum carboxylation and electron transport rates on pine foliage at the site to parametrize the model (Law et al., 2003), and time series of sap flow data to parametrize the hydraulic characteristics of the trees (Schwarz et al., in press). The close correspondence of SPA predictions of daily transpiration with sap flow estimates (R^2 varied from 0.82 to 0.87 over the three years) means that we have confidence in the SPA predictions of stomatal opening (Schwarz et al., 2004). The SPA model, in effect, translates sap flow observations into GPP estimates, which are then assimilated. This means that NEE data, respiration data, and GPP pseudo-observations can all be assimilated into the analysis and their consistency can be determined.

Stock measurements

We estimated LAI (one-sided LAI) from optical measurements with an LAI-2000 plant canopy analyzer (LI-COR, Lincoln, NE, USA), on a 10 m grid within the 100 m × 100 m plot. We also measured needle clumping within shoot, and clumping at scales larger than shoot to correct for these effects on LAI estimates. LAI data were collected at four times through the 3-year period of this study – once in 2000 and in 2001, and twice in 2002. Foliar mass was determined from LAI and direct measurements of C mass per unit leaf area on foliage samples.

On five 10 m radius subplots within the main plot, we recorded dimensions of all trees (stem height, and DBH, 1.37 m) and shrubs (length, width, height, and diameter at shrub base). Aboveground biomass and coarse root mass were then determined from allometric relationships derived from destructive harvest of five trees covering the range of sizes present (Law et al., 2001b). Likewise, shrub biomass was determined from the destructive harvest of five to nine shrubs per species and scaled to the site by the number of shrubs per size class (Law et al., 2001b). Dimensions were sampled at three times during the 3-year period of the study, once in 2000 and twice in 2002.

Litterfall production (<1 cm diameter) was determined from monthly collections of litterfall in 20 trays (0.13 m² each); litter was separated into foliage and woody material (Law et al., 2001b). Litterfall was only collected during the latter half of each of the 3 years, when the majority of litterfall occurs, resulting in around 18 months of data over 3 years. We estimated daily litterfall for both foliage and woody material on the basis of a constant rate across the month.

Fine root (≤2 mm diameter) biomass was estimated from soil cores collected on 2 days during 2002 (days 134 and 219). On each occasion, cores were extracted from 18 different locations (six sampling points along three transects) using a 7.2 cm diameter corer. At each sampling location, samples were divided into three layers: 0–20, 20–50, and 50–100 cm depth. All samples were refrigerated until processed in the laboratory, at which time the roots were separated from the surrounding soil and sorted according to diameter class and then further separated into live and dead categories based on visual inspection of each root segment. Following the separation and sorting, the root samples were dried at 70 °C for 48 h and then weighed to measure biomass.

C storage in soil was determined from 27 soil cores to 1 m depth. Live vegetation and roots were removed and total C was determined to 1 m depth (Law et al., 2003). Coarse woody debris (CWD) was measured in four 75 m line transects, and fine wood debris was recorded on four 15 m transects per subplot for all four subplots, following the protocol of Harmon & Sexton (1996). Six samples of forest floor fine litters

Table 1 Estimated errors associated with each component of the model and with each set of observations, and the total number of daily observations available

Symbol	Model error	Observational error	Number of observations	Pool/flux description
C_f	20%	10%	4	Foliar C mass
C_w	20%	10%	3	Wood C mass
C_r	20%	30%	2	Fine root C mass
C_{lit}	20%	30%	1	Fresh litter C mass
$C_{SOM/WD}$	20%	30%	1	Soil organic matter plus woody debris C mass
A_f	20%	N/A		Foliage allocation rate
A_w	20%	N/A		Wood allocation rate
A_r	20%	N/A		Fine root allocation rate
L_f	$0.5\,g\,C\,m^{-2}\,day^{-1}$	20%	18	Foliage litter production rate
L_w	$0.5\,g\,C\,m^{-2}\,day^{-1}$	20%	18	Wood litter production rate
L_r	$0.5\,g\,C\,m^{-2}\,day^{-1}$	N/A		Fine root litter production rate
D	20%	N/A		Decomposition rate of litter
G	20%	30%	1096	Gross primary production
R_{tot}		20%	401	Total respiration rate
R_a	20%	50%	4	Autotrophic respiration rate
R_{h1}	20%	N/A		Heterotrophic respiration rate of fresh litter
R_{h2}	20%	N/A		Heterotrophic respiration rate of SOM/WD
NEE	N/A	$0.5\,g\,C\,m^{-2}\,day^{-1}$	684	Net ecosystem exchange rate of C

For each time step, the predictive error on each pool and flux is drawn from a normal distribution of mean zero and with a standard deviation defined as a percentage of the mean pool or flux at that time. For modelled litter fluxes and observed net ecosystem exchange (NEE), an absolute and constant value is set instead (see text).

were collected to determine nonwoody litter mass (Law *et al.*, 2003).

Confidence limits on observations

Correct estimation of observational error is crucial to the quality of the analysis (Table 1), because the magnitude of observational errors determines to what extent the simulated fields will be corrected to match the observations. Observation error variances are best specified by knowledge of instrumental characteristics, and by comparing replicated samples. Observation error correlations are usually assumed to be zero – distinct measurements are assumed to be affected by physically independent errors. In most cases, we specify the standard deviation of model error as a percentage of the mean.

Errors in the total respiration estimates arise from the limited sampling of the automated soil chambers in space, and the use of empirical relationships to scale foliage and stem respiration in time. Because soil respiration is the largest component of ecosystem respiration, we set the coefficient of variation in total respiration in accordance with the reported values for the soil chamber data at the study site, ~ 0.2 (Irvine & Law, 2002). The uncertainty of autotrophic respiration

estimates must be relatively larger than that of total respiration, because the autotrophy is estimated by a highly invasive approach, so errors are set to 50%. A detailed uncertainty analysis of modelling of GPP suggests errors of up to 30% (Williams *et al.*, 2001c), and this value was assigned to SPA model estimates of GPP. An analysis of the uncertainties in calculating daily NEE from eddy covariance (Anthoni *et al.*, 1999) suggests an error on daily estimates of 12%. The uncertainty because of potential advection at night and during early morning and late evening transition periods could not be assessed. In addition, recently discovered software error in the LI7500 flux instrument suggests that the flux estimates of NEE are biased high by about 20% (biased towards a stronger sink for CO_2; LI-COR Inc.). However, the nature of NEE (it can be positive or negative) means that defining errors by a coefficient of variation is unsuitable, so, instead, errors are set at $0.5\,g\,C\,m^{-2}\,day^{-1}$.

Errors on pool measurements (Table 1) were determined from the variation in replicated measurements. The coefficient of variation on LAI observations at the site was 10%, and stem sampling indicated a standard error on wood biomass equal to 10% of the mean (Law *et al.*, 2001b). Based on replicated soil cores, we estimate the error on fine root biomass at 30% (data not shown).

