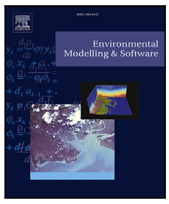
Exploring a hybrid ensemble–variational data assimilation technique (4DEnVar) with a simple ecosystem carbon model

Environmental Modelling and Software 186 (2025) 106361

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



Exploring a hybrid ensemble–variational data assimilation technique (4DEnVar) with a simple ecosystem carbon model

Natalie Douglas *, Tristan Quaife, Ross Bannister

National Centre for Earth Observation, Department of Meteorology, University of Reading, Reading, RG6 6AH, United Kingdom

ARTICLE INFO

Keywords:
Data assimilation
Ecosystem model
Land surface
Parameter estimation
4DEnVar

ABSTRACT

The study presented here evaluates the ability of the 4DEnVar data assimilation technique to estimate the parameters from synthetically generated observations from a simple carbon model. The method is particularly attractive in its speed and ease of use, and its avoidance in construction of adjoint or tangent linear model code. Additionally, the assimilation analysis step can be performed independently of ensemble generation; there is no need to integrate the 4DEnVar code with that of the underlying model, assuming parameters are static in time. The 4DEnVar method is capable of closely estimating the model parameters with increased certainty given that the ensemble produces a sufficient number of trajectories exhibiting behaviour seen in the observations. We find that the root mean squared error between trajectories and observations is significantly reduced when compared with the prior — in one case a 96% and 99% reduction in the biomass and soil pools respectively.

1. Introduction

Climate models provide prognostic information about future Earth System change and the likelihood of associated impacts on human life based on modelled trajectories of atmospheric CO_2. Land surface models, such as JULES (the Joint UK Land Environment Simulator) (Best et al., 2011; Clark et al., 2011) and ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) (Krinner et al., 2005), are used to provide surface boundary conditions for the atmospheric transport component of such Earth system models. Land surface models are designed to simulate a range of processes such as the fluxes of heat, water, momentum and CO_2 from the Earth's surface into the atmosphere. Some of these processes are governed by physical principles, in particular energy balance, but for many biological processes, such as those that govern carbon fluxes, there do not exist fundamental equations for these processes. Consequently, land surface models often rely on simplified or empirical representations of phenomena and the true values of the model parameters are often unknown.

DA and associated model-data fusion techniques seek to identify an optimal state or parameter set based on two sources of information: the prior knowledge of the state and/or parameters and a set of measured observations which, in general, will not belong to the same mathematical space. Uncertainty information corresponding to both priors and observations is required along with an observation operator that maps state/parameters to observation space. Current operational DA techniques include variational approaches such as 4DVar (Le Dimet and Talagrand, 1986; Rawlins et al., 2007) or sequential filtering procedures such as the Ensemble Kalman Filter Evensen (2003). There has been notable success in using DA with the earth observation (EO) datasets that are becoming increasingly available and we are starting to see hybrid techniques that combine the advantages to each of the aforementioned DA approaches emerge (Bannister, 2017).

Sujay et al. (2022) summarise progress made, and document major gaps that remain, utilising EO measurements in data assimilation methods in a state estimation context. Here, we highlight some of the relevant studies that used real data with various data assimilation techniques for the purposes of land surface model calibration. The pioneering activities in land surface model parameter estimation involved constraining the Simple Diagnostic Biosphere Model against in situ CO_2 flux measurements: Knorr and Heimann (1995) did this via simple loss minimisation quantifying the mismatch between the model and observations while Kaminski et al. (2002) added a term embodying a mismatch from the prior estimate of the parameters and including complex tangent and adjoint calculations. Several studies using similar techniques and observations for different land surface models followed. Rayner et al. (2005) used the Carbon Cycle Data Assimilation System (CCDAS) to constrain parameters of the Biosphere Energy Transfer Hydrology model (BETHY) while Kuppel et al. (2012) calibrated the ORCHIDEE model. Raoult et al. (2016)

* Corresponding author.
E-mail address: n.douglas@reading.ac.uk (N. Douglas).

