Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 112 (2008) 1347–1364
Remote Sensing of Environment
www.elsevier.com/locate/rse

Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter

Tristan Quaife [a], Philip Lewis [a,*], Martin De Kauwe [a], Mathew Williams [b], Beverly E. Law [c], Mathias Disney [a], Paul Bowyer [a]

[a] NERC Centre for Terrestrial Carbon Dynamics and Department of Geography, University College London, Gower St., London WC1H 0AP, UK
[b] NERC Centre for Terrestrial Carbon Dynamics and School of GeoSciences, Crew Building, University of Edinburgh, Edinburgh EH9 3JN, UK
[c] College of Forestry, Oregon State University, Corvallis, OR 97331, USA

Received 31 July 2006; received in revised form 9 February 2007; accepted 5 May 2007

Abstract

An Ensemble Kalman Filter (EnKF) is used to assimilate canopy reflectance data into an ecosystem model. We demonstrate the use of an augmented state vector approach to enable a canopy reflectance model to be used as a non-linear observation operator. A key feature of data assimilation (DA) schemes, such as the EnKF, is that they incorporate information on uncertainty in both the model and the observations to provide a best estimate of the true state of a system. In addition, estimates of uncertainty in the model outputs (given the observed data) are calculated, which is crucial in assessing the utility of model predictions.

Results are compared against eddy-covariance observations of CO_2 fluxes collected over three years at a pine forest site. The assimilation of 500 m spatial resolution MODIS reflectance data significantly improves estimates of Gross Primary Production (GPP) and Net Ecosystem Productivity (NEP) from the model, with clear reduction in the resulting uncertainty of estimated fluxes. However, foliar biomass tends to be over-estimated compared with measurements. Issues regarding this over-estimate, as well as the various assumptions underlying the assimilation of reflectance data are discussed.

Keywords: Data assimilation; MODIS; Ecosystem modelling; Terrestrial Carbon Dynamics

1. Introduction

Quantifying the magnitude and dynamics of carbon (C) sources and sinks across regions is of major importance to Earth System Science because of the relevance to the dynamics of atmospheric carbon dioxide (CO_2) (Schimel et al., 2001) and hence climate, and national and international carbon accounting policies (IPCC, 2001). A major uncertainty in the analysis of terrestrial C dynamics exists over whether, and under what circumstances, particular regions or landscapes act as sources or sinks. Advances in this area have been made through the use of process-based ecosystem models (Law et al., 2001a; Rastetter et al., 2003) but large variations remain between different models (Churkina et al., 2005) and their parameterisation over large areas is often uncertain. Additional knowledge has been gained through networks of C flux tower measurements over a range of ecosystems (Valentini et al., 2000). But whilst these aid understanding in conjunction with models, the results are difficult to directly scale up to provide regional or global estimates (Running et al., 1999).

This paper develops from the work of Williams et al. (2005), who argue that the Bayesian assimilation of measurements into a terrestrial ecosystem model can improve estimates of ecosystem C stocks and fluxes over model runs alone. Importantly, uncertainty in the predicted stocks and fluxes is reduced compared to the original measurements. Williams et al. (2005) use a simple box representation (the Data Assimilation Linked Ecosystem Carbon model, DALEC, outlined below) that represents the major pools and fluxes of C in an ecosystem,

* Corresponding author.
E-mail address: plewis@geog.ucl.ac.uk (P. Lewis).

0034-4257/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.05.020

and assimilate a time series of flux and occasional stock measurements into the model using the Ensemble Kalman Filter (EnKF) (Evensen, 2003). This variant of the Kalman Filter (KF) tracks the model error statistics using an ensemble of model state vectors and has the desirable advantage of being applicable to non-linear models. The ecosystem model parameters (initial carbon pool sizes and rate parameters) are estimated through optimisation of a merit function against the observations available. This model can then be run to provide estimates of ecosystem dynamics and C fluxes when driven by climate data. Williams et al. (2005) demonstrate that with measurements over a young ponderosa pine forest in central Oregon, USA over a 3-year period, assimilation provides improved estimates of net ecosystem exchange of C (NEE) compared to model runs alone (-419 ± 29 g C m^{-2} compared to -251 ± 197 g C m^{-2}). In addition, assimilation can aid the identification of model deficiencies through examining model-data biases (Cosby, 1984; Rastetter, 2003). Such an approach can be relatively easily applied to other 'data rich' ecosystems and regions. It can potentially be extended in spatial extent by applying the calibrated model to regions with similar ecosystems to those under which the model optimisation was performed. The fact that such a large reduction in uncertainty can potentially be achieved through the assimilation of time series observations implies that to estimate regional to global carbon stocks and fluxes accurately, such data would ideally be utilised both spatially and temporally. However, measurements of C fluxes over ecosystems are only available at sparse tower sites and obviously cannot directly fulfil this role. The use of atmospheric measurements from so-called 'tall towers' (Davis et al., 2003) and air- and spaceborne sensors (Gerbig et al., 2003; Raupach et al., 2005) could potentially fill this data gap on C flux measurements in the future, although this will require linking the ecosystem models to models of atmospheric transport to allow assimilation into 'spatialised' ecosystem data assimilation (DA) schemes. Whilst direct measurements of C fluxes clearly provide strong constraints to modelling C fluxes, there are a range of other measurements available relating to ecosystem processes that can also be considered as candidates for assimilation and reducing uncertainty in model estimates. Probably the only source of such data available as time series over wide geographical extents are those obtained from Earth Observation (EO) satellites.

The synoptic coverage of moderate to coarse spatial resolution EO sensors makes them ideal candidates for generation of data suitable for integration into ecosystem models. Typical moderate spatial resolution (250 m–1 km) sensors designed to study the biosphere, such as the MODIS sensors on NASA's Terra and Aqua satellites (Heinsch et al., 2006), or the VEGETATION sensor on board SPOT-4 and SPOT-5 (Maisongrande et al., 2004), have near daily coverage of the whole globe. Measurements in the optical spectrum are responsive to vegetation state variables such as leaf area index (LAI) (Cohen et al., 2006) that can be linked to ecosystem model variables such as leaf foliar biomass through simple linear operators (multiplication by specific leaf area in this case). In addition, such observations can be used to track vegetation dynamics to estimate vegetation phenology to calibrate temperature responses of plant functional types as well as test ecosystem model behaviour by comparison with satellite observations (McCloy & Lucht, 2004). An appreciation of the importance of such data in driving, parameterising and testing ecosystem models has resulted in much effort in the past decade being put into the generation of suites of global high-level products from EO measurements (Justice et al., 2002). Among those most directly relevant to C modelling in terrestrial ecosystems derived from optical sensors are: land cover and land cover change; LAI; the fraction of absorbed photosynthetically-active radiation (fAPAR); snow cover; tree and shrub cover and biomass (Justice et al., 2002; Law & Waring, 1994). Such data are derived via a range of methods ranging from pattern recognition (classification of land cover) through assumed or calibrated empirical relationships with satellite-derived spectral indices (snow cover, some fAPAR products), to the inversion of physically-based canopy reflectance models (LAI, fAPAR) (Knyazikhin et al., 1998). Further derived products exist, such as estimates of daily Gross Primary Production (GPP) and annual Net Primary Production (NPP) (Running et al., 2000), which are generated via production efficiency models, driven by satellite-derived estimates of fAPAR. Other relevant forms of information, particularly related to canopy structure, such as above-ground biomass or tree height are starting to become available from other sensor technologies such as Synthetic Aperture RADAR (SAR) and LiDAR (Treuhaft et al., 2004). Such products are of great importance in mapping the dynamics of terrestrial vegetation over recent decades and also in reducing the vast amount of data available from satellites to more easily manageable quantities of information. In addition, many of these products either directly produce information on components of the terrestrial C cycle (GPP, NPP) or give information that could be assimilated into terrestrial ecosystem models within the framework described by Williams et al. (2005). We can however identify a number of features of such datasets that can practically limit their utility in such an application:

1. Despite many efforts at validation of these high-level, derived products it is generally difficult to accurately characterise their error and uncertainty (Heinsch et al., 2006). Characterisation of bias and uncertainty are key to effective DA; without this information assimilation will at best achieve sub-optimal results;
2. A range of estimates of what are nominally the same product (e.g. fAPAR from MODIS, MERIS, AVHRR (Gobron et al., 2002; Knyazikhin et al., 1998; Tian et al., 2000)) are provided from different sensors processed by different methods and relying on different assumptions. Whilst this may be of value in obtaining ensemble statistical characterisations of such quantities, the generation of various (somewhat different) products from individual sensors can make it difficult to exploit such information in a DA scheme i.e. how should such products be used/combined;
3. Products are generated using specific models based on particular (implicit or explicit) sets of assumptions and often

particular sets of driving (e.g. climate) data, which may in turn be inconsistent with the assumptions made in the ecosystem model being used for assimilation. GPP, for example, is typically derived from EO data using a production efficiency model driven by estimates of fAPAR and climate data (Potter et al., 1993). Ecosystem models often use more sophisticated schemes to model GPP, which may result in inconsistencies between observed and modelled GPP values over a given site. This can be further exacerbated through use of different driving climate datasets;

4. For efficiency and for other reasons such as the need for cloud-free observations, high-level EO products are generally derived from a set of observations composited over a limited time period (e.g. 8 days for the MODIS LAI/fAPAR product). Most products are derived independently for each temporal window over which samples are taken but this can result in artificially large high frequency variations in the products due to the temporal compositing. Although such variations can be smoothed out (e.g. Plummer et al., 2006), the degree of smoothing needs to be imposed, and the products take no direct advantage of expectations of relatively low temporal frequency variations which would be expected for the parameters under observation (e.g. LAI). A further disadvantage of this approach is that if insufficient cloud-free observations are available during a given compositing period for a full model-based retrieval to be made, typically either some form of empirical 'back up' algorithm is used, or no parameter value is generated at all.

Given the issues surrounding the high-level EO-derived ecosystem products, it is in many ways advantageous to directly assimilate satellite observations of surface-leaving spectral radiance, reflectance or albedo. This has the specific advantage, from a DA perspective, of allowing uncertainties in the observations to be tracked, as they can be far more easily characterised for satellite radiance than for higher level products. In the case of optical EO, the satellite measures top-of-atmosphere (TOA) spectral radiance. A desirable strategy, therefore, would be to assimilate these data directly into the ecosystem model. However, these measurements are affected by the atmospheric composition, to which the ecosystem model is only very slightly sensitive (e.g. the ratio of direct to diffuse global illumination, water vapour and CO_2 in the planetary boundary layer). It may be feasible to attempt to assimilate TOA radiance measurements into an ecosystem model coupled with an atmospheric circulation model, but that presents many technical challenges and is beyond the scope of this paper. A compromise then, is to assimilate relatively low-level EO-derived observations of an intrinsic surface property. The most relevant observation here being at-ground spectral bidirectional reflectance factor (BRF), i.e. satellite radiance corrected for atmospheric scattering and absorption (Vermote et al., 2002) and normalised by the incident radiation field.