Models

Description of C dynamic model

We represent the C cycle with a simple box model of pools connected via fluxes (Fig. 1). There are five pools – C content of foliage (C_f), woody stems and coarse roots (C_w) and fine roots (C_r), and of fresh leaf and fine root litter (C_{litter}) and soil organic matter (SOM) plus WD ($C_{SOM/WD}$). There is one pseudopool, the daily accumulation of photosynthate (GPP) that is entirely utilized each day. The fluxes among pools are based on the following assumptions:

1. All C fixed during a day is either expended in autotrophic respiration or else allocated to one of three plant tissue pools – foliage, wood, or fine roots.
2. Autotrophic respiration is a fixed fraction of total photosynthetic fixation (Thornley & Cannell, 2000), and it is not directly temperature sensitive.
3. Plant allocation and litterfall are donor-controlled functions with no direct environmental influence and constant rate parameters.
4. Soil transformations are sensitive to temperature, with a Q_{10} of 2.0. Otherwise, the only environmental forcing in the C model is on GPP, via solar radiation, air temperature, and soil moisture.
5. All C losses are via mineralization; there is no dissolved loss term.

The aggregated canopy model (ACM) of photosynthesis (Williams et al., 1997) provides the forecast estimate of daily C inputs to the system. The ACM is a big-leaf, daily time step model that estimates GPP as a function of LAI, foliar nitrogen, total daily irradiance, maximum and minimum daily temperature, day length, atmospheric CO_2 concentration, soil–plant water potential, and total soil–plant hydraulic resistance. The model has 10 parameters, and we use a local model calibration (see Appendix). Using measurements of leaf C mass per area (111 g C m^{-2} leaf area), estimates of LAI for the ACM are determined directly from the C_f pool. For simplicity we assume a constant value for foliar N concentration (2.7 g N m^{-2} leaf area). To characterize the changes in moisture limitation through the summer, we drive the ACM model with estimates of soil–leaf water potential difference and hydraulic resistance, both derived from a detailed ecophysiological study undertaken at the site using the SPA model (Schwarz et al., 2004).

Initial conditions and rate constants

The C box model has nine unknown parameter constants (Table 2), and the initial values of the five C pools are also unknown (Table 3). To generate estimates of parameters and initial conditions, we nested the EnKF within an optimization routine. This routine varied the unknown parameters and initial conditions to find the
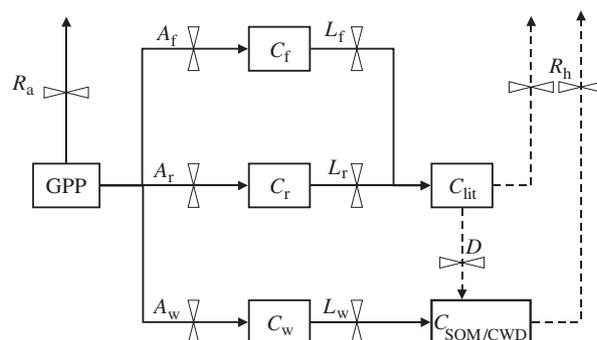

Fig. 1 C dynamic model, showing pools (boxes) and fluxes (arrows). Allocation fluxes are marked A, litterfall fluxes by L, and respiration by R (split between autotrophic (a) and heterotrophic (h)). D is a decomposition flux and GPP is gross primary production. Temperature-controlled fluxes are marked by dashed lines.

Table 2 Estimated values of the model parameters as calculated by the optimization routine

Parameter	Description	Model value	Low/high bounds
t_1	Decomposition rate constant	4.4×10^{-6}	$1 \times 10^{-6}/0.01$
t_2	Autotrophic respiration as a fraction of GPP	0.47	0.2/0.7
t_3	Fraction of NPP allocated to foliage	0.31	0.01/0.5
t_4	Fraction of NPP allocated to fine roots	0.43	0.01/0.5
t_5	Turnover rate of foliage	2.7×10^{-3}	$2 \times 10^{-4}/0.02$
t_6	Turnover rate of woody matter	2.06×10^{-6}	$2 \times 10^{-6}/0.02$
t_7	Turnover rate of fine roots	2.48×10^{-3}	$2 \times 10^{-4}/0.02$
t_8	Mineralization rate of fresh litter	2.28×10^{-2}	$5 \times 10^{-5}/0.5$
t_9	Mineralization rate of soil organic matter and woody debris	2.65×10^{-6}	$1 \times 10^{-6}/0.5$

Low and high bounds for each parameter, which define the search space are also shown.
NPP, net primary production; GPP, gross primary production.

Table 3 Estimated values of the initial conditions in each dynamic C pool

Symbol	C pool	Initial value ($g\,C\,m^{-2}$)	Low/high bounds
C_f	Foliage	58	30/70
C_w	Wood (stems and coarse roots)	770	750/900
C_r	Fine roots	102	50/300
C_{lit}	Fresh foliar and fine root litter	40	20/3000
$C_{SOM/WD}$	Soil organic matter plus woody debris	9897	6000/10 000

Low and high bounds for each pool, which define the search space.

values that minimized the sum-of-squared differences of the innovations (i.e., the difference between model forecast/prediction and observations) for all available observations (Table 1). This approach, in effect, undertakes numerous implementations of the EnKF with varied initial values and parameters, and identifies in which implementation the predictions require the minimum correction. This implementation is then assumed to have the optimal parameter set and initial conditions.

The optimization routine uses the quasi-Newton method and a finite difference gradient, UMINF, IMSL Math Library. We set specific bounds to the parameters and initial conditions to ensure realistic values (Tables 2 and 3). We used the optimized parameters in all the subsequent analyses, although we also undertook a sensitivity analysis to find how variation in parameters affected the analysis, and also affect model forecast bias.

Model error variances and ensemble size

The model errors determine how rapidly the uncertainty on the forecast of the state vector grows over time. The errors are determined by random disturbances to the system, errors in measured inputs (drivers), and the error inherent in representing a complex system as a simple model (Cosby *et al.*, 1984). The standard deviations of model error can be expressed as a fraction of mean values, and this is useful as it largely restricts the ensembles to positive, and thus meaningful, values when the mean of the particular ensemble is close to zero (e.g., GPP during winter). We set the model uncertainties to 20% (Table 1), and undertook sensitivity analyses to explore the implications of changes in these values. As an exception, we assigned a relatively large, constant error to litterfall simulations (Table 1). The modelling of litterfall uses a simple rate constant, and so does not simulate stochastic events, such as storms, or periodic events, such as leaf senescence, that cause pulses of litter. By setting a high and constant error on litterfall modelling, any available observations are used to form the majority of the analysis, and we are explicitly recognizing the weakness of our litterfall submodel.

The ensemble size of the EnKF is the number of model states that are concurrently predicted and analysed. The differences between the analyses form a statistical sample of analysis errors. We set ensemble size to 200, large enough to ensure the correct estimate of the error variances in the predicted model state (Allen *et al.*, 2002).

Results

Parameter estimates and initial conditions

The fitting process generated estimates of the nine model parameters and five initial conditions using all the available flux and stock data. The optimization estimated that 47% of GPP was respired by plants each day (Table 2), and 53% remained for net primary production (NPP). Of NPP, 31% was allocated to foliage each day, 43% to fine roots, and 25% to woody stems and coarse roots. The parametrized daily turnover rates of plant tissues (Table 2) suggest a leaf life span of 1.0 years, a fine root life span of 1.1 years, and a life span of woody materials of 1323 years. The parametrized estimate of fresh fine litter turnover, via mineralization and decomposition combined, was 0.1 years at a constant 10 °C. Parametrized turnover of SOM and WD was 1033 years, at a constant 10 °C.