https://doi.org/10.1016/j.envsoft.2025.106361
Received 16 August 2024; Received in revised form 22 January 2025; Accepted 1 February 2025
Available online 10 February 2025

conditioned the parameters of an early version of JULES (version 2.2) using site measurements of gross primary productivity (GPP) and latent heat. Williams et al. (2005) avoided the use of a cost function and calibrated the DALEC model using an embedded ensemble Kalman filter which required manipulation of model code to apply the DA algorithm. Multiple other parameter estimation studies using CO_2 flux measurements adopted Markov Chain Monte Carlo approaches (simplified photosynthesis and evapotranspiration model (SIPNET): Braswell et al., 2005; Zobitz et al., 2014; community land model (CLM): Post et al., 2017). These methods, although more computationally demanding, avoid complex adjoint calculations and model code manipulation and quantify uncertainty on the estimated parameters in the form of full posterior parameter distributions. The emergence of soil moisture observation datasets saw a shift in data assimilation activities and further attempts to avoid the construction of complex adjoint calculations. Pinnington et al. (2018) were able to improve estimates of shallow soil moisture modelled by JULES at a regional scale via an update to the model's soil parameters using the ESA Climate Change Initiative's satellite observations of soil moisture. Raoult et al. (2021) performed a similar activity calibrating ORCHIDEE parameters against in situ surface soil moisture measurements from the International Soil Moisture Network. These studies avoided the use of complex adjoint construction in the minimisation of their cost functions by implementing the Nelder-Simplex algorithm and performing a genetic random search respectively.

All of the studies mentioned above, and many other land surface parameter estimation studies, each have their own inherent difficulties pertaining to the data assimilation method chosen. All of these issues, in varying degrees, could be circumvented by the implementation of the 4DEnVar hybrid data assimilation method which was originally developed to incorporate flow-dependent background error covariances in numerical weather prediction (Liu et al., 2008; Desroziers et al., 2014). 4DEnVar combines the comparatively low computational requirements of gradient-descent optimisation of a cost function with the use of an ensemble to estimate its derivative information thus avoiding the construction of adjoint model code (which is usually required in standard 4DVar and which often needs to be updated with new model versions). Other methods that also avoid the construction of an adjoint, such as the MCMC method tend to require large computational power and can be slow in execution. In the parameter estimation problem, the DA aspect of 4DEnVar can be applied separately from the model runs, so there is no need for manipulation of model code to embed a DA algorithm. The ensemble itself can also be updated following optimisation and the updated ensemble can be used to approximate the posterior error covariance matrix. Recently we have seen applications of the 4DEnVar data assimilation method in land surface parameter estimation activities; 4DEnVar was used to calibrate JULES crop parameters at a single Nebraskan site using observations of GPP, leaf area index, and canopy height (Pinnington et al., 2020 - here the method is called 'LAVENDAR'). Subsequently, 4DEnVar was used to calibrate the pedotransfer functions in JULES to improve soil moisture estimates over East Anglia using NASA Soil Moisture Active Passive (SMAP) soil moisture data (Pinnington et al., 2021).

Two previous intercomparison studies have assessed the ability of different DA techniques to estimate parameters in simplified land surface carbon models. The OptIC (Optimisation InterComparison) project described in Trudinger et al. (2007) was a study to compare the ability of methods to successfully estimate the 4 parameters of the idealised two-store carbon model as presented by Raupach (2007). Participants were supplied with noisy synthetic observations generated from model runs in order to assess their ability in finding those parameters. A similar intercomparison project — the REFLEX project (Fox et al., 2009), was conducted where the DALEC model was used requiring the estimation of more parameters than those of the model used in the OptIC project. The additional value added by the REFLEX project was in its inclusion of real data where true parameter values are unknown, as would be the case when using real data to constrain the parameters of a complex realistic TBM. However, without known values for the true parameters, it is only possible to assess the consistency across participants in terms of the estimated parameter values and their associated confidence intervals.

The OptIC study had contributions from ten participants that included results from using the Kalman filter, Markov chain Monte Carlo and various variational DA methods. Regardless of their similar levels of ability to estimate parameters, the methods varied considerably in the number of model integrations required. Computational requirements varied from a few hundred to over a hundred thousand model runs; the usability or effectiveness of methods requiring a high number of iterations would be affected in an application to more realistic TBMs that have more parameters and model complexity and therefore much higher computational expense.