A typical ecosystem model does not directly provide estimates of canopy spectral reflectance. Indeed, the ecosystem model will not generally be greatly sensitive to all of the factors that affect canopy BRF (e.g. arrangement and distribution of leaves). In order to be able to assimilate such data then, an 'observation operator' is required that relates one or more ecosystem state variables to canopy BRF. This is achieved here with a physically-based model of canopy BRF. A key requirement of the EnKF is that the observations must be predicted from the ecosystem model state vector by means of a linear observation operator. Canopy BRF is not, however, a linear function of the state vector of a typical ecosystem model and hence the EnKF scheme must be extended to deal with this non-linearity.

Several studies have presented data assimilation schemes that use EO data to improve the functioning of ecosystem or plant growth models. Knorr and Lakshmi (2001) assimilate fAPAR, derived from AVHRR, and skin surface temperature derived from TOVS into the Biosphere Energy-Transfer Hydrology (BETHY) model. Assimilating the fAPAR and temperature data separately the study demonstrated the improvement in the modelled estimates of the un-assimilated variable. The fAPAR was used to determine the model parameters by minimising against the whole data time series using a merit function weighted by observational uncertainty and the temperature data was assimilated using a nudging technique. The fAPAR was predicted from the BETHY state vector using a semi-discrete canopy BRF model (Gobron et al., 1997; Knorr, 1998). This BRF model does not take into account various factors, such as multi-scale clumping, that are important for predicting BRFs in areas where the vegetation does not provide continuous cover. Furthermore the minimisation scheme does not account for uncertainty in the predicted fAPAR and so BRF model inadequacies are not carried through uncertainties in the model outputs. The fAPAR assimilation described by Knorr and Lakshmi (2001) forms part of the Carbon Cycle Data Assimilation System (CCDAS; Rayner et al., 2005). The CCDAS also assimilates observations of atmospheric CO_2.

Nouvellon et al. (2001) couple the Markov chain canopy reflectance model of Kuusk (1995) to a grassland ecosystem model to predict the Normalised Difference Vegetation Index (NDVI). NDVI data were assimilated from a time series of TM and ETM+ images by minimising the mean squared difference between the predictions and observations. Neither uncertainties in the model nor in the observations were taken into account, hence the assimilation scheme is sub-optimal and the reduction in the error of the model predictions could not be quantified. The authors recommend the use of the Kalman Filter to address this issue.

Vivoy et al. (2001) have used the Kalman Filter to assimilate NDVI data into the ORCHIDEE (Organizing Carbon and Hydrology In Dynamic Ecosystems Environment) model. The SAIL (Scattering from Arbitrarily Inclined Leaves) model of Verhoef (1984) was used to predict the NDVI from the ORCHIDEE state vector and assimilation was performed using the Extended Kalman Filter (EKF). The authors overcome the restriction that the observation operator must be linear in the EKF assimilation scheme by using a Jacobian matrix to represent the SAIL model at the required point in parameter space. A problem with this approach is that if uncertainty in the

parameters of the SAIL model is large then the Jacobian is not an adequate description of the SAIL model. The authors address this problem by running the observation operator in an ensemble mode, which they concede will only partially overcome the inadequacies of using the Jacobian (Vivoy et al., 2001). In addition the error covariance matrix used in the assimilation scheme is diagonal and so cross correlations between the errors of the different parameters cannot develop. Both of these issues may be addressed using the Ensemble Kalman Filter.

We demonstrate the feasibility of the assimilation of canopy BRF data into the simple ecosystem model, DALEC, described by Williams et al. (2005), using a non-linear observation operator and an augmented state vector. The DALEC model assimilation scheme has been shown to work well in reducing uncertainty in C flux estimates under situations where large quantities of field (C flux and stock) data are available for assimilation. DALEC operation is examined here for a test region in Oregon, USA, using the calibrated DALEC parameterisation of Williams et al. (2005), and the assimilation of 500 m resolution Terra MODIS surface BRF data at red and near infrared (NIR) wavelengths over a 3-year period. Predicted carbon fluxes are compared to those measured in the field during the same time period. Assumptions underlying the assimilation scheme are discussed and reductions in uncertainty in the predictions of C fluxes are demonstrated.

2. Method

2.1. The Ensemble Kalman Filter

The EnKF was designed to facilitate data assimilation into non-linear models within a Kalman gain scheme. Its formulation is largely due to Evensen (1994, 2003). It has the same basic form as the KF but instead of propagating a model error covariance matrix the EnKF uses an ensemble of model states to represent error statistics in the model:

$$\mathbf{A}^{a} = \mathbf{A} + \mathbf{A}'\mathbf{A}'^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{A}'\mathbf{A}'^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_{e})^{-1}(\mathbf{D} - \mathbf{H}\mathbf{A}) \quad (1)$$

where, $\mathbf{A}$ is a matrix containing the ensemble of model state vectors; $\mathbf{A}'$ is the matrix of ensemble perturbations (the ensemble minus its mean); $\mathbf{R}_e$ is the error covariance matrix of observation errors; $\mathbf{D}$ is the observation ensemble and $\mathbf{H}$ is the observation operator. The superscript 'a' denotes the analysed state vector ensemble (i.e. the result of applying the EnKF to the model state and available observations) and the superscript 'T' denotes a matrix transpose. As the size of the ensemble reaches infinity the EnKF is mathematically identical to the Kalman Filter.

Unlike the Extended Kalman Filter (EKF), which is also designed for application to non-linear models, the EnKF does not require a tangent linear model to propagate error covariances. These are represented within the ensemble itself and hence propagated during the course of the normal forward run of the model. As construction of a tangent linear model can be prohibitively time consuming and complex the EnKF offers an attractive alternative.

The operation $\mathbf{D}-\mathbf{H}\mathbf{A}$ imposes the condition that the observation operator $\mathbf{H}$ itself is linear. In this form it is not possible to couple a canopy reflectance model to an ecosystem model as top of canopy BRF is not a linear transform of a typical ecosystem model state vector. This may be overcome by forming an augmented state vector (Evensen, 2003). Eq. (1) then becomes:

$$\mathbf{A}^{a} = \mathbf{A} + \mathbf{A}'\hat{\mathbf{A}}'^{\mathrm{T}}\hat{\mathbf{H}}^{\mathrm{T}}(\hat{\mathbf{H}}\hat{\mathbf{A}}'\hat{\mathbf{A}}'^{\mathrm{T}}\hat{\mathbf{H}}^{\mathrm{T}} + \mathbf{R}_{e})^{-1}(\mathbf{D} - \hat{\mathbf{H}}\hat{\mathbf{A}}) \quad (2)$$

Where $\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}'$ are formed by augmenting the state vector ensemble with predictions of reflectance and $\hat{\mathbf{H}}$ is a new linear observation operator that transforms between the augmented state vector and the observation. In effect the predictions of reflectance become part of the model at observation times.

The augmented state vector ensemble is formed by:

$$\hat{\mathbf{A}} = \mathbf{h}(\mathbf{A}) \quad (3)$$

where $\mathbf{h}$ is the non-linear observation operator. In the work presented here $\mathbf{h}$ is a canopy reflectance model that is used to predict the BRF derived from TOA radiances measured at-sensor.

Having a BRF model as an observation operator necessitates the assignment of values to all parameters of the model that are not exploited directly in the assimilation scheme ('ancillary parameters'). In this paper, the only parameter that is exploited directly is LAI. Estimates of uncertainty in the BRF model predictions are treated by using a variance–covariance matrix to map through random variations in the model ancillary parameters at each time step for each ensemble member. Although mean ancillary parameters are assumed constant over the time period (see below), their variation is tied together in this way, but is assumed independent of the ensemble LAI value.

2.2. Study area

The study area (Metolius) is a temperate coniferous forest site in Oregon, Northwestern US which has been studied extensively (Law et al., 2001a). The Metolius site is located within a section of privately owned land east of the Cascades Range, Oregon, USA (44° 26′ N, 121° 34′ W, elevation 1188 m) and its position relative to the coast and mountains induces a semi-arid climate (Williams et al., 2005). The site is characterised primarily by young coniferous pine forest, dominated by ponderosa pine (*Pinus ponderosa*), as well as a sparse understory of bitterbush (*Purshia tridentata*) and manzanita (*Arctostaphylos patula*) (Law et al., 2001a,b). The forest was formerly a mature forest which was felled in 1978 and allowed to regrow naturally, however small patches of older growth forest still persist (Williams et al., 2005). Fig. 1 shows the 4.5 km × 4.5 km region centred on the flux tower site, as a Normalised Difference Vegetation Index (NDVI) image from July 2002. There is a clear variation in tree density over the region, with generally higher densities towards the North of the flux tower site.

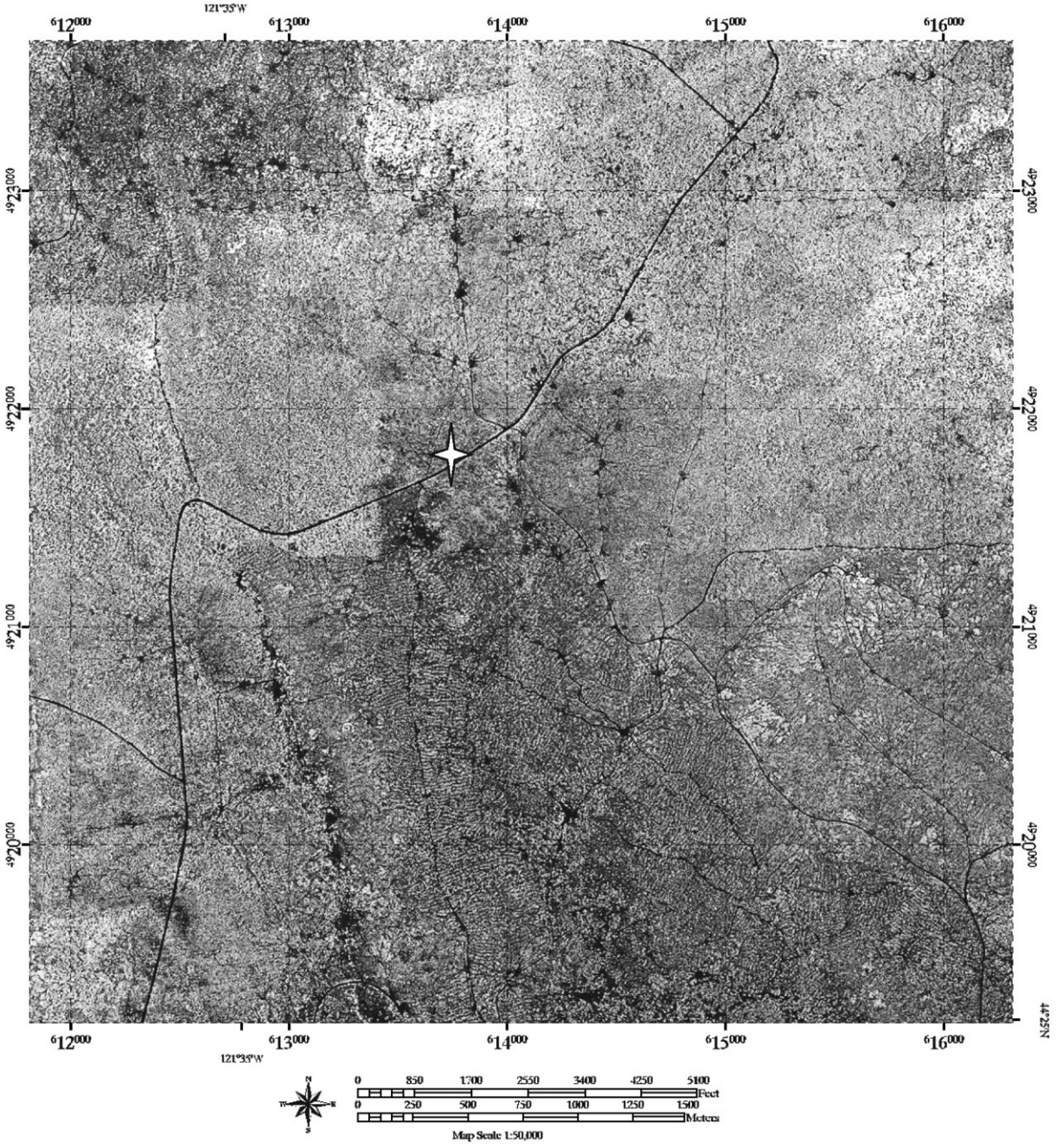


Fig. 1. IKONOS NDVI image of the 4.5 km × 4.5 km area around the flux tower site. The data are from July 2002.