The estimated initial conditions (Table 3) assigned a smaller C mass to foliage than to fine roots, but 83% of the initial vegetation C pool was allocated to woody biomass. The C mass of CWD and SOM are more than an order of magnitude greater than total plant C biomass.

Analysis of carbon ecosystem exchanges

Over the 3 years, 2000–2002, the analysis suggests a total NEE of $-419 \pm 29\,g\,C\,m^{-2}$ (i.e., a clear C sink) (Table 4). The total estimated GPP over the 3 years was $2172 \pm 18\,g\,C\,m^{-2}$, with 2002 being the most productive year (Table 4). The analysis suggests that autotrophic sources contributed to 58% of total respiration (Table 4)

Table 4 Analysis of annual C fluxes ($g\,C\,m^{-2}\,yr^{-1}$) for 3 years, and their sum total

Flux	2000	2001	2002	Total	SD	R^2	N	RMSE
Gross primary production	716	702	753	2172	18	0.95	1096	0.32
Total respiration	560	585	608	1753	33	0.95	401	0.23
Net ecosystem exchange	−156	−117	−145	−419	29	0.95	684	0.20
Net primary production	378	369	409	1156	20	N/A	N/A	N/A
Litter production	235	242	265	742	51	N/A	N/A	N/A
Autotrophic respiration	339	333	334	1016	18	N/A	N/A	N/A
Heterotrophic respiration	221	252	264	737	36	N/A	N/A	N/A

Standard deviation on the 3-year analysis in the ensemble of model runs (SD), the fraction of the daily variation of the observations explained by the analysis (R^2), the number of daily observations (N), and the root-mean-square error (RMSE) of the daily analysis over 3 years.
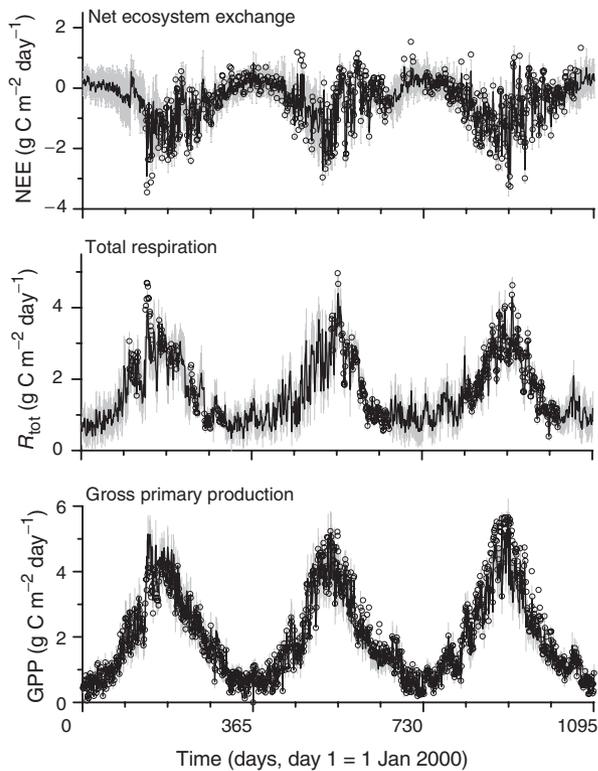

Fig. 2 Daily observations (∘) and analyses (lines) of net ecosystem exchange of CO_2 (NEE, negative flux = net ecosystem uptake), total respiration (R_{tot}), and gross primary production (GPP) for 3 years, 2000–2002. Grey vertical lines indicate the standard deviation around the mean of the ensembles, a measure of the confidence limits of the analysis.

Fig. 3 Daily analysis of total heterotrophic respiration (R_h, top panel), and intermittent observations (symbols, ∘, with standard error bars) and complete analyses (lines) of autotrophic respiration (R_a, bottom panel) over 3 years, 2000–2002. Grey lines indicate the standard deviation around the mean of the ensembles.

and that litter production and heterotrophic respiration were of similar magnitude each year.

The daily analysis of NEE, GPP, and total respiration generally closely track their corresponding observations (Fig. 2) and have root-mean-square errors (RMSE) in the range 0.20–0.32 $g\,C\,m^{-2}\,day^{-1}$ (Table 4). Most of the NEE data lie within one standard deviation of the analysis of NEE (Fig. 2, top panel). Similarly, measure-
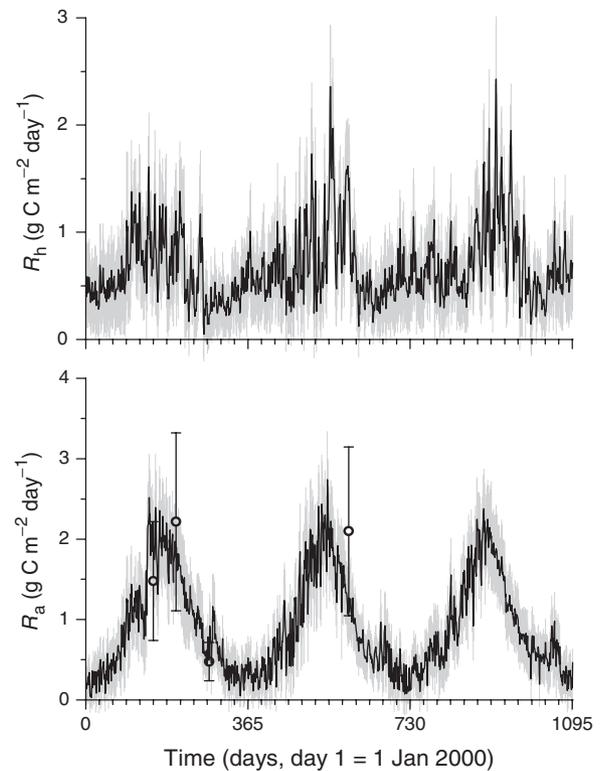
ments of autotrophic respiration all lie within one standard deviation of the associated analysis (Fig. 3). Cosby *et al*. (1984) suggest tests for model adequacy (i.e., that all the component matrices of the KF, Eqns (1)–(4), are correctly specified). The tests analyse the innovation sequence specified by

$$d_j - H\psi_j^f$$

and require that the sequence has a zero mean and is serially uncorrelated. Innovations from the model forecast of NEE had a zero mean ($P > 0.05$), but the innovations were serially correlated ($P < 0.01$).