As new DA methods for tackling the parameter estimation problem emerge, revisiting the OptIC study provides means for assessing their utility. The purpose of the study presented here is to apply some aspects of the OptIC experiment to test the ability of the 4DEnVar DA technique to effectively estimate parameters with high confidence and relatively low computational power. We investigated the 4DEnVar method's ability to retrieve the parameters used to generate synthetic noisy truths for parameter sets that exhibit different behaviours in the model trajectories. These parameter sets are the same as some of those used in the OptIC experiment and here we only experiment with additive Gaussian noise. After some preliminary experiments and testing, and examining the root mean square error (RMSE) for different ensemble sizes, we found that the method is capable of estimating the toy carbon model parameters well and with increased certainty for some experiments. However, the method only performs well in cases where parameters do not generate extreme behaviour in the model trajectories. We also found that successful parameter estimation requires careful selection of a prior ensemble and with an effective parameter ensemble, 4DEnVar – for its ease and lack of computational expense – provides an attractive alternative DA method for use in parameter estimation.

2. Simple carbon model

The simple two-store carbon model that we use here is documented by Raupach (2007) and describes basic plant and soil carbon dynamics. All biomass carbon, including leaf, wood and root carbon, is treated as a single pool and all litter and soil carbon is treated in a second pool. This is an example of a producer-utiliser system in the sense that the plants are producers of organic carbon compounds in the form of biomass (when extracting atmospheric CO_2 during photosynthesis) and the soil is a utiliser feeding from plant litter (through heterotrophic organisms) and respiring carbon back to the atmosphere. This producer-utiliser system is in processor mode since the soil receives carbon passively as opposed to actively seeking carbon such as in a harvester type producer-utiliser system (e.g. biosphere-human carbon system).

2.1. Model equations

The dynamics between the biomass store x_1 and the soil store x_2 is defined by the two following idealised equations:

$$\frac{dx_1}{dt} = \underbrace{F(t)\left(\frac{x_1}{x_1+p_1}\right)\left(\frac{x_2}{x_2+p_2}\right)}_{\text{NPP}} - \underbrace{k_1 x_1}_{\text{litterfall}} + \underbrace{s_0}_{\text{seed}} \tag{1}$$

$$\frac{dx_2}{dt} = \underbrace{k_1 x_1}_{\text{litterfall}} - \underbrace{k_2 x_2}_{\text{resp.}} \tag{2}$$

where p_i is the limitation parameter for production by lack of x_i, k_i is the decay rate for store i, s_0 is a seed production term independent of x_1 and x_2 and $F(t)$ is a forcing factor dependent on time. We adopt the notation used in the OptIC project (Trudinger et al., 2007) as opposed

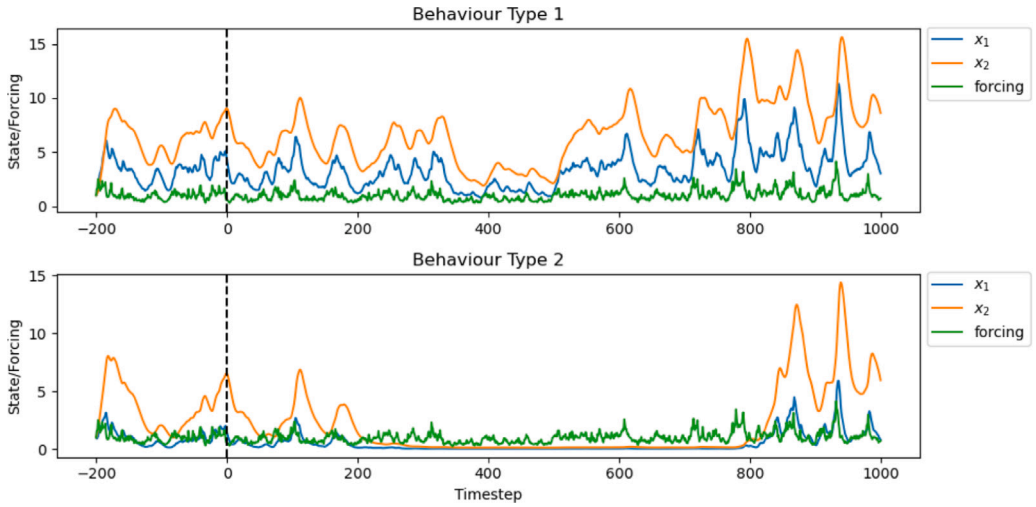


Fig. 1. The top plot shows fluctuation of the biomass pool around the "active-biosphere" steady state against time in days with parameters set to $p_1 = 1$, $p_2 = 1$, $k_1 = 0.2$, and $k_2 = 0.1$. The bottom plot shows flipping between the "active-biosphere" and "dormant-biosphere" steady states with parameters $p_1 = 1$, $p_2 = 1$, $k_1 = 0.4$, and $k_2 = 0.1$.