The site is part of the AmeriFlux network of flux towers (http://public.ornl.gov/ameriflux/), with measurements recorded at the study site between 1999 and 2003, by an eddy-covariance system (Anthoni et al., 2002). The eddy-covariance system is mounted on top of a tower overlooking the canopy and samples a variable footprint dictated by the wind direction. Measurements of wind speed and gaseous concentrations from the eddy-covariance system allow the net exchange of carbon, water vapour and energy between the terrestrial land surface and the atmosphere to be quantified. Soil respiration estimates were derived from six automated chambers (Irvine & Law, 2002). Relevant meteorological data have also been collected at the site. These including incident PAR, vapour pressure deficit and temperature (diurnal range and mean), which provide the driving variables for the DALEC model.

Net ecosystem exchange of CO_2 (NEE) (the negative of Net Ecosystem Production, NEP) was computed from micrometeorological data using the eddy-covariance technique. The data were collected with a Campbell Scientific sonic anemometer and LI7500 open-path infrared gas analyzer (LiCor, Lincoln, NE, USA) at 12 m height for the 3 m tall canopy. Additional measurements included vertical profiles of CO_2 concentration

(LI6262), temperature, and wind speed, and soil temperature at 2 cm depth. A quality control process removes spikes from the fast-response data on wind and sonic temperature from the CSAT3, and humidity and CO_2 from the LI7500. A point-by-point WPL correction was applied to the LI7500 output. A tilt correction based on the long-term average wind direction dependence of the tilt angle was applied to the wind components (Mahrt et al., 2000). Fluxes were averaged hourly after applying an automated algorithm to identify the averaging time-scale within each 1-h record that captures turbulence while excluding poorly sampled meso-scale motions (using the variable averaging time method of Vickers & Mahrt, 1997). Net ecosystem exchange was calculated from the CO_2 flux and change in CO_2 storage in the canopy air space.

Periodic biological measurements were made to estimate NPP (foliage, live stemwood, root), above- and below-ground biomass (foliage, wood, litter, root), rooting depth, seasonal LAI, foliar nitrogen, and soil carbon (Schwarz et al., 2004). Ecosystem respiration was determined from measurements of soil, foliage and wood respiration and site-specific regressions to obtain hourly respiration following Law et al. (1999).

LAI was estimated in 1999, 2001, and 2002 using the optical approach described in Law et al. (2001b). At each site, light transmittance was measured at 39 pre-determined locations over a 100 m × 100 m area around the site using a LAI-2000 Plant Canopy Analyzer (Li-Cor, Inc., Lincoln, NE). Mean LAI was calculated and then corrected for wood interception and for clumping of shoots and needles within shoots according to the methods outlined in Law et al. (2001a). In 2002, LAI was measured in early spring, prior to the expansion of current year needles, and again in late summer during maximum seasonal leaf area to estimate the seasonal changes in leaf area. In 1999 and 2001, LAI was measured at the seasonal maximum. Examination of the measurements of maximum seasonal LAI suggested that leaf area at the site increased by about 17% per year over the four year period (Schwarz et al., 2004). The LAI data have been scaled up to foliar biomass using measurements of the grams of carbon per leaf area (GCLA).

Tree transpiration was estimated by measuring sap flux density continuously between April and November using a heat dissipation technique. Site-level estimates of transpiration were computed with tree transpiration, site-specific sapwood area relationships, tree size and density (Granier, 1987; Irvine et al., 2002). Photosynthetic parameters were determined from $A–C_i$ curves on 1-year old foliage (J_{max}, V_{cmax}) (Irvine et al., 2004).

2.3. Ecosystem model

The DALEC model (Williams et al., 2005) represents the ecosystem as a series of carbon stores linked together by fluxes, the magnitude of which are controlled by a set of rate parameters (Fig. 2). The carbon pools are C_f, C_r, C_w, C_{lit}, and $C_{som/cwd}$ which represent foliar, root, woody, litter and soil organic matter and coarse woody debris carbon respectively. In addition there is a pool that represents the daily GPP. Fluxes are denoted R_a and R_h (the autotrophic and heterotrophic respiration), D (the decomposition from litter to soil organic matter), L_f, L_r and L_w (the loss of foliar, fine root and woody carbon respectively) and A_f, A_r and A_w (the allocation of foliar, fine root and woody carbon from GPP). Microbially-driven fluxes (R_h and D) are temperature dependent.

GPP is modelled by the Aggregated Canopy Model (ACM) of Williams et al. (1997). ACM is driven by climate and canopy data including incoming photosynthetically active radiation, soil water potential (in this case determined from soil moisture measurements collected at the site), mean daily temperature and diurnal temperature range, and canopy nitrogen. ACM requires LAI to determine light interception within the canopy, and this is determined as a function of the total foliar biomass (C_f).

The Net Ecosystem Productivity (NEP) is the difference between the GPP and the combined autotrophic and heterotrophic respiration. NEP is the total carbon retained by an

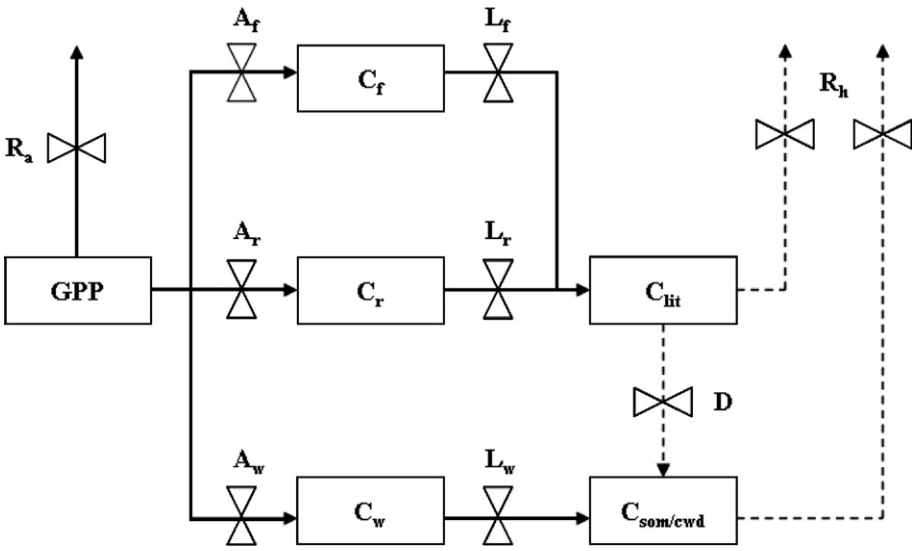


Fig. 2. A schematic diagram of the DALEC model. Boxes represent stocks and bow-ties represent fluxes. Dotted lines represent temperature dependent fluxes.

Table 1a
Parameters for Ponderosa pine GPP predictions in the ACM model (from Williams et al., 2005)

Parameter	Term	Value
a_1	Nitrogen use efficiency (NUE) parameter	2.155
a_2	Day length coefficient	0.0142
a_3	CO_2 compensation point	217.9
a_4	CO_2 half saturation point	0.980
a_5	Midsummer coefficient	0.155
a_6	Coefficient of hydraulic resistance	2.653
a_7	Maximum canopy quantum yield	4.309
a_8	Temperature coefficient of NUE	0.060
a_9	LAI-canopy quantum yield coefficient	1.062
a_{10}	Soil–leaf water potential exponent	0.0006

ecosystem (in vegetation and soil) after plant and microbial respiration. A negative value of NEP indicates a source of carbon.

Parameters for both DALEC and ACM for the Metolius site have been determined previously (Williams et al., 2005) and are used here. DALEC and ACM parameters from Williams et al. (2005) are given in Tables 1a and b.

2.4. Satellite data

The MODIS data employed in this study were the Level 2 gridded 500 m MODIS Terra daily surface reflectance (MOD09GHK, v004). MOD09 contains surface reflectance measured at 7 wavebands ranging from blue to middle infrared wavelengths, although only data from the red and near infrared parts of the spectrum (centred at 646 nm and 848 nm respectively) were used in this study. The data are derived from TOA measurements of spectral radiance, with atmospheric scattering and absorption effects removed (Vermote et al., 2002). The uncertainty in these data is assumed to be 0.004 and 0.015 for the red and NIR bands respectively (Roy et al., 2005).

One reason for using the MODIS surface reflectance is that the surface reflectance products of many other sensors treat components of atmospheric radiative transfer with less rigour. The SPOT-VEGETATION surface reflectance products (Maisongrande et al., 2004), for example, contain gross assumptions about aerosol distribution and hence errors in the derived BRFs may be prone to bias. MODIS data have undergone wide testing and validation (Vermote et al., 2002, 1997) and should provide unbiased estimates of the surface BRF with stable error characteristics. Biased observations will cause the model to be pulled away from its correct state and observations whose uncertainties are incorrectly assigned will not affect the model trajectory to the correct extent.

The geo-location accuracy of MODIS is around 50 m (1/10th of a pixel) at nadir (Wolfe et al., 2002). We do not explicitly model the uncertainty due to the geo-location and we do not expect it to have a major effect on the results of this study given the accuracy of the MODIS data.

MODIS data granules were obtained from the NASA Goddard Space Flight Centre (GSFC) Distributed Active Archive Centre (GDAAC). Daily data for the years 2000–2002 were extracted for the 9×9 pixels that surround the study site. Only reflectance data flagged indicated to be of high quality by the quality control (QC) data plane were utilised in this study. These are deemed to be free of clouds, snow and other issues. Reflectance values which passed the QC checks were extracted for the study area in the two wavebands along with the associated sun-sensor geometry information required for modelling BRF.

The data were further filtered for residual cloud and cloud shadow effects by detecting outliers after applying a Ross-Thick Li-Sparse linear kernel-driven model of BRF (Wanner et al., 1995) to the data in the red waveband. This approach tended to remove around 10% of the data which had passed the QC checks.