Analysis of C stock dynamics

The sparse nature of stock data means that the model forecast plays a greater role in the analysis, and observations correct the model for drift, particularly in slow cycling pools. The analysis of foliar C suggests a strong seasonal cycle (Fig. 4, top panel), and this is partly because of the direct link in the ACM model between foliar C (proportional to LAI) and GPP. The strong seasonal cycle in observed GPP is transmitted by the EnKF to the analysis of foliar C, and the high frequency of GPP data generates narrow confidence intervals on the analysis of foliar C, compared with intervals on the analyses of root or wood C. In three out of the four cases, the observed LAI lie within the uncertainty of the forecast. The analysis of wood C mass indicates a slow gain over 3 years. Each of the three observations of wood C lies close to or within one standard deviation of the forecast of the previous time step (Fig. 4). After an observation is assimilated, the uncertainty on the analysis is reduced, and then grows slowly with time. The analysis of fine root C suggests that mass doubles over the 3 years, and shows how reductions in uncertainty when observations are
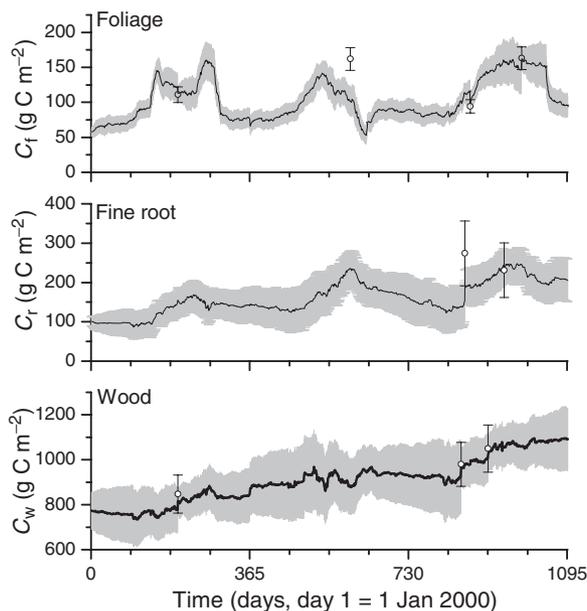
assimilated are less than for wood C (Fig. 4), because measurement uncertainty on fine roots is larger (Table 1). For fresh litter, no observations were available, but the analysis suggested an annual cycle varying between 10 and $60 \, g \, C \, m^{-2}$, with a peak in early spring and a minimum in late summer (data not shown). For SOM and woody litter, only a single observation was available, so the analysis is primarily governed by the model forecast (data not shown). The model suggests an increase in the SOM/WD over the 3 years. However, the limited data mean that confidence limits on the analyses for the litter pools are broad.

Analysis of litterfall

The analysis suggests that the total production of litter over 3 years is $742 \, g \, C \, m^{-2}$ (Table 4), with fine root litter dominating this flux ($436 \, g \, C \, m^{-2}$). Because the litterfall model is governed by simple turnover rates, and the uncertainty on the model was set high, the analyses of foliar and woody litterfall are dominated by the observations (Fig. 5) and have relatively broad confidence intervals (Table 4). Lacking any observations,
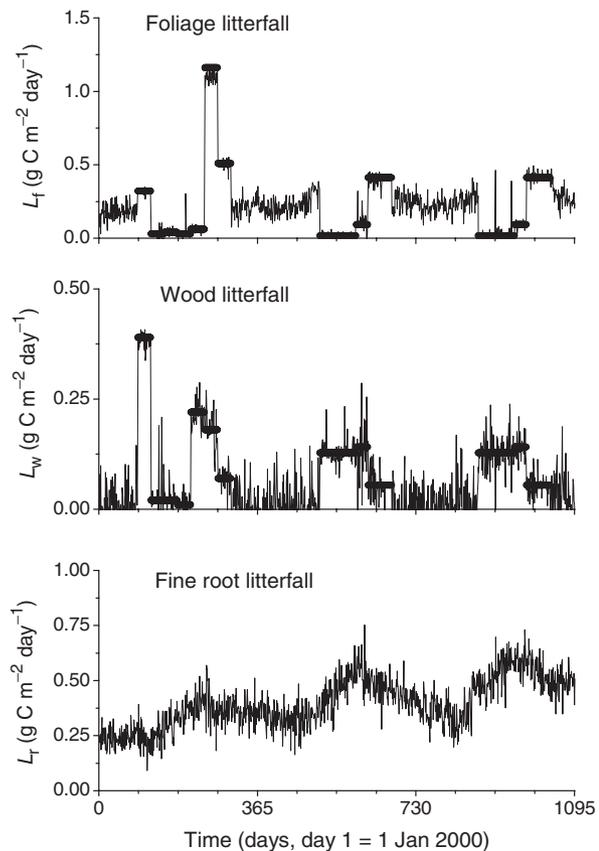
Fig. 4 Daily analysis (lines) and intermittent observations (symbols) of selected C stocks. Grey lines indicate one standard deviation around the mean of the ensembles. Errors bars (one standard error) are also given for each observation.

Fig. 5 Daily analysis (lines) and intermittent observations (symbols) of litterfall. For clarity, errors bars are omitted.

the analysis of fine root litter production is driven largely by the model, and is, thus, less seasonal.

Sensitivity analysis

We explored the sensitivity of the analysis to three different factors: (1) the number of time series observations assimilated; (2) variation in model error covariance; and (3) variation in model parameters.

We explored the sensitivity to the number of assimilated time series by reanalysis using the following: (a) the model without any DA, so that model forecasts are never corrected with data; (b) assimilating only GPP data into the model forecasts; and (c) assimilating GPP data and total respiration data only into the forecast (Fig. 6). Without any DA, the model forecasts were of NEE were poor (mean of the innovations was significantly different from zero, $P > 0.05$), and the total NEE prediction over 3 years was $-251 \pm 197 \, \mathrm{g\,C\,m^{-2}}$. The NEE prediction was a clear underestimate of observed summer sink strength. The confidence interval of one standard deviation encompasses most of the observations (Fig. 6a), suggesting that the model error estimate is reasonable. Once GPP data were assimilated into the forecast (Fig. 6b), the analysis was significantly improved. The estimate of NEE over 3 years was $-413 \pm 107 \, \mathrm{g\,C\,m^{-2}}$, and the mean forecast innovations were not significantly different from zero ($P < 0.05$). The analysis reproduced the seasonal pattern of NEE, but was poorer at predicting the high-frequency variation. When both GPP and total respiration data were assimilated (Fig. 6c), forecasts of NEE remained unbiased ($P < 0.05$), and there was a further reduction in the uncertainty on the NEE estimates (over 3 years $-472 \pm 56 \, \mathrm{g\,C\,m^{-2}}$).