Table 1
Parameters present in the simple two-store carbon model along with their reference values and bounds given in the OptIC project.

Var./Par.	Description	Ref. value	Lower bound	Upper bound
p_1	Production limitation for biomass	1	0.5	5
p_2	Production limitation for litter/soil	1	0.5	5
k_1	Decay rate constant for biomass	0.2	0.03	0.9
k_2	Decay rate constant for litter/soil	0.1	0.01	0.12
s_0	Constant production term	0.01	–	–

to that given in Raupach (2007). The carbon pool fluxes are often measured in grams per square metre per day and so typical units are (1) g m^{-2} day^{-1} for $\frac{dx_1}{dt}$, $\frac{dx_2}{dt}$, $F(t)$ and s_0, (2) g m^{-2} for x_1, x_2, p_1 and p_2 and (3) day^{-1} for k_1 and k_2. The parameters in this simple carbon model with their reference values and bounds supplied in the OptIC project can be seen in Table 1.

The first term in Eq. (1) representing NPP is a product of the forcing $F(t)$ and two Michaelis–Menten form factors. The forcing term represents light and water availability and the p_1 and p_2 parameters prevent the Michaelis–Menten factors from taking unit value indicating a limitation to net primary productivity (NPP) through lack of biomass investment and nutrient availability respectively. The second term in Eq. (1) represents the loss of carbon from x_1, proportional to x_1 through the decay constant k_1, which we see entering x_2 in Eq. (2). A similar term showing loss via heterotrophic respiration can also be seen in the latter equation. The seed production term in Eq. (1), s_0, represents a small constant input to the biomass store, independent of the stores, generated from a long-term store of carbon (for example, in the form of seed propagules), and prevents extinction of the carbon stores in this model.

2.2. Equilibria and stability

The steady states that arise in this model in the case of constant forcing, i.e. the solutions to $\frac{dx_1}{dt} = \frac{dx_2}{dt} = 0$, and subsequently their stability properties, are determined by the parameter combinations. These are investigated and detailed in Raupach (2007). For the purposes of this study, we are interested in the solutions that arise when (i) random forcing is applied ($F(t)$ dependent on time) to reflect the effects of water and light availability from varying weather conditions and (ii) when $s_0 > 0$ to avoid depletion of the carbon pools that occurs when the stable steady state at $(x_1, x_2) = (0, 0)$ is reached irrespective of the forcing applied. We fix $s_0 = 0.01$ throughout and the forcing is determined by $F(t) = p_0 e^{m(t)}$ where $m(t)$ is a dimensionless Markov process with zero mean and standard deviation σ_m. In finite difference form $m_i = am_{i-1} + b\sigma_m w_i$ with $a = e^{\frac{-\Delta t}{T_m}}$ and $b = \sqrt{1 - a^2}$ and where w_i is a Gaussian random number with zero mean and unit variance. We used the values $p_0 = 1$, $\sigma_m = 0.5$, $T_m = 10\Delta t$ with the step size $\Delta t = 1$. This ensures that forcing remains positive and resembles the environmental limitations on NPP. Under these forcing conditions, there are two types of behaviour exhibited by the system:

1. the carbon pools fluctuate around the only "active-biosphere" steady state
2. the carbon pools flip between two steady states — the "active-biosphere" point and the (lower but non-zero) "dormant-biosphere" point. The forcing term induces a flip between the basins of attraction for each steady state, however, in some cases, the carbon pools can spend prolonged periods at the near zero steady state known here as mortality events.

Fig. 1 shows these two types of behaviours — the top plot shows fluctuation around the "active-biosphere" steady state with parameters set at the reference values whereas the bottom plot shows flipping between the "active-biosphere" and "dormant-biosphere" steady states when k_1 only is changed to 0.4.

3. Data assimilation

The aim of traditional variational data assimilation methods is to find an optimal initial state vector $\mathbf{x} \in \mathbb{R}^{n_x}$ often termed the analysis (or posterior estimate) that minimises a cost function that sums the distance to the prior estimate with the distance of the simulated observations to the actual observations. The errors are assumed to be normally distributed and the terms in the cost function are weighted by the inverses of their respective covariance matrices.