In this study, we make no use of MODIS observations that are thought to be contaminated with snow. This will clearly result in a lack of observations in the winter period, but the inclusion of such data would require the observation operator to be able to model the effect of snow on canopy BRF, which cannot be achieved with the current version of the model. However, the site experiences periodic rather than continuous snowfall through the winter, so the data gap is not as large as might be expected in other northern regions.

2.5. Canopy reflectance model

The canopy reflectance model used in this work is the hybrid Geometric Optic Radiative Transfer (GORT) model described by Ni et al. (1999) and is a modification of the model of Li et al. (1995). The model uses geometric optics to describe the proportions of shaded and sunlit crown and ground viewed at a given sun-sensor geometry and assumes that the tree crowns can be adequately represented as spheroids (Li & Strahler, 1992). The spectral properties of the scene are determined using an analytical approximation to a solution of the radiative transfer equation.

Forerunners of this model have been used in previous studies over the same region to successfully characterise the canopy-leaving radiation field (Abuelgasim & Strahler, 1994; Wu & Strahler, 1994). The contrast between shaded and sunlit scene components in the model allows it to describe the BRF shape of

Table 1b
Parameters for DALEC for ponderosa pine (from Williams et al., 2005)

Parameter	Term	Value
T_1	Decomposition from litter to SOM	4.41×10^{-6}
T_2	Proportion of GPP lost to respiration	0.4733
T_3	Proportion of NPP sent to foliage	0.3150
T_4	Proportion of NPP sent to roots	0.4344
T_5	Rate of leaf loss	0.002665
T_6	Rate of wood loss	2.06×10^{-6}
T_7	Rate of root loss	2.48×10^{-3}
T_8	Rate of respiration from litter	2.28×10^{-2}
T_9	Rate of respiration from litter SOM	2.65×10^{-6}
C_f	Initial foliar carbon pool	57.705
C_w	Initial woody carbon pool	769.86
C_r	Initial root carbon pool	102.00
C_{lit}	Initial litter carbon pool	40.449
C_{som}	Initial SOM carbon pool	9896.7

sparse tree canopies more accurately than classical turbid medium approaches.

GORT is parameterised by the ratio of the mean crown centroid height to the crown vertical radius, h/b, the ratio of the vertical to horizontal crown radius, b/r, the projected crown cover, $\lambda\pi r^2$ (where λ is the tree stem density in m^{-2}), mean leaf area per unit area of tree crown FAVD$\cdot b\cdot 4/3$ where FAVD is the foliage area volume density (m^{-1}), leaf reflectance and transmittance and soil reflectance. Leaf optical properties are determined by the PROSPECT model (Jacquemoud & Baret, 1990; Jacquemoud et al., 1995) which is parameterised by leaf water content, chlorophyll concentration, the dry matter content, and a structural term, the effective number of leaf layers. Soil reflectance is assumed to be Lambertian here and is modelled using the empirical spectral functions described by Price (1990).

A component of the GORT model relies on computationally expensive integrations over the height of the canopy to determine the probability of a photon travelling from above the canopy to the ground (or vice-versa) without being scattered given that it a) does not enter a tree crown or, b) enters at least one tree crown. Case b) implies that the photon enters the geometric primitive that represents the crown but is uncollided. These cases are referred to $P(n=0|\theta)$ and $P(n>0|\theta)$ respectively where θ is the view or illumination geometry (depending on whether the photon is travelling upward or downward).

For the GORT model to be used as part of an EnKF scheme, where many runs of the model must be made each time an observation is available, the computational cost of these integrations is prohibitive. To avoid this cost a functional approximation to the probability terms has been developed:

$$P(n=0|\theta)=e^{-c/\mu} \tag{4}$$

$$P(n>0|\theta)=e^{-a/\mu}-P(n=0|\theta) \tag{5}$$

Where μ is the cosine of the appropriate zenith angle (i.e. viewing or illumination) and expressions for a and c are given below:

$$c=\lambda\pi r^2 \tag{6}$$

$$a=c\left(e^{0.0014166c^2}-e^{-k\mathrm{LAI}}\right) \tag{7}$$

The value for k is given by:

$$k=0.348535c^{-1.08069-0.0874595c} \tag{8}$$

and LAI$=\lambda\pi r^2\cdot$FAVD$\cdot b\cdot 4/3$ (described above). The constants in Eqs. (7) and (8) were derived empirically by fitting to the outputs of the GORT probability routines. The values for $P(n=0|\theta)$ and $P(n>0|\theta)$ may be plugged directly into the GORT model equations given in Ni et al. (1999) in place of the same values derived by numerical integration. The execution time of these simple expressions is typically several hundred times faster than the integration procedures used to produce the probabilities in the original model.

The GORT model was coupled to DALEC by assuming that the foliar biomass predicted by DALEC is related to the FAVD of the GORT model by:

$$C_f/\mathrm{GCLA}=\lambda\pi r^2.\mathrm{FAVD}\cdot b\cdot 4/3 \tag{9}$$

where GCLA is the grams of carbon per leaf area which has been determined in the field to be 111.0 g C m^{-2} (Williams et al., 2005). Canopy reflectance is sensitive to many more parameters than this single point of contact with the ecosystem model. Ideally, more of these parameters would be linked into DALEC, but for the present we must consider them simply as 'ancillary parameters' that are required to model spectral BRF, but which play no role in the ecosystem model. An estimate of these parameters is therefore required to drive the observation operator.

The look-up table (LUT) approach described by Weiss et al. (2000) was used to invert the GORT model against the time series of MODIS BRF data for the central pixel in the scene (that containing the tower site). The GORT model was used to construct a LUT of around 3×10^5 records, where for each simulation the model parameter values were drawn from transformed random noise varied within predefined ranges (Table 2a). The transformation approximately linearises the parameter with respect to reflectance. This reduces bias in the sampling of the LUT. The relative root mean squared error ($\mathrm{RMSE_{rel}}$) was used as the merit function:

$$\mathrm{RMSE_{rel}}=\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\frac{\rho_j(\Omega_v,\Omega_o)-\hat{\rho}_j(\Omega_v,\Omega_o,P)}{\rho_j(\Omega_v,\Omega_o)}\right)^2} \tag{10}$$

where N is the number of spectral bands, $\rho_j(\Omega_v, \Omega_o)$ is the measured reflectance at given viewing (Ω_v) and sun directions (Ω_o), $\hat{\rho}_j(\Omega_v, \Omega_o, P)$ is the modelled reflectance at the same viewing and sun directions, and P is the set of canopy and soil model parameters. As the wavebands used are not sensitive to canopy water content, the PROSPECT leaf water parameter was

Table 2a
GORT model parameter ranges and transformation used in generating the look-up table and inversion results

Parameter		Units	Min	Max	Transformation
FAVD$\cdot b\cdot 4/3$	Crown leaf area index	Unitless	0.1	5.0	$e^{-0.5\mathrm{FAVD}\cdot b\cdot 4/3}$
$\lambda\pi r^2$	Projected crown cover	Unitless	0.01	1.0	None
b/r	Crown height to width ratio	Unitless	0.5	1.5	None
h/b	Crown depth to vertical radius ratio	Unitless	0.5	2.0	None
C_{ab}	Leaf chlorophyll content	μg cm^{-2}	50	120	$e^{-0.01[\mathrm{Cab}]}$
ρ_{soil1}	Price's first soil basis function	Unitless	0.05	0.4	None
ρ_{soil2}	Price's second soil basis function	Unitless	0.0	0.1	None

All other model parameters are fixed.

Table 2b
GORT mean model parameter values, standard deviations and correlation matrix derived from the model inversion

	FAVD·b·4/3	$\lambda\pi r^2$	b/r	h/b	C_{ab}	ρ_{soil1}	ρ_{soil2}
Mean	2.77	0.63	0.74	1.36	68.35	0.28	0.05
Standard deviation	0.46	0.08	0.21	0.23	17.96	0.01	0.03
Correlation matrix							
FAVD·b·4/3	1.00	−0.83	0.08	0.19	−0.57	0.34	0.02
$\lambda\pi r^2$	−0.83	1.00	−0.20	−0.55	0.44	−0.56	−0.20
b/r	0.08	−0.20	1.00	0.26	−0.38	0.03	−0.07
h/b	0.19	−0.55	0.26	1.00	0.36	0.56	0.02
C_{ab}	−0.57	0.44	−0.38	0.36	1.00	−0.22	0.12
ρ_{soil1}	0.34	−0.56	0.03	0.56	−0.22	1.00	0.53
ρ_{soil2}	0.02	−0.20	−0.07	0.02	0.12	0.53	1.00

not adjusted. The assumed GCLA value of 111.0 g C m^{-2} was used to set the dry matter per unit leaf area term in the PROSPECT model. The number of leaf layers used in PROSPECT was estimated to be 2.0, based on consideration of the ratio of spectral reflectance to transmittance for pine needles. Estimates of the remaining parameters were obtained by finding the set of model realisations which resulted in the minimisation of the $RMSE_{rel}$ between modelled and measured reflectances, assuming the canopy variables to be constant over the (3-year) time period considered. The final parameters were selected by taking the mean parameters from a set of 50 lowest $RMSE_{rel}$ parameter combinations, with these samples also used to characterise the variance/covariance matrix of the inversion (Table 2b).

Observed MODIS reflectance values compare favourably with those forward modelled by GORT (Fig. 3), using the selected mean parameter set, for bands 1 and 2 as a function of viewing zenith angle. The characteristic BRDF shapes are well reproduced by GORT; a strong downward bowl in the red and a shallow, tilted, upward bowl in the near infrared. $RMSE_{rel}$ values are around 10% of the observed values.

3. Results

The DALEC model was run in an ensemble of 200 members for the years 2000 to 2002 inclusive using the meteorological data collected at the field site and the parameterisation previously determined by Williams et al. (2005) for each of the 81 pixels in the study area. On days when MODIS observations were available the EnKF was used to adjust the ensemble state so as to represent the best estimate of the true system state (the ensemble mean) and the second order error statistics. Data for comparison with the modelled fluxes and pools are available only at the central pixel, although the footprint of the flux tower will vary with wind direction.