We made changes to the values of the model and observational error covariances to explore the sensitivity of NEE analyses to these parameters (Table 5). A halving or a doubling of model error had little impact on the analysis. An order of magnitude reduction in model error slightly increased sink strength, reduced the estimated analysis error, and almost doubled the size of the RMSE of analysed NEE vs. observed NEE (RMSE). A tripling of model error resulted in a small reduction in sink strength and the RMSE. A halving or doubling of observational error had only a minor impact on NEE predictions and on the comparison of analysed with observed NEE. Reducing observational error to 10% of the nominal value decreased sink strength by only $19 \, \mathrm{g\,C\,m^{-2}}$ over 3 years. Tripling observational error increased sink strength by just $6 \, \mathrm{g\,C\,m^{-2}}$ over 3 years. The sensitivity analyses suggest
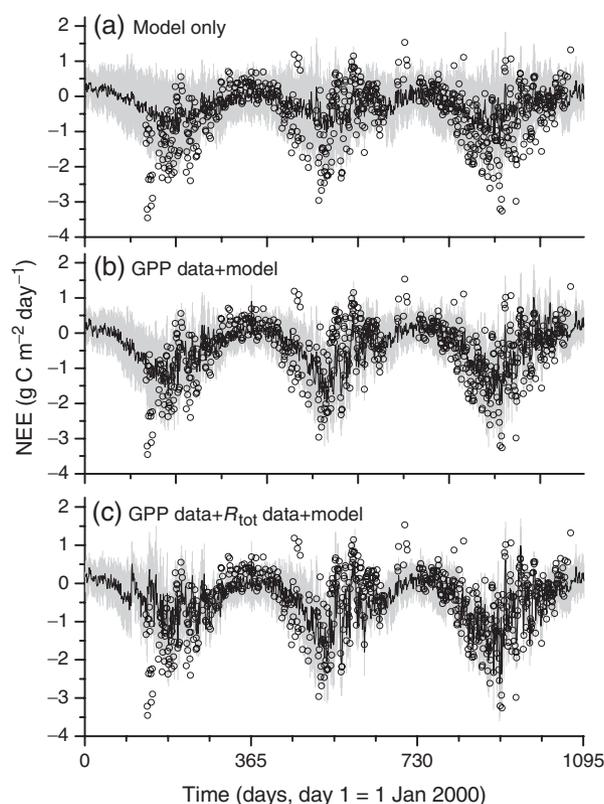

Fig. 6 Daily analysis (lines) of net ecosystem exchange (NEE) generated using (a) model only, no observations; (b) model plus gross primary production (GPP) data only; and (c) model plus GPP and total respiration data (R_{tot}). NEE observations are shown as open circles. Grey vertical lines indicate the standard deviation around the mean of the ensembles.

that there is little sensitivity in the NEE analysis to variations in estimated model and observational errors.

The nominal parameters and initial conditions used in the model (Tables 2 and 3) are not unique, and it is important to quantify how uncertainty in these values affects the analysis. We used a Monte Carlo technique to select alternative parameter sets and initial conditions, and used these in a multidimensional sensitivity analysis. For the model parameters ($N = 9$) and initial conditions ($N = 5$), we first generated, for each, distributions with means equal to the nominal values and variances equal to 20% or 50% of the mean. All parameters and initial conditions were sampled randomly from these distributions to generate 400 new, unique sets of parameters and initial conditions. We generated a new analysis with each of these unique parameter sets and initial conditions, and determined whether NEE forecasts were unbiased (innovations have a zero mean, $P > 0.05$), using the innovation test outlined in Results. Of the 400 analyses generated with

Table 5 Sensitivity analyses quantifying the impact of changes in estimated model and observation error covariances

Model error (% of nominal)	Observation error (% of nominal)	NEE (± SD)	R^2	RMSE
100	100	-419 ± 29	0.95	0.20
10	100	-440 ± 18	0.81	0.41
50	100	-423 ± 24	0.93	0.25
150	100	-417 ± 33	0.96	0.19
300	100	-414 ± 43	0.96	0.17
100	10	-400 ± 23	0.96	0.17
100	50	-414 ± 26	0.96	0.17
100	150	-422 ± 32	0.94	0.23
100	300	-425 ± 38	0.91	0.29

Changes were made to either all observational or all model errors, using a percentage adjustment to the nominal values (see Table 1). The table shows the predicted mean net ecosystem exchange (NEE) over 3 years $(g\,C\,m^{-2}) \pm 1$ standard deviation of the ensemble. The R^2 and root-mean-square-error (RMSE) $(g\,C\,m^{-2})$, are for the comparison of the analysis with the NEE observations. The first case is the default analysis.

20% variance, 189 had unbiased estimates of NEE, and, for the analyses with 50% variance, 75 were valid. For the 20% variance analyses, the mean NEE was $-421 \pm 17\,g\,C\,m^{-2}$, and for the 50% variance analyses the mean NEE was $-449 \pm 60\,g\,C\,m^{-2}$. Inspection of the 50% variance cases at the extremes of the range of NEE estimation revealed that, while NEE forecasts were unbiased, analyses of C stocks (e.g., wood C) were poor. The tests of bias and the analyses of NEE were most sensitive to changes in the parameter t_2, the parameter that determines autotrophic respiration as a fraction of GPP.

Discussion

The quality of the analysis

The tests on the innovations show that the model forecast is an unbiased predictor of NEE, although autocorrelations in the innovations show that the errors are not white. These autocorrelations suggest that Eqns (1)–(4) are not perfectly specified, although this might be expected given the extreme simplicity of the C model we used. Iterative improvements to the model should be undertaken in the light of these statistical tests of adequacy. For the current analysis, and carefully avoiding any prediction into the future, the lack of bias in the forecasts means that the results should still provide a valid basis for discussion.

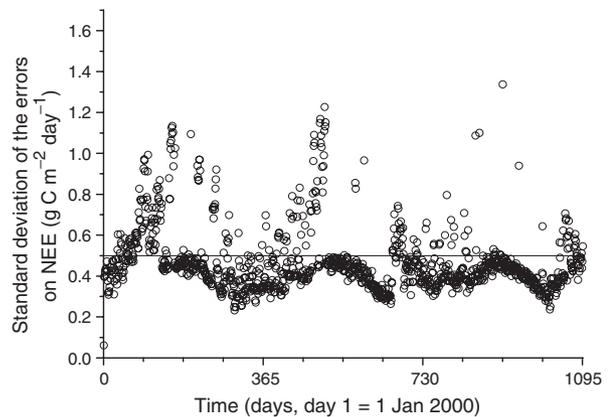The sensitivity analysis shows how the addition of extra data sets, particularly time series of photosynth-

Fig. 7 Daily variation in the standard deviation around the mean of the ensembles on net ecosystem exchange (NEE). Symbols show the error on the assimilated estimate, and the line indicates the uncertainty assigned to measurement data.

esis and ecosystem respiration, improves the estimates of NEE (compare top panel of Fig. 2 with Fig. 6). The confidence interval is approximately halved with the assimilation of each additional flux time series. With model alone the uncertainty is $196\,g\,C\,m^{-2}$; the uncertainty drops to $109\,g\,C\,m^{-2}$ once GPP data are assimilated, to $56\,g\,C\,m^{-2}$ with the assimilation of respiration data, and to $29\,g\,C\,m^{-2}$ with the assimilation of NEE data. DA improves on the estimates generated by observations alone. The uncertainty on daily measurements of NEE was estimated at $0.5\,g\,C\,m^{-2}\,day^{-1}$, but the uncertainty on assimilated estimates average $0.47\,g\,C\,m^{-2}\,day^{-1}$, and vary depending on whether NEE data were available for assimilation (Fig. 7). Higher uncertainties occurred on days when neither NEE nor R_{tot} observations were available for assimilation.