3.1. 4DVar and incremental 4DVar

In 4DVar, the cost function for state estimation, in compact form, is given by:

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}_\mathrm{b}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}_\mathrm{b}\right) + \frac{1}{2}\left(\hat{\mathbf{h}}(\mathbf{x})-\hat{\mathbf{y}}\right)^T \hat{\mathbf{R}}^{-1}\left(\hat{\mathbf{h}}(\mathbf{x})-\hat{\mathbf{y}}\right) \quad (3)$$

where prior errors are normally distributed with mean $\mathbf{x}_b \in \mathbb{R}^{n_x}$ and error covariance matrix $\mathbf{B} \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x}$. The measured and simulated observation vectors are

$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{pmatrix} \in \mathbb{R}^{n_o} \quad \text{and} \quad \hat{\mathbf{h}}(\mathbf{x}) = \begin{pmatrix} \mathbf{h}_0(\mathbf{x}) \\ \mathbf{h}_1\left(\mathbf{m}_{0\to 1}(\mathbf{x})\right) \\ \vdots \\ \mathbf{h}_N\left(\mathbf{m}_{0\to N}(\mathbf{x})\right) \end{pmatrix} \in \mathbb{R}^{n_o}$$

respectively where $\mathbf{y}_t \in \mathbb{R}^{n_{o,t}}$ are the observed observations and $\mathbf{h}_t \in \mathbb{R}^{n_{o,t}}$ is the observation operator at time t. The forward model $\mathbf{m}_t$ maps the state vector at time t to time $t+1$:

$$\mathbf{x}_{t+1} = \mathbf{m}_{t\to t+1}(\mathbf{x}_t)$$

and $\mathbf{x}_t = \mathbf{m}_{0\to t}(\mathbf{x}_0)$ maps the state from time 0 to time t. The observational error covariance matrix is

$$\hat{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_1 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_N \end{pmatrix} \in \mathbb{R}^{n_o} \times \mathbb{R}^{n_o}$$

where $\mathbf{R}_t \in \mathbb{R}^{n_{o,t}} \times \mathbb{R}^{n_{o,t}}$ corresponds to the observation error covariance matrix for the observations at time t. For the case when observations are uncorrelated, $\hat{\mathbf{R}}$ is a diagonal matrix and its inverse is trivial.

The cost function in Eq. (3) is generally non-quadratic due to the nonlinearity of the observation operator and the forward model. Optimising cost functions of this nature often results in locating a local rather than a global minimum. This problem can be overcome by adopting an incremental 4DVar (Courtier et al., 1994) approach whereby the optimised state is assumed to be of the form $\mathbf{x} = \mathbf{x}_b + \delta\mathbf{x}$, i.e. a small increment to the prior knowledge of the state. Subsequently, via a Taylor expansion of the forward model: $\mathbf{m}_{0\to t}(\mathbf{x}_b + \delta\mathbf{x}) \approx \mathbf{m}_{0\to t}\left(\mathbf{x}_b\right) + \mathbf{M}_{t\to 0}\delta\mathbf{x}$, where the tangent linear model $\mathbf{M}_{t\to 0} = \mathbf{M}_{t\to t-1}\mathbf{M}_{t-1\to t-2}\cdots\mathbf{M}_{1\to 0} \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x}$, the cost function in Eq. (3) becomes quadratic:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}\left(\hat{\mathbf{h}}\left(\mathbf{x}_b\right) + \hat{\mathbf{H}}\delta\mathbf{x} - \hat{\mathbf{y}}\right)^T \hat{\mathbf{R}}^{-1}\left(\hat{\mathbf{h}}\left(\mathbf{x}_b\right) + \hat{\mathbf{H}}\delta\mathbf{x} - \hat{\mathbf{y}}\right) \quad (4)$$

where

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_{1\to 0} \\ \vdots \\ \mathbf{H}_N\mathbf{M}_{N\to 0} \end{pmatrix} \in \mathbb{R}^{n_o} \times \mathbb{R}^{n_x}$$