3.1. Foliar biomass

The DALEC model induces variability in predicted reflectance via variability in its foliar biomass and thus C_f is the main parameter that is directly adjusted by the EnKF scheme. Without assimilating any data the model predicts a shallow annual variation in the foliar biomass and the associated uncertainty grows steadily. Fig. 4a shows data for the tower site, with Fig. 4b demonstrating spatial results over 9×9 pixels surrounding the site. The field measurements all lie within the uncertainty bounds, suggesting that the DALEC modelled uncertainties have been assigned reasonable values. Assimilating the MODIS BRF data, however, the modelled trajectory is pulled away at the point where the first observation is available, predicting values approximately 3 times higher than the model values without assimilation, and the seasonal cycle becomes more pronounced (Fig. 4a). The mean of the ensemble state moves closer to two of the field measurements, but further away from the other two. Only one measurement lies with a standard

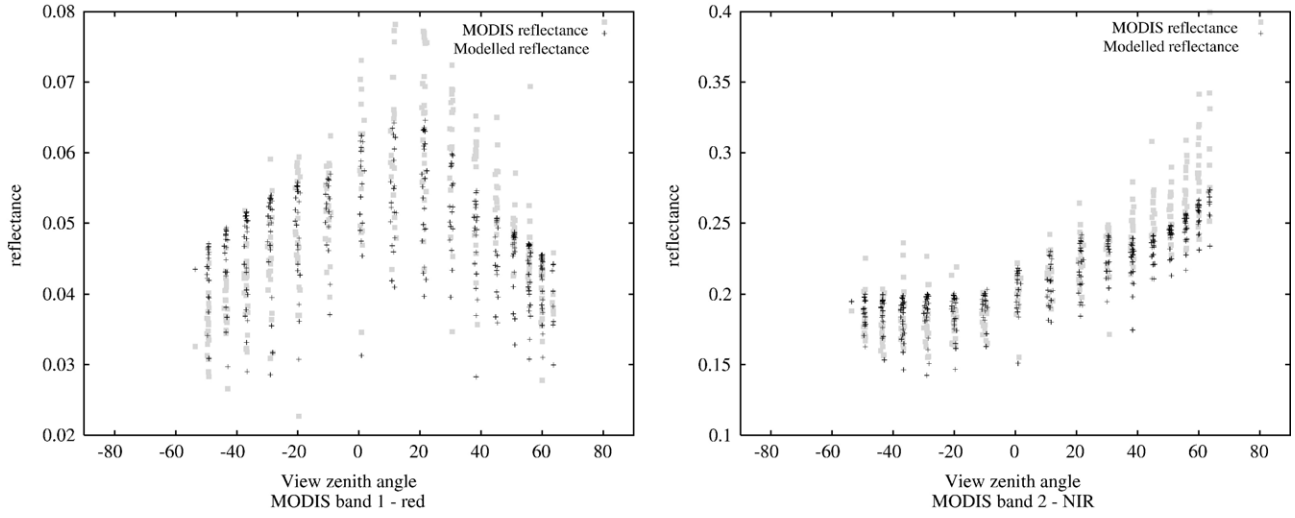
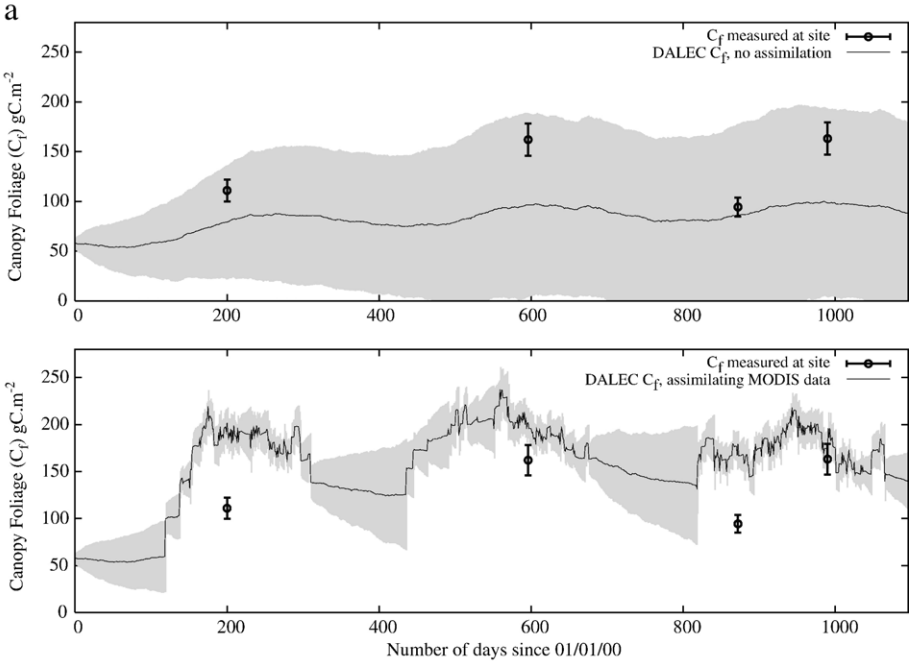
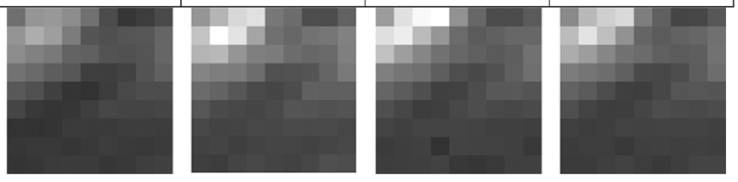


Fig. 3. Observed and modelled reflectances for the study area as a function of view zenith angle for bands 1 and 2 of MODIS. RMSEs (absolute) are 0.0076 in band 1 and 0.028 in band 2.

a



b



mean	188.1	230.0	218.5	212.2
std. dev.	55.0	75.5	84.0	70.6
min	128.1	152.3	128.1	143.7
max	351.4	515.7	534.0	442.2

Fig. 4. a. Foliar biomass predicted by the DALEC model with no assimilation (top panel) and assimilating reflectance data from MODIS bands 1 and 2 (bottom panel) for the central pixel (tower site). The grey spread about the line shows the uncertainty in the foliar biomass estimates. Field measurements taken by a LAI-2000 plant canopy analyzer and scaled up to biomass using estimates of grams of carbon per leaf area also plotted. b. Spatial results for mean foliar carbon (C_f) g C m^{-2}. Images are scaled linearly from 0 to 550 g C m^{-2}.

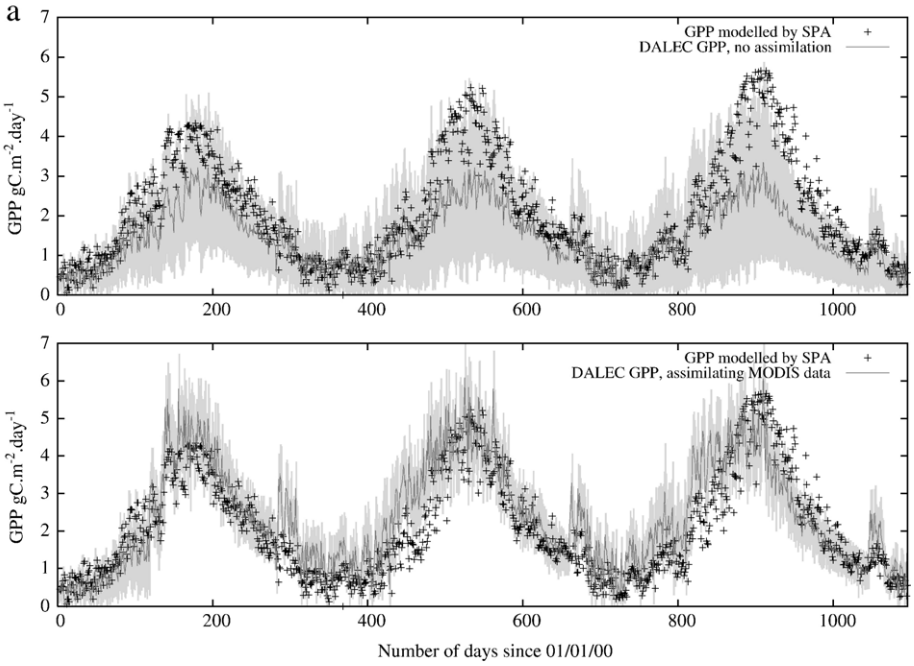
deviation of the modelled mean. The spatial variation of mean foliar biomass is large, ranging from 62.1 g C m^{-2} to 550.8 g C m^{-2}. The values (g C m^{-2}) for the central pixel are 131.6, 177.5, 164.3 and 157.8 for the years 2000, 2001, 2002 and three-year mean respectively.

One feature of assimilating the MODIS BRF that is evident in the foliar biomass results is the lack of samples over winter months. At the beginning of each year there is a period of approximately 100 days where all samples have been excluded because of clouds or snow cover (as determined from the MODIS QA data). The effect of this is that uncertainty in the ensemble mean of C_f grows steadily throughout the winter months. This in turn will affect the uncertainty in the rest of the model state. It might be reasonable to expect the data to induce a stronger dynamic across the annual cycle (i.e. a greater loss of foliage during the winter) if they were available.

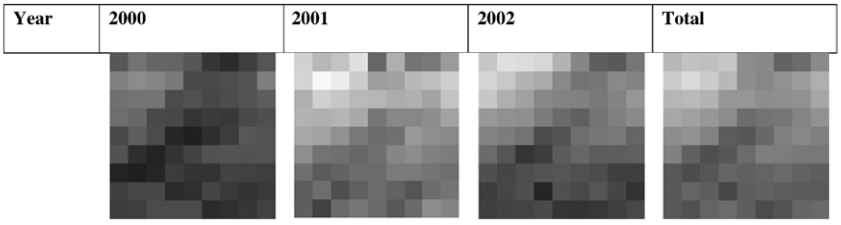
3.2. Gross primary productivity

The GPP is adjusted directly by the EnKF in response to the reflectance observation but is also increased by the extra foliage that is available to intercept incoming PAR. Fig. 5 shows the impact of assimilating MODIS data on the GPP over the three years, with Fig. 5a giving results for the central pixel and Fig. 5b the spatial results. For comparison modelled estimates of GPP from the Soil Plant Atmosphere (SPA) model (Williams et al., 1996) are also plotted. SPA is taken as a nominally accurate model of GPP and has been shown to produce reliable estimates

a



b



mean	876.1	1000.8	944.5	2820.7
std. dev.	50.0	68.4	79.5	184.6
min	776.7	862.7	794.7	2527.3
max	997.3	1181.2	1136.9	3275.3

Fig. 5. a. Gross Primary Production predicted by the DALEC model with (bottom) and without (top) assimilating MODIS reflectance data for the central pixel. The light grey spread around the ensemble mean is one standard deviation of the ensemble. The black crosses have been modelled by the SPA model. b. Spatial results for annual and total 3-year Gross Primary Production g C m^{-2}. Images are scaled linearly from 500 to 1200 g C m^{-2} (annual) and 1500 to 3600 g C m^{-2} (total).

of GPP for ponderosa pine ecosystems (Schwarz et al., 2004; Williams et al., 2001).

Assimilating the MODIS reflectance data moves the DALEC-predicted GPP toward the values predicted by the SPA model. During the first halves of 2001 and 2002 the GPP is higher than those values predicted by the SPA model.

Annual and three year summations of the GPP were compared for the model running with no data assimilation, assimilating MODIS reflectance data and the results presented by Williams et al. (2005) where all available field data (including NEP and the SPA modelled GPP) are assimilated into the DALEC model (Table 3). In 2000, the GPP estimate is clearly improved by the assimilation. In the following two years, however, the GPP predicted by assimilating the MODIS reflectance data is over-estimated by the same order of magnitude as it is underestimated when there is no data assimilation. The total GPP for the three years is over-estimated by around 400 g m^{-2} of carbon in comparison to the SPA modelling. The spatial results show a large degree of variability, with mean annual GPP ranging from 567.3 to 1177.7 g C m^{-2}. The central pixel value is towards the low end of the range for 2000, but close to the centre of the range for other years.