The assimilation process improves the quality of the estimates of C dynamics, but the calculated uncertainty needs to take account of the full range of uncertainties in the model, which can be challenging to quantify. Our approach was to use an ensemble of analyses to explore how uncertainty in the parameters and initial conditions of the model affected the quality of the analysis. We found that, if all parameters and initial conditions were sampled from a normal distribution around their nominal values, with a variance set to 20% of the nominal values, the mean of the NEE analyses over 3 years for unbiased models $(-421 \pm 17\,g\,C\,m^{-2})$ was little different from the nominal analysis $(-419 \pm 29\,g\,C\,m^{-2})$. Increasing the variance to 50% broadened the confidence interval to $\pm 60\,g\,C\,m^{-2}$, but this level of uncertainty in all the model parameters and initial conditions is unlikely. Overall, sensitivity analysis indicated that the uncertainty in model error

estimates, and in model parameter estimates, caused minor modifications to the analysis. This low sensitivity to model errors is important because these errors are generally poorly defined.

For simplicity, we used a daily model and assimilated daily data. However, generation of daily data was problematical, because flux data are rarely available in continuous time series. For NEE data, some hourly gap-filling was generally required, and this process necessarily introduced further uncertainty into the estimate, and potentially increased any inherent bias.

Ecological insights

The estimates of NEE from our analysis (a net sink) differ from those generated for 2001 using more conventional approaches for combining flux data (Law et al., 2003), which suggested that the forest was a small net source of C. NPP estimates from this and previous studies are similar, but the analysis of heterotrophic respiration for 2001 ($184\,C\,m^{-2}\,yr^{-1}$) is less than that derived from a more direct estimate based on the same automated chamber measurements of soil respiration ($505\,C\,m^{-2}\,yr^{-1}$; Irvine & Law, 2002) and heterotrophic fractions of soil respiration plus woody detritus decomposition (detritus, $67\,C\,m^{-2}\,yr^{-1}$; total, $301\,C\,m^{-2}\,yr^{-1}$). This difference is large enough to determine whether the forest is a source or sink.

The strength of the KF approach is that it combines a mass balance model with all available data, and thus provides a more precise assessment of C exchange than pure budget approaches. However, the EnKF relies on the model being an unbiased representation of the system, and on the data being unbiased. For example, the analysis may be degraded by bias in the NEE data, because of unreliable night-time measurements. One way to remove the night-time bias of eddy flux data would be to assimilate hourly data from well-mixed periods only. This approach would also avoid the problem of using gap-filling to generate daily NEE, but would require a model complex model. DA can also identify bias, by using the model and other data sources to reveal inconsistencies in the analysis. For instance, a bias in NEE data would necessarily alter the GPP and respiration analyses. Through the interconnections in the model, the NEE bias would be transmitted into the analyses of GPP and respiration. The transmission of bias will thereby increase the differences between the analyses of GPP and respiration and the unbiased observations of these data. The analysis presented here indicates a high degree of consistency among all flux data (see R^2 estimates in Table 4), which argues against a significant bias. Systematic bias through all observations will cause greater problems, as this may conceal any inconsistencies. In any DA exercise a careful investigation for evidence of bias is vital.

The model representation of litterfall and turnover is oversimplified, and the scarcity of data relating to CWD/SOM pools and their turnover means that the analysis of these processes is problematic. The difference between the EnFK estimate of R_h and that provided by mass balance suggests that process understanding and measurement capability remain weak in this area. Time series data describing the dynamics of the soil C and WD pools must be a critical part of any C study, so that this area of uncertainty is minimized. The broad confidence limits on the modelling of litter ensured that the observations determined the values of litter fluxes in the analysis. The current model formulation is valid for diagnosis, but, until an improved litter model is incorporated, it is less suitable for prognosis.

The parameter-fitting exercise suggests an NPP/GPP ratio of 53%, slightly larger than the range suggested by a survey of multiple forest types (Waring et al., 1998), where NPP/GPP $= 0.47 \pm 0.04$. The parameterized estimates of leaf life span are reasonable (1.0 years), as the site is a mix of pine, ferns, and deciduous shrubs. The model estimate of fine root life span (1.1 years) is close to that suggested by mini-rhizotron data, 1.7 years (Law et al., 2001b). However, the turnover rates of woody stems (1330 years) and SOM/WD (1033 years) are unrealistically large, as a result of the relatively long time constants on the dynamics of these pools, compared with the length of the observational period in this study, and also because of the aggrading nature of the forest stand.

A comparison of NEE estimated by fluxes against the overall change in C stocks revealed that the assimilation process can break the strict mass balance of the model forecast. The analysis suggested that C stocks increased by $1990 \pm 436\,g\,C\,m^{-2}$, a larger, more uncertain sink than suggested by summing daily NEE ($419 \pm 19\,g\,C\,m^{-2}$). The poor match arose because of a gradual, but significant accumulation of C in the SOM pool, which is poorly constrained and acts as a noise sink. We found that increasing the model confidence in the SOM prediction improved the mass balance, and reduced the accumulation of C in the SOM pool. While the mismatch identified here would be problematical over very long model runs, in this analysis the SOM pool has such a slow turnover that these mass changes have no noticeable impact on the predictions of fluxes. For prognostic applications, an improved model parametrization is required, and decadal data on woody biomass, from inventories and tree ring studies, should be assimilated to improve predictions of wood turnover. Field manipulation studies on soil processes will

provide insights and constraints into their rates of change.

Operating the model without assimilation, so that predictions are uncorrected throughout the 3 years (Fig. 6a), led to a reduced estimate of the strength of the C sink. The model-only simulations revealed a bias in NEE forecasts compared with observations. The bias largely resulted from the constant, low allocation to foliage that leads to a low GPP over the summer months. The cumulative effect becomes significant. When the model is linked to data via assimilation, the GPP and LAI observations ensure a boost of allocation to foliage in spring, increased GPP, and a stronger summer sink. To improve the prognostic capability of the simple model, a detailed leaf phenological model with temporal variation in C allocation is required. A more effective phenological model should ensure a better forecast of maximum LAI in 2001 (Fig. 4).

The future of DA in ecosystem studies

Our study has demonstrated why DA provides a powerful tool for analysing ecosystem processes, such as C cycling. Data alone are often insufficient for this task, or at least problematical, because of gaps in time series, or methodological uncertainties. Gap-filling techniques are generally highly statistical, and fail to take account of related data sets or process-based models. A conventional modelling approach to ecosystem analysis can also be problematical. Models are often tuned to fit a single data set, and that means the model provides little extra information. And if models have large numbers of parameters, there may be several different combinations of these parameters that can be tuned to match the data. Finally, many models fail to provide meaningful confidence limits around predictions.

Model DA, as we have demonstrated here with the EnKF, solves many of these problems. Assigning uncertainties to all observations determines the relative importance of data in forming the analysis. All predictions have associated confidence limits. The EnKF uses a model to generate complete time series, so gaps are eliminated. But the gap-filling makes use of all other available data at that time, and a modelled estimate embodying all prior information from the data set that is to be filled. Any increase in uncertainty for gap-filled estimates is explicitly calculated. DA makes use of all available data sets, so none are wasted in model parametrization. In our application here, we have estimated the key components of the C cycle from a combination of respiration data from chambers, NEE data from an eddy flux tower, C mass changes in vegetation and soil pools, a detailed model of photo-synthesis calibrated from sap flow data, and a mass balance model of C dynamics. By attaching error estimates to all model components and data sets, the KF produces a statistically optimal estimate of NEE (and, of course, each state variable in the model). The process is an efficient and ordered way of dealing with so much data.