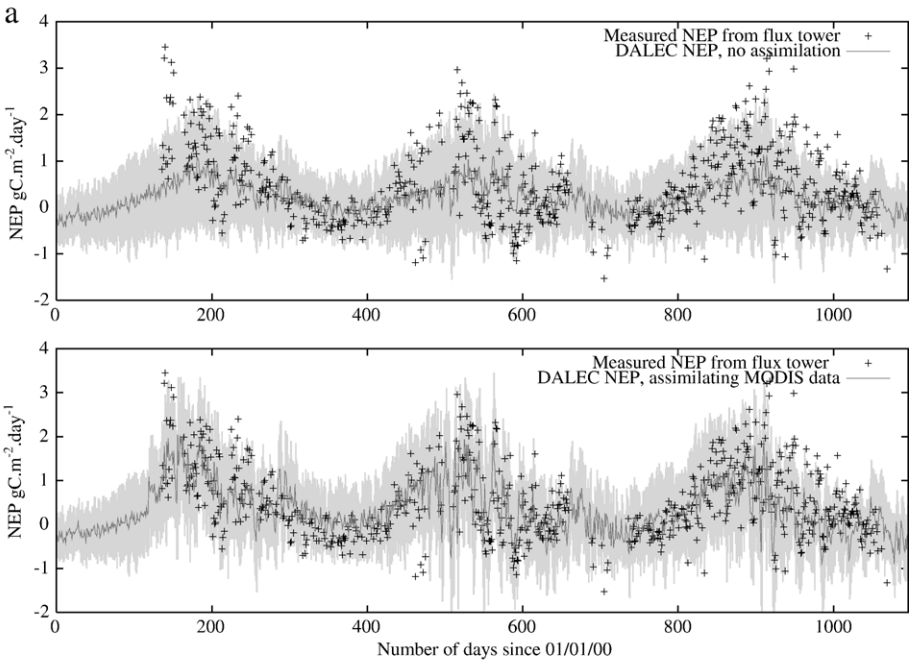
3.3. Net Ecosystem Production

There is a clear improvement in estimated NEP when reflectance data are assimilated into the ecosystem model (as shown in Fig. 6a). NEP is directly affected by the assimilation due to the effect on GPP, but it is also indirectly affected due to changes in heterotrophic respiration. (NEP is the difference

Table 3
Temporally averaged statistic for the major fluxes predicted by the DALEC model

Flux (g C m^{-2})	Assimilated data	2000	2001	2002	Total	Standard deviation
NEP	No assimilation	85.9	76.7	77.5	240.2	212.2
	MODIS B1 and B2	160.5	125.0	87.0	373.0	151.3
	Williams et al. (2005)	150.0	115.0	140.9	406.0	27.8
GPP	No assimilation	524.8	561.5	560.5	1646.4	834.5
	MODIS B1 and B2	780.7	941.6	898.8	2620.3	96.8
	Williams et al. (2005)	713.3	702.7	754.6	2170.3	18.1
Total respiration	No assimilation	438.8	484.8	483.1	1406.2	684.8
	MODIS B1 and B2	620.2	812.3	811.2	2247.1	155.3
	Williams et al. (2005)	563.6	587.7	613.7	1764.3	30.8

between the GPP and the total ecosystem respiration). The amount of C in the fine litter pool is controlled partly by the amount of leaf litter entering it, and this in turn is a function of the amount of foliar biomass. Changes in the foliar biomass have a gradual effect on the amount of the leaf litter pool. Consequently increases in foliar biomass tend to increase the

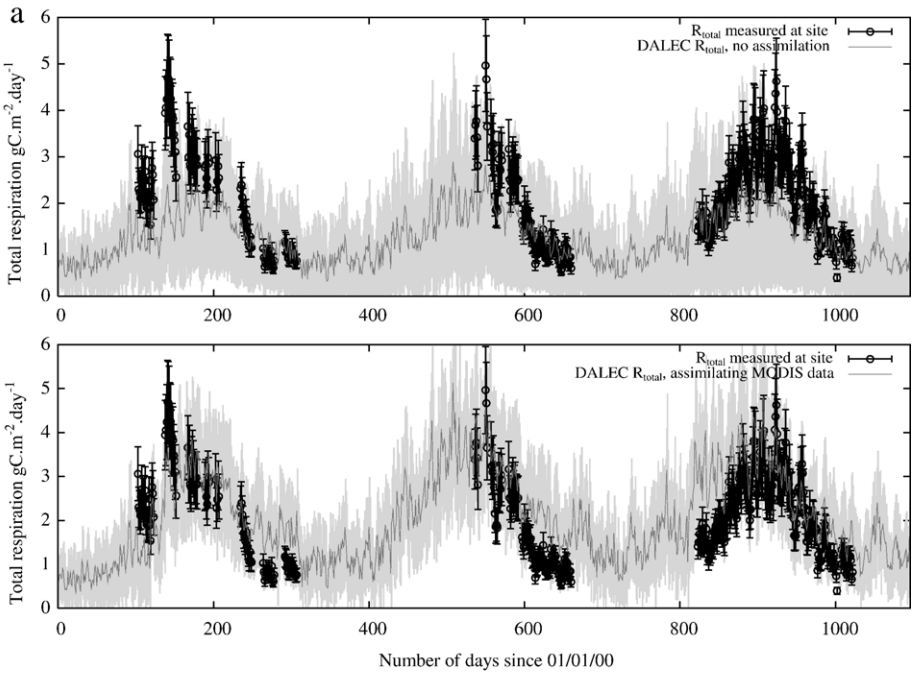


b

Year	2000	2001	2002	Total
mean	150.3	91.5	57.1	299.2
std. dev.	25.8	58.6	76.7	145.7
min	76.5	-181.6	-249.4	-239.6
max	198.0	189.7	193.2	524.7

Fig. 6. a. Net Ecosystem Production predicted by the DALEC model with (bottom) and without (top) assimilating MODIS reflectance data for the central pixel. The light grey spread around the ensemble mean is one standard deviation of the ensemble. The black crosses are data acquired by the eddy-covariance tower. b. Spatial results for annual and total 3-year Net Ecosystem Production g C m^{-2}. Images are scaled linearly from −300 to 300 g C m^{-2} (annual) and −500 to 500 g C m^{-2} (total).

a



b

Year	2000	2001	2002	Total
mean	725.7	909.3	887.4	2521.5
std. dev.	64.5	111.6	142.8	305.7
min	613.7	761.0	641.4	2140.2
max	920.8	1362.8	1369.7	3514.9

Fig. 7. a. Total ecosystem respiration predicted by the DALEC model with (bottom) and without (top) assimilating MODIS reflectance data for the central pixel. The light grey spread around the ensemble mean is one standard deviation of the ensemble. The black crosses are data acquired in the field, with error bars of one standard deviation. b. Spatial results for annual and total 3-year total ecosystem respiration g C m^{-2}. Images are scaled linearly from 500 to 1500 g C m^{-2} (annual) and 1500 to 4500 g C m^{-2} (total).

amount of carbon in the fine litter pool and, in turn, increase the respiration from this pool.

The plots show that assimilating the reflectance data improves the estimate of NEP and this is borne out by the statistics given in Table 3. The summed NEP for the entire time series is much closer to the values taken from Williams et al. (2005) than the model with no assimilation and the uncertainty in the estimate is greatly reduced. With the exception of 2001 the same is true of the annual totals. Fig. 6b shows spatial estimates of mean annual and total NEP. According to this model, the areas of high foliar biomass towards the top left of the images are losing C (NEP is negative), the maximum annual loss being of a similar magnitude to the maximum gain in the region.

3.4. Total ecosystem respiration

Given the general over-estimation of the GPP, the improvement in the estimates of NEP must result from over-estimation of ecosystem respiration (i.e. the sum of the auto- and heterotrophic components). Although field observations of the total respiration are sparser than those of NEP, respiration predicted by the model is generally too high when the reflectance data have been assimilated (Fig. 7a). Autotrophic respiration is raised proportionally with GPP and heterotrophic respiration increases as additional carbon is fed into the fine litter pool. The flux totals for the respiration in Table 3 show a constant over-estimate for each year.

Although the measurements of the total respiration are generally restricted to the second half of each year, it appears from the data points in the first half of 2002 (Fig. 7a) that the over-estimation is occurring in the early part of the year, as is the case with GPP.

Fig. 7b shows spatial results for total ecosystem respiration. Again, a large range of values is observed over the region, with the maximum being of the order of twice the minimum.

4. Discussion

The DALEC model is a simplified description of C dynamics, which makes it suitable for DA studies of this type. In the current use of the model, EO measurements of BRF data at moderate spatial resolution are assimilated into DALEC and the state variables modified sequentially i.e. based on current estimates of the state variables at the previous time step and a current set of observations. Whilst this provides the optimal estimate of model predictions at the next time step of operation, it does not guarantee an optimal estimate of fluxes considered over annual (or over other) time periods. For this reason alone, it may be necessary to consider other assimilation schemes in the future, such as the Ensemble Kalman Smoother or 4D-VAR, which allow observations to influence the entire model trajectory.

We propose that the assimilation of 'low-level' EO products is preferable to the use of standard-derived products of biophysical parameters for a number of reasons. In particular it is more straightforward to provide uncertainty estimates on lower-level products and hence they lend themselves to more self-consistent modelling systems in which the same assumptions about canopy properties affecting radiative behaviour can be made at all stages of modelling. This consistency is not entirely achieved here, although progress is made towards a general framework. One particular inconsistency that should be attended to in future work is the fact that radiative scattering is treated by the GORT model, based on a particular set of geometric assumptions about the canopy structure, and radiation interception is achieved by a simpler sub-model of ACM. Ideally these approaches should be unified.

Technical difficulties such as the need for a non-linear observation operator have been overcome in the development of this scheme by the use of an augmented state vector. A summary of the results would suggest that foliar C from the MODIS assimilation is over-estimated at the tower site (with respect to four ground measurements made on 100 m × 100 m plots), and GPP and total ecosystem respiration over-estimated as a result (compared to SPA modelled results driven by ground-measured LAI data). The over-estimation in the latter arises from the decomposition of extra leaf material after leaf fall. The NEP estimates arrived at here, the difference between these two fluxes, are close to those of Williams et al. (2005) after assimilation. A significant reduction in the uncertainties of the modelled fluxes is observed as a result of the EO data assimilation. All fluxes show a relatively strong spatial variability over the 4.5 km × 4.5 km area surrounding the site, which is attributable to variation in LAI over the region (all other parameters being constant over all pixels).

These results are generally very encouraging for the further application of the system developed here to the reduced uncertainty of spatialised estimates of C fluxes, although there are clearly some areas that require further testing and refinement. Perhaps the most significant of these are the issues relating to the apparently large discrepancy between the ground-measured LAI data of Williams et al. (2005) and the LAI modelled here (see Fig. 4 for the related quantity, foliar C). As noted above, all other major concerns about the model results result from this feature. Field data indicate that the stand increased canopy cover over the years 2000–2002, so that over the seasonal signal of LAI change there is also an aggrading canopy signal. It is of some concern that this is not visible in the results of the MODIS data assimilation.

We identify a number of routes through which such a discrepancy could occur, and consider each in turn in the following sections.

4.1. Errors associated with the GORT BRF model and MODIS BRF data

4.1.1. The non-applicability of model assumptions and model approximations

Various abstractions and assumptions go into the creation of any model. The utility of any model depends to a large degree on how appropriate such assumptions are. Examples of assumptions in the GORT model that might receive further attention are the treatment of lower boundary scattering as Lambertian. Further examples include assuming no understory vegetation, or the representation of trees having constant leaf area density within bounded spheroidal crowns. In fact, the landscape is relatively complex — the system has an evergreen open forest canopy with a deciduous understory. Scattering from this latter term is therefore not modelled here. The model was chosen to best mimic the conditions encountered in the study area (sparse to dense tree crowns on a bright background), but the impacts of these assumptions and further scales of clumping should be included in further studies.

In building a relatively complex BRF model such as GORT, various approximations are made within the formulation to keep the computational demand of the model within tolerable limits. A requirement for the EO operator in this scheme is that it is rapid enough to be able to run it for each day when an observation occurs, over a set of ensemble members (200 here) to obtain results for a single pixel. Other approximations within GORT, such as the relatively simplistic treatment of multiple scattering, have an unquantified impact on model error. This form of error can be assessed through the use of radiative transfer model testing and inter-comparison exercises (Pinty et al., 2004; Widlowski et al., 2007).

It is noted above that no MODIS observations are used when the MODIS QC snow flag is set. This limits the availability of data during the winter months (the regions of high uncertainty in Fig. 4a). The reason for this is that the lower boundary of the canopy (the ground) is currently treated as being Lambertian with an associated soil spectral function. The use of winter data

in future work will require the incorporation of a model of snow scattering into the observation operator.

4.1.2. Errors not characterised in the MODIS BRF data

This is an important factor for the practical use of such data. Although significant efforts are put into detecting clouds and cloud shadows, a proportion (estimated to be around 10% here) of 'clean' observations were perceived to have some contamination. We must therefore develop robust ways in which to filter contaminated data or quantify the additional uncertainty in the observations over the published values. In this paper, we achieve this filtering through the detection of outliers after modelling directional effects with a semi-empirical BRF model. The impact of contaminated data on the results of assimilation has not been investigated in detail here, but will clearly make the DALEC parameters more variable than they should be and may introduce bias.

4.1.3. Errors in the ancillary model parameters

The use of the GORT model as an EO operator necessitates estimation of ancillary model variables that are not directly linked to DALEC parameters. We have chosen to approach this issue in a pragmatic manner here by attempting to obtain estimates of these ancillary parameters from the inversion of the GORT model using MODIS BRF data. Assuming these parameters to be constant over time is clearly a rather simplistic assumption, but it is preferable to be simply assigning arbitrary values e.g. through a land cover map, as parameter values can at least be adapted for local conditions. The results for the ancillary parameters (Table 2b) show a high degree of (negative) correlation between crown LAI and both crown cover and leaf chlorophyll, suggesting that these terms cannot be reliably estimated from the model and red and NIR BRF data alone. The effects of such uncertainties on modelled reflectance are partially accounted for in the assimilation scheme, but this does not directly include a consideration of parameter bias or the fact that these 'fixed' parameters may vary over time. There are several ways in which this method might be improved. An iterative approach might be considered in which the updated estimate of LAI that results from one pass of the assimilation could be fed back into improving estimates of the static model parameters. Alternatively, the ancillary parameters could be included in the assimilation through the use of zero-order process models where no linkage can be found to DALEC state variables.

The results for the 'tree shape' terms, ratios h/b and b/r, describe prolate spheroids that are close to the ground. Ponderosa pine trees are however not easily considered as spheroids and are more conical in form. It is not thought that this difference in form would have a large impact on the modelling, but it makes it more difficult to compare the derived tree shape terms to any measured data.

The mean crown cover for the central pixel in the study area was estimated from the BRF inversion to be around 63%. Canopy cover from measurements in one of the 10 m radius subplots at the young pine site in 2000 was 66% vegetation, 34% soil (Law, unpublished data), so this seems a reasonable estimate for the tower pixel. However to facilitate processing for the results presented here this same crown cover was assumed for all other pixels in the spatial assimilations. This is arguably a weakness of the particular approach taken here as the IKONOS data in Fig. 1 show a clear spatial variation in crown cover. This could be simply overcome by performing a BRF model inversion on all pixels independently, although the cost of such processing is potentially significant if a dense LUT is used as in this case.

The product of the crown LAI and the projected crown cover obtained from the BRF inversion assuming all parameters to remain static over the three year period suggests a scene LAI of around 1.75, which equates to foliar C of 193.7 g C m^{-2}. The mean LAI over the 9×9 pixels of interest can be found from Fig. 4b to vary between 1.69 and 2.09. All of these are rather higher than any of the measurements of Williams et al. (2005) (measured on a 100×100 m grid), but not surprisingly they mostly lie around the mean of the C_f estimates arising from assimilating the MODIS BRF data (Fig. 4a). Law et al. (2001a) observed a mean LAI of 1.6 for 20 plots of different aged stands over 5×15 km in 2000 encompassing this study area. Turner et al. (2005) used spectral regressions between field observations of LAI from Law et al. (2001a) with tassled cap spectral vegetation indices from Landsat data and found a mean LAI of around 2.0 for a 5×5 km area around the tower. There is clearly some difference between the various LAI estimates over the site, but since they are at different scales and times, it is difficult to draw firm conclusions.

4.2. Errors associated with the DALEC model operation

Having considered errors in the BRF model and data, the next area that merits attention is the DALEC model itself. Issues associated with this are: (i) those resulting from model structure; and (ii) those resulting from error in the model parameters.

The first of these is relatively easy to address. The intention of DALEC is to define a simple ecosystem model, not a detailed model of plant processes. It is designed to allow the incorporation of observations to keep the model on track. Whilst it is therefore desirable to keep DALEC as simple as possible, the underlying model must be adequate to describe observed behaviour. Previous analyses have for example shown that the simple phenology components of DALEC do not accurately simulate the timing and allocation to, and losses from, the foliar biomass pool. A strength of the approach demonstrated here is that observational time series provided by EO can potentially be used to correct for such model drift, and thus observations of phenology compensate for model weaknesses. The value of this is clearly limited in the study as winter (snow) observations are not used and so phenological timing cannot be reliably estimated. In the analysis of Williams et al. (2005) leaf loss was forced by strongly assimilating litter fall data. Since there are no data here to constrain model leaf fall the behaviour of the model is entirely controlled by the underlying ecosystem model, which does not heavily induce leaf loss, resulting in a large over-estimation of GPP in the winter months, particularly in the winter of 2000–2001.

The absence of any valid measurements during winter periods shows up clearly in the profile for the foliar biomass (Fig. 4a). At the beginning of each year there is a period of around one hundred days where the model is running without assimilating any data; this is accompanied by the characteristic growth of uncertainty around the ensemble mean. It is possible that the seasonal cycle could be more pronounced if winter observations adjusted the foliar biomass estimates downwards. Field observations of seasonal LAI and summer maximum foliar biomass suggest that foliar biomass of the trees in winter is 30–40% lower than the summer maximum (foliar retention time is ~3 years; Law et al., 1999; Schwarz et al., 2004), whereas DALEC (Williams et al., 2005) suggests a winter decrease of ~25%. It would seem therefore that DALEC (without further data to constrain it) underestimates this decrease. This in turn leads to an over-estimate of LAI at the start of the growth season.

The other area of concern is the setting of the (initial state and rate) parameters that define DALEC. The parameter values for DALEC in this study are taken from Williams et al. (2005) and assumed to be constant over the study area. These were derived through a minimisation of the DALEC model outputs and the large number of variables available for assimilation. It is rather costly to attempt this type of parameterisation at a large number of other sites because of the quantity and type of measurements required. Any errors in the model prediction are treated only via a zero mean uncertainty, so biases introduced by incorrect model parameters are not considered. Alternative routes to the estimation of DALEC parameters should be considered for wider spatial application of this method. One approach currently being investigated is to include the model rate parameters in the assimilation and allow them to vary over time in response to the observations. Another approach would be to assess the uncertainty of these parameters and explicitly model them during the assimilation.

5. Conclusion

This paper aims to assess the feasibility of the assimilation of canopy BRF data into the simple ecosystem model, DALEC as an aid to improving spatial estimates of C fluxes and pools. The study has shown the approach to be feasible and can be considered successful in many respects, in particular in developing and linking the theory and tools required to allow assimilation of moderate resolution BRF data from MODIS into an ecosystem model through the use of an ensemble scheme with an augmented state vector allowing the treatment of a non-linear observation operator. It has also been demonstrated that BRF model inversion allows a practical method for the estimation of 'ancillary' parameters (those not directly linked to ecosystem model parameters) required by the observation operator. Although the ensuing parameters are derived from a simplistic assumption that they are constant over a three year period, a mechanism has been included to allow for uncertainty in these terms to be incorporated into the estimation of model error.

The performance of the assimilation of EO data in reducing uncertainty in C flux estimates has been demonstrated, although the quantity of foliar carbon (the one direct link between the EO data and the ecosystem model here) seems to be rather over-estimated (compared to published measurements). Particular problems in this regard are identified in winter when no EO observations were used. The over-estimate of foliar carbon leads to a general over-estimate of GPP and total ecosystem respiration, although these appear to approximately balance out to give a reasonable estimate of NEP. The reasons for this over-estimate, with the range of possible causes noted above, must be examined in further studies. Further constraints to the modelling from other EO data sources and other links to the model should also be followed up.

Acknowledgements

We would like to acknowledge the support of NERC in this work through funding of the NERC Centre for Terrestrial Carbon Dynamics (CTCD). We are also most grateful to Dr. Wenge Ni-Meister for the provision of source code for the GORT model, the supporting work of Paul Schwarz, James Irvine and Meredith Kurpius at OSU and to the anonymous reviewers for their helpful comments.

References

Abuelgasim, A. A., & Strahler, A. H. (1994). Modeling bidirectional radiance measurements collected by the Advanced Solid-State Array Spectro-radiometer (ASAS) over Oregon transect conifer forests. *Remote Sensing of Environment*, *47*, 261−275.

Anthoni, P. M., Unsworth, M. H., Law, B. E., Irvine, J., Baldocchi, D., & Moore, D. (2002). Seasonal differences in carbon and water vapor exchange in young and old-growth ponderosa pine ecosystems. *Agricultural and Forest Meteorology*, *111*, 203−222.

Churkina, G., Schimel, D., Braswell, B. H., & Xiao, X. (2005). Spatial analysis of growing season length control over net ecosystem exchange. *Global Change Biology*, *11*, 1777−1787.

Cohen, W. B., Maiersperger, T. K., Turner, D. P., Ritts, W. D., Pflugmacher, D., Kennedy, R. E., et al. (2006). MODIS land cover and LAI Collection 4 product quality across nine sites in the Western Hemisphere. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1843−1857.

Cosby, B. J. (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.

Davis, K. J., Bakwin, P. S., Yi, C., Berger, B. W., Zhao, C., Teclaw, R. M., et al. (2003). The annual cycles of CO_2 and H_2O exchange over a northern mixed forest as observed from a very tall tower. *Global Change Biology*, *9*, 1278−1293.

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*(C5), 10143−10162.

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*, 343−367.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., et al. (2003). Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *Journal of Geophysical Research*, *108* (D24), 4757−4783.