Arguably, the subjective component of the assimilation process is in choosing the predictive model. By keeping this model simple, and by generating the model parameters directly from the available data sets, with their associated uncertainties, we keep the model construction process transparent and reproducible. The DA process and the statistical analysis of innovations offer the potential to compare alternative models, which would improve objectivity. For example, it would be intriguing to test a model that predicted autotrophic respiration based on the kinetics of metabolic reaction, rather than indirectly through GPP. Examining the analysis to locate when and where major corrections occur provides information on model reliability. The model we have used here clearly oversimplifies seasonal dynamics in litter turnover, and foliar and root phenology (Figs 4 and 5), and improvements to the model in these areas would be worth investigating. Using more detailed models in place of the simple model used here should provide enhanced analyses. The DA framework provides a means of quantifying any such improvement.

DA offers another important service, the determination of minimum data sets required to achieve specified confidence levels. The assimilation procedure can quantify how changes in the frequency of eddy flux data, chamber data, or measures of C pools affect the confidence limits on NEE predictions, for example. The very dense data collected at the young ponderosa pine site are not easily or rapidly reproducible or extensible. But it is not currently clear to what degree measurements could be reduced while still providing a reasonable estimate of NEE. DA offers the capability to make this determination, and thus play an important role in designing efficient regional ecological monitoring systems. The results of this initial analysis emphasize the importance of time series as constraints. Occasional, rare measurements of stocks, such as SOM or CWD pools, have limited use in constraining the estimates of other components of the C cycle. Long time series are particularly crucial for improving the analysis of pools with long time constants, such as SOM, woody biomass, and WD. Long-running forest stem surveys, and tree ring data, offer a rich resource that could be assimilated to provide an important constraint on C cycling of slow pools. Detailed time series of below-ground processes are more challenging to obtain. But

both on-line mass spectrometry, to determine continuously C isotopes of soil CO_2 effluxes, and tunable diode laser technology, for monitoring whole ecosystem isotopic signatures, may provide powerful sets of dynamic data that can be assimilated into models. And for extending estimates of NEE across regions, DA can play a further important role, by assimilating remote-sensing data into the analysis of C cycles. We have shown, via sensitivity analysis, how assimilating an estimate of photosynthesis – which might be provided indirectly by remotely sensed data – improves the analysis of NEE. Long-term remote-sensing data sets (from the past 30 years) could provide estimates of disturbance that could also be assimilated into a model.

Acknowledgements

References

Allen JI, Eknes M, Evensen G (2002) An Ensemble Kalman Filter with a complex marine ecosystem model: hindcasting phytoplankton in the Cretan Sea. *Annales Geophysicae*, **20**, 1–13.

Anthoni PM, Law BE, Unsworth MH (1999) Carbon and water vapor exchange of an open-canopied ponderosa pine ecosystem. *Agricultural and Forest Meteorology*, **95**, 115–168.

Anthoni PM, Unsworth MH, Law BE *et al.* (2002) Seasonal differences in carbon and water vapor exchange in young and old-growth ponderosa pine ecosystems. *Agricultural and Forest Meteorology*, **111**, 203–222.

Baldocchi DD (2003) Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology*, **9**, 479–492.

Burgers G, van Leeuwen PJ, Evensen G *et al.* (1998) On the analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, **126**, 1719–1724.

Cosby BJ (1984) Dissolved oxygen dynamics of a stream: model discrimination and estimation of parameter variability using an extended Kalman filter. *Water Science and Technology*, **16**, 561–569.

Cosby BJ, Hornberger GM, Kelly MG *et al.* (1984) Identification of photosynthesis – light models for aquatic systems. II. Application to a macrophyte dominated stream. *Ecological Modelling*, **23**, 25–51.

Ehman JL, Schmid HP, Grimmond CSB *et al.* (2002) An initial inter-comparison of micrometeorological and ecological inventory estimates of carbon exchange in a mid-latitude deciduous forest. *Global Change Biology*, **8**, 575–589.

Eknes M, Evensen G (2002) An ensemble Kalman filter with a 1-D marine ecosystem model. *Journal of Marine Science*, **36**, 75–100.

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99**, 10143–10162.

Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367.

Farquhar GD, von Caemmerer S (1982) Modelling of photosynthetic response to the environment. In: *Physiological Plant Ecology II. Encyclopedia of plant physiology, New Series*, **Vol. 12B** (eds Lange OL, Nobel PS, Osmond CB, Ziegler H), pp. 549–587. Springer-Verlag, Berlin.

Finnigan JJ, Clement R, Malhi Y *et al.* (2003) A re-evaluation of long-term flux measurement techniques – part I: averaging and coordinate rotation. *Boundary Layer Meteorology*, **107**, 1–48.

Goulden ML, Munger JW, Fan S-M *et al.* (1996) Measurements of carbon storage by long-term eddy correlation: methods and a critical evaluation of accuracy. *Global change biology*, **2**, 169–182.

Grewal MS (1993) *Kalman Filtering: Theory and Practice*. Prentice-Hall, Englewood Cliffs, NJ.

Harmon ME, Sexton J (1996) *Guidelines for measurement of woody detritus in forest ecosystems*. 20. US LTER Network Office, University of Washington, Seattle.

Heuvelink GBM, Webster R (2001) Modelling soil variation: past, present, and future. *Geoderma*, **100**, 269–301.

Irvine J, Law BE (2002) Contrasting soil respiration in young and old-growth ponderosa pine forests. *Global Change Biology*, **8**, 1183–1194.

Jarvis PG, Miranda HS, Muetzenfeldt RI (1985) Modelling canopy exchanges of water vapour and carbon dioxide in coniferous forest plantations. In: *The Forest–Atmosphere Interaction* (eds Hutchison BA, Hicks BB), pp. 521–542. D. Reidel, Dordrecht.

Kalman RE (1960) A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, **82**, 35–45.

Law BE, Baldocchi DD, Anthoni PM *et al.* (1999a) Below-canopy and soil CO2 fluxes in a ponderosa pine forest. *Agricultural and Forest Meteorology*, **94**, 171–188.

Law BE, Kelliher F, Baldocchi DD *et al.* (2001a) Spatial and temporal variation in respiration in a young ponderosa pine forest during summer drought. *Agricultural and Forest Meteorology*, **110**, 27–43.

Law BE, Ryan MG, Anthoni PM *et al.* (1999b) Seasonal and annual respiration of a ponderosa pine ecosystem. *Global Change Biology*, **5**, 169–182.

Law BE, Sun OJ, Campbell J *et al.* (2003) Changes in carbon storage and fluxes in a chronosequence of ponderosa pine. *Global Change Biology*, **9**, 510–524.

Law BE, Thornton P, Irvine J *et al.* (2001b) Carbon storage and fluxes in ponderosa pine forests at different developmental stages. *Global Change Biology*, **7**, 755–777.

Lorenc A (1986) Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112**, 1177–1194.

Lorenc AC (1995) *Atmospheric data assimilation*. Scientific Paper No. 34, Meteorological Office, Bracknell.

Malhi Y, Phillips OL, Lloyd J *et al.* (2002) An international network to monitor the structure, composition and dynamics of Amazonian forests (RAINFOR). *Journal of Vegetation Science*, **13**, 439–450.

Maybeck PS (1979) *Stochastic Models, Estimation and Control*, **Vol. 1**. Academic Press, New York.

McGuire AD, Joyce LA, Kicklighter DW *et al.* (1993) Productivity response of climax temperate forests to elevated temperature and carbon dioxide: a North American comparison between two global models. *Climatic Change*, **24**, 287–310.

Parton WJ, Stewart JWB, Cole CV *et al.* (1988) Dynamics of C, N, P and S in grassland soils: a model. *Biogeochemistry*, **5**, 109–131.

Phillips OL, Malhi Y, Higuchi N *et al.* (1998) Changes in carbon balance of tropical forests: evidence from long-term plots. *Science*, **282**, 439–442.

Rastetter EB, Aber JD, Peters DPC *et al.* (2003) Using mechanistic models to scale ecological processes across space and time. *Bioscience*, **53**, 68–76.

Running SW (1994) Testing forest – BGC ecosystem process simulations across a climatic gradient in Oregon. *Ecological Applications*, **4**, 238–247.

Running SW, Baldocchi DD, Turner DP *et al.* (1999) A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modelling and EOS satellite data. *Remote Sensing of Environment*, **70**, 108–127.

Schmid HP, Lloyd CR (1999) Spatial representatives and the location bias of flux footprints over inhomogeneous areas. *Agricultural and Forest Meterology*, **93**, 195–209.

Schwarz PA, Law BE, Williams M *et al.* (2004) Climatic vs. biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems. *Global Biogeochemical Cycles*, **18**, GB4006, doi: 10.1029/2004GB002227.

Thornley JHM, Cannell MGR (2000) Modelling the components of plant respiration: representation and realism. *Annals of Botany*, **85**, 55–67.

Turner DP, Koerper GJ, Harmon ME *et al.* (1995) A carbon budget for the forests of the conterminous United States. *Ecological Applications*, **5**, 421–436.

Waring RH, Landsberg JJ, Williams M *et al.* (1998) Net primary production of forests: a constant fraction of gross primary production? *Tree Physiology*, **18**, 129–134.

White MA, Thornton PE, Running SW *et al.* (2000) Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls. *Earth Interactions*, **4**, 1–85.

Williams M, Bond BJ, Ryan MG *et al.* (2001a) Evaluating different soil and plant hydraulic constraints on tree function using a model and sap flow data from ponderosa pine. *Plant, Cell and Environment*, **24**, 679–690.

Williams M, Law BE, Anthoni PM *et al.* (2001b) Use of a simulation model and ecosystem flux data to examine carbon–water interactions in ponderosa pine. *Tree Physiology*, **21**, 287–298.

Williams M, Rastetter EB, Fernandes DN *et al.* (1996) Modelling the soil–plant–atmosphere continuum in a *Quercus-Acer* stand at Harvard Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. *Plant, Cell and Environment*, **19**, 911–927.

Williams M, Rastetter EB, Fernandes DN *et al.* (1997) Predicting gross primary productivity in terrestrial ecosystems. *Ecological Applications*, **7**, 882–894.

Williams M, Rastetter EB, Shaver GR *et al.* (2001c) Primary production in an arctic watershed: an uncertainty analysis. *Ecological applications*, **11**, 1800–1816.

Wofsy SC, Goulden ML, Munger JW *et al.* (1993) Net exchange of CO_2 in a mid-latitude forest. *Science*, **260**, 1314–1317.

Appendix

Carbon dynamic model equations. T is a temperature-sensitive rate parameter.

Fluxes:

For G, see ACM model below
$R_a = t_2 G$
$A_f = (1-t_2)\, t_3 G$
$A_w = (1-t_2)(1-t_3-t_4)G$
$A_r = (1-t_2)t_4 G$
$L_f = t_5 C_f$
$L_w = t_6 C_w$
$L_r = t_7 C_r$
$R_{h1} = t_8 C_{lit} T$
$R_{h2} = t_9 C_{SOM} T$
$D = t_1 C_{lit} T$

Pools:

$\Delta C_f = A_f - L_f$
$\Delta C_w = A_w - L_w$
$\Delta C_r = A_r - L_r$
$\Delta C_{lit} = L_f + L_w + L_r - R_{h1} - D$
$\Delta C_{SOM/WD} = D - R_{h2}$
$T = 0.5 \times \exp(0.0693A)$ (equivalent to a curve with $Q_{10} = 2$, $T = 1$ when $A = 10$)

where A is the mean daily air temperature (°C).

ACM equations to estimate G

$$g_c = \frac{|\psi_d|^{a_{10}}}{0.5 T_r + a_6 R_{tot}}, \tag{A1}$$

where ψ_d is the maximum soil–leaf water potential difference (MPa), T_r is the daily temperature range (°C), and R_{tot} is the total plant–soil hydraulic resistance (MPa m^2 s mmol^{-1}).

$$p = \frac{a_1 N L}{g_c}\exp(T_{max}a_8), \tag{A2}$$

where N is the average foliar N (g N m^{-2} leaf area), L is the LAI, estimated by $L = C_f/111$, T_{max} is the maximum daily temperature

$$q = a_3 - a_4, \tag{A3}$$

$$C_i = \tfrac{1}{2}\Big[C_a + q - p$$

$$+ \sqrt{(C_a + q - p)^2 - 4(C_a q - p a_3)}\,\Big], \qquad (A4)$$

$$e_0 = \frac{a_7 L^2}{L^2 + a_9}, \qquad (A5)$$

$$\delta = -0.408 \cos\left(\frac{360(D + 10)}{365}\,\frac{\pi}{180}\right), \qquad (A6)$$

where δ is the solar declination (radians) and D the day of year.

$$s = 24 \cos^{-1}(-\tan(\Delta)\tan(\delta))/\pi, \qquad (A7)$$

where s is the day length in hours. If $\tan(\Delta)\tan(\delta) \geq 1$ then $s = 24$, Δ the latitude (radians)

$$G = \frac{e_0 I g_c (C_a - C_i)}{e_0 I + g_c (C_a - C_i)} (a_2 s + a_5), \qquad (A8)$$

where G is the GPP in $g\,C\,m^{-2}\,day^{-1}$, I the irradiance ($MJ\,m^{-2}\,day^{-1}$).

Parameters valid for Ponderosa pine GPP predictions in ACM model (calibrated from SPA model predictions at the young site, Schwarz *et al.*, 2004)

Parameter	Value
a_1	2.155
a_2	0.0142
a_3	217.9
a_4	0.980
a_5	0.155
a_6	2.653
a_7	4.309
a_8	0.060
a_9	1.062
a_{10}	0.0006