Gobron, N., Pinty, B., Verstraete, M. M., & Govaerts, Y. (1997). A semidiscrete model for the scattering of light by vegetation. *Journal of Geophysical Research*, *102*(D8), 9431−9446.

Gobron, N., Pinty, B., Verstraete, M., & Taberner, M. (2002). Medium resolution Imaging Spectrometer (MERIS): An optimised fAPAR algorithm theoretical basis document. Institute for Environment and Sustainability, Joint Research

Centre, TP 440 I-21020 Ispra (VA), Italy (Rev. 2.0 Jan. 22, 2002). Available from http://www-gvm.jrc.it/stars/Docs/atbd_meris_v3_gen.pdf

Granier, A. (1987). Evaluation of transpiration in a Douglas-fir stand by means of sap flow measurements. *Tree Physiology*, *4*, 309−320.

Heinsch, M. Z., Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J., Bolstad, P. V., et al. (2006). Evaluation of remote sensing based terrestrial production from MODIS using AmeriFlux eddy tower flux network observations. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1908−1925.

IPCC (2001). Climate Change 2001: The Scientific Basis. In J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, & C. A. Johnson (Eds.), *Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge: Cambridge University Press.

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

Irvine, J., Law, B. E., Anthoni, P. M., & Meinzer, F. C. (2002). Water limitations to carbon exchange in old-growth and young ponderosa pine stands. *Tree Physiology*, *22*, 189−196.

Irvine, J., Law, B. E., Kurpius, M. R., Anthoni, P. M., Moore, D., & Schwarz, P. A. (2004). Age-related changes in ecosystem structure and function and effects on water and carbon exchange in ponderosa pine. *Tree Physiology*, *24*, 753−763.

Jacquemoud, S., & Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, *34*, 75−91.

Jacquemoud, S., Ustin, S. L., Verdebout, V., Schmuck, G., Andreoli, G., & Hosgood, B. (1995, January 23–26). Prospect redux. *Summaries of the Fifth Annual JPL Airborne Earth Science Workshop*, *vol. 95–1*. (pp. 99−104) Pasadena, CA: Jet Propulsion Laboratory.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., et al. (2002). An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, *83*(1–2), 3−15.

Knorr, W. (1998). Constraining a global mechanistic vegetation model with satellite data. *Geoscience and Remote Sensing Symposium Proceedings, IGARSS'98, 6–10 Jul 1998* (pp. 971–973). doi:10.1109/IGARSS.1998.699643

Knorr, W., & Lakshmi, V. (2001). Assimilation of fAPAR and surface temperature into a land surface and vegetation model. In V. Lakshmi, J. Albertson, & J. Schaake (Eds.), *Land Surface Hydrology, Meteorology and Climate: Observations and Modelling* (pp. 177−200). Washington: AGU.

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Diner, D. J., & Running, S. (1998). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, *103*(D24), 32257−32276.

Kuusk, A. (1995). A fast, invertible canopy reflectance model. *Remote Sensing of Environment*, *51*(3), 342−350.

Law, B. E., Ryan, M. G., & Anthoni, P. M. (1999). Seasonal and annual respiration of a ponderosa pine ecosystem. *Global Change Biology*, *5*, 169−182.

Law, B. E., Thornton, P., Irvine, J., Van Tuyl, S., & Anthoni, P. (2001). Carbon storage and fluxes in ponderosa pine forests at different developmental stages. *Global Change Biology*, *7*, 755−777.

Law, B. E., Van Tuyl, S., Cescatti, A., & Baldocchi, D. D. (2001). Estimation of leaf area index in open-canopy ponderosa pine forests at different successional stages and management regimes in Oregon. *Agricultural and Forest Meteorology*, *108*, 1−14.

Law, B. E., & Waring, R. H. (1994). Combining remote sensing and climatic data to estimate net primary production across Oregon. *Ecological Applications*, *4*, 717−728.

Li, X., & Strahler, A. H. (1992). Geometric–optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 276−292.

Li, X., Strahler, A. H., & Woodcock, C. E. (1995). A hybrid geometric optical-radiative transfer approach for modeling albedo and directional reflectance of discontinuous canopies. *IEEE Transactions on Geoscience and Remote Sensing*, *33*, 466−480.

Mahrt, L., Lee, X., Black, A., Neumann, H., & Staebler, R. M. (2000). Vertical mixing in a partially open canopy. *Agricultural and Forest Meteorology*, *101*, 67−78.

Maisongrande, P., Duchemin, B., & Dedieu, G. (2004). VEGETATION/SPOT: An operational mission for Earth monitoring; presentation of new standard products. *International Journal of Remote Sensing*, *25*, 9−14.

McCloy, K. R., & Lucht, W. (2004). Comparative evaluation of seasonal patterns in long time series of satellite image data and simulations of a global vegetation model. *IEEE Transactions on Geoscience and Remote Sensing*, *42*, 140−153.

Ni, W. G., Li, X. W., Woodcock, C. E., Caetano, M. R., & Strahler, A. H. (1999). An analytical hybrid GORT model for bidirectional reflectance over discontinuous plant canopies. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(2), 987−999.

Nouvellon, Y., Moran, M. S., Seen, D. L., Bryant, B., Rambal, S., Ni, W., et al. (2001). Coupling a grassland ecosystem model with Landsat imagery for a 10-year simulation of carbon and water budgets. *Remote Sensing of Environment*, *78*, 131−149.

Pinty, B., Widlowski, J. -L., Taberner, M., Gobron, N., Verstraete, M. M., Disney, M., et al. (2004). Radiation Transfer Model Intercomparison (RAMI) exercise: Results from the second phase. *Journal of Geophysical Research*, *109*, D06210. doi:10.1029/2003JD004252

Plummer, S., Arino, A., Simon, M., & Steffen, W. (2006). Establishing an Earth Observation Product service for the terrestrial carbon community: The Globcarbon initiative. *Mitigation and Adaptation Strategies for Global Change*, *11*, 97−111.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., et al. (1993). Terrestrial ecosystem production — A process model-based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*, 811−841.

Price, J. C. (1990). Using spatial context in satellite data to infer regional scale evapotranspiration. *IEEE Transactions on Geoscience and Remote Sensing*, *28*, 940−948.

Rastetter, E. B. (2003). The collision of hypotheses: What can be learned from comparisons of ecosystem models? In C. D. Canham, J. J. Cole, & W. K. Lauenroth (Eds.), *Models in Ecosystem Science* (pp. 211−224). Princeton (NJ): Princeton University Press.

Rastetter, E. B., Aber, J. D., Peters, D. P. C., Ojima, D. S., & Burke, I. C. (2003). Using mechanistic models to scale ecological processes across space and time. *Bioscience*, *53*, 68−76.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., et al. (2005). Model-data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Global Change Biology*, *11*, 378−397.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H. (2005). Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochemical Cycles*, *19*. doi:10.1029/2004GB002254

Roy, D. P., Jin, Y., Lewis, P., & Justice, C. O. (2005). Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sensing of Environment*, *97*, 137−162.

Running, S. W., Baldocchi, D. D., Turner, D. P., Gower, S. T., Bakwin, P. S., & Hibbard, K. A. (1999). A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modelling and EOS satellite data. *Remote Sensing of Environment*, *70*, 108−127.

Running, S. W., Thornton, P. E., Nemani, R., & Glassy, J. M. (2000). Global terrestrial gross and net primary productivity from the Earth Observing System. In O. E. Sala, R. B. Jackson, H. A. Mooney, & R. W. Howarth (Eds.), *Methods in Ecosystem Science* (pp. 44−57). New York: Springer-Verlag.

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., et al. (2001). Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature*, *414*, 169−172.

Schwarz, P. A., Law, B. E., Williams, M., Irvine, J., Kurpius, M., & Moore, D. (2004). Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems. *Global Biogeochemical Cycles*, *18*, GB4006. doi:10.1029/2004GB002227

Tian, Y., Zhang, Y., Knyazikhin, Y., Myneni, R., Glassy, J. M., Dedieu, G., et al. (2000). Prototyping of MODIS LAI and FPAR algorithm with LASUR and LANDSAT data. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 2387−2401.

Treuhaft, R. N., Law, B. E., & Asner, G. P. (2004). Forest attributes from radar interferometric structure and its fusion with optical remote sensing. *BioScience*, *54*, 561−572.

Turner, D. P., Ritts, W. D., Cohen, W. B., Maeirsperger, T. K., Gower, S. T., Kirschbaum, A., et al. (2005). Site-level evaluation of satellite-based global terrestrial gross primary production and net primary production monitoring. *Global Change Biology*, *11*, 666−684.

Valentini, R., Matteucci, G., Dolman, A. J., Schulze, E. D., Rebmann, C., Moors, E. J., et al. (2000). Respiration as the main determinant of carbon balance in European forests. *Nature*, *404*, 861−865.

Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment*, *16*, 125−141.

Vermote, E. F., El Saleous, N. Z., & Justice, C. O. (2002). Atmospheric correction of MODIS data in the visible to middle infrared: First results. *Remote Sensing of Environment*, *83*, 97−111.

Vermote, E. F., El Saleous, N. Z., Justice, C. O., Kaufman, Y. J., Privette, J. L., Remer, L., et al. (1997). Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. *Journal of Geophysical Research*, *102*(D14), 17131−17141.

Vickers, D., & Mahrt, L. (1997). Quality control and flux sampling problems for tower and aircraft data. *Journal of Atmospheric and Oceanic Technology*, *14*, 512−526.

Vivoy, N., Francois, C., Bondeau, A., Krinner, G., Polcher, J., Kergoat, L., et al. (2001). Assimilation of remote sensing measurements into the ORCHIDEE/STOMATE DGVM biosphere model. *Proceedings: 8th International Symposium Physical Measurements & Signatures in Remote Sensing. Aussois, FRANCE* (pp. 713−718).

Wanner, W., Li, X., & Strahler, A. (1995). On the derivation of kernel-driven models of bidirectional reflectance. *Journal of Geophysical Research*, *100*(D10), 21077−21090.

Weiss, M., Baret, F., Myneni, R. B., Pragnère, A., & Knyazikhin, Y. (2000). Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie*, *20*, 3−22.

Widlowski, J. -L., Taberner, M., Pinty, B., Bruniquel-Pinel, V., Disney, M., Fernandes, R., et al. (2007). The third RAdiation transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance models. *Journal of Geophysical Research*, *112*, D09111. doi:10.1029/2006JD007821

Williams, M., Law, B. E., Anthoni, P. M., & Unsworth, M. (2001). Using a simulation model and ecosystem flux data to examine carbon–water interactions in ponderosa pine. *Tree Physiology*, *21*, 287−298.

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

Williams, M., Rastetter, E. B., Fernandes, D. N., Goulden, M. L., Wofsy, S. C., Shaver, G. R., 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., Schwarz, P. A., Law, B. E., Irvine, J., & Kurpius, M. (2005). An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, *11*, 89−105.

Wolfe, R. E., Nishihama, M., Fleig, A. J., Kuyper, J. A., Roy, D. P., Storey, J. C., et al. (2002). Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of Environment*, *83*, 31−49.

Wu, Y., & Strahler, A. H. (1994). Remote estimation of crown size, stand density, and biomass on the Oregon transect. *Ecological Applications*, *4*, 299−312.