and $\hat{\mathbf{H}}_t \in \mathbb{R}^{n_{o,t}} \times \mathbb{R}^{n_x}$ is the linearised observation operator at time t. One should ensure that $\mathbf{m}_{0\to t}(\mathbf{x}_b + \delta\mathbf{x}) - \mathbf{m}_{0\to t}\left(\mathbf{x}_b\right)$ is a good approximation for the evolution of $\mathbf{M}_{t\to 0}\delta\mathbf{x}$ over the time window when implementing incremental 4DVar. The gradient of the cost function in Eq. (4) is given by

$$\nabla J(\delta\mathbf{x}) = \mathbf{B}^{-1}\delta\mathbf{x} + \hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\left(\hat{\mathbf{h}}\left(\mathbf{x}_b\right) + \hat{\mathbf{H}}\delta\mathbf{x} - \hat{\mathbf{y}}\right) \quad (5)$$

where $\hat{\mathbf{H}}^T \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_o}$ is the adjoint of $\hat{\mathbf{H}}$.

The problem of finding the optimised state such that $\nabla J(\delta\mathbf{x}) = 0$ is usually tackled using gradient descent computational methods. Arguably the biggest problem in implementing these methods is determining the adjoint $\hat{\mathbf{H}}^T$ as it requires the calculation of model and observation operator derivatives. It is also important to note here that even though incremental 4DVar derives a quadratic cost function, it is possible that the minimisation locates a nonunique minimum. In addition to these common pitfalls the definition and conditioning of the prior error covariance matrix $\mathbf{B}$ is often an issue in implementing these methods.

3.2. 4DEnVar

Some of the drawbacks of 4DVar and incremental 4DVar, such as the need for the tangent linear or adjoint of the model and the lack of posterior uncertainty information, can be addressed by making use of an ensemble. In 4DEnVar, we take an ensemble of n_e state vectors $\{\mathbf{X}_{b,i}\}_{i=1,\dots,n_e}$ drawn from a $\mathcal{N}(\mathbf{x}_b, \mathbf{B})$ distribution and define the scaled perturbation matrix:

$$\mathbf{X}'_\mathrm{b} = \frac{1}{\sqrt{n_e-1}}\left(\mathbf{X}_{b,1} - \bar{\mathbf{x}}_b, \mathbf{X}_{b,2} - \bar{\mathbf{x}}_b, \dots, \mathbf{X}_{b,n_e} - \bar{\mathbf{x}}_b\right) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_e} \quad (6)$$

where $\bar{\mathbf{x}}_b \in \mathbb{R}^{n_x}$ is the mean of the ensemble. This offers a natural choice for the prior error covariance matrix that is preconditioned:

$$\mathbf{B} = \mathbf{X}'_b\left(\mathbf{X}'_b\right)^T. \quad (7)$$

We adopt the assumptions of incremental 4DVar and make a transformation of variables from $\delta\mathbf{x}$ in Eq. (4) to $\mathbf{w}$:

$$\mathbf{x}_a = \bar{\mathbf{x}}_b + \delta\mathbf{x} = \bar{\mathbf{x}}_b + \mathbf{X}'_b\mathbf{w} \quad (8)$$

where $\mathbf{w}$ is normally distributed with zero mean and unit covariance so that the analysis $\mathbf{x}_a$ is obtained from the mean of the ensemble updated via a weighted combination of the ensemble members. The cost function and gradient are

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\left(\hat{\mathbf{h}}\left(\bar{\mathbf{x}}_\mathrm{b}\right) + \mathbf{Y}'_b\mathbf{w} - \hat{\mathbf{y}}\right)^T \hat{\mathbf{R}}^{-1}\left(\hat{\mathbf{h}}\left(\bar{\mathbf{x}}_\mathrm{b}\right) + \mathbf{Y}'_b\mathbf{w} - \hat{\mathbf{y}}\right) \quad (9)$$

and

$$\nabla J(\mathbf{w}) = \mathbf{w} + \left(\mathbf{Y}'_b\right)^T \hat{\mathbf{R}}^{-1}\left(\hat{\mathbf{h}}\left(\bar{\mathbf{x}}_\mathrm{b}\right) + \mathbf{Y}'_b\mathbf{w} - \hat{\mathbf{y}}\right) \quad (10)$$

respectively where

$$\mathbf{Y}'_b = \frac{1}{\sqrt{n_e-1}}\left(\hat{\mathbf{h}}\left(\mathbf{X}_{\mathrm{b},1}\right) - \hat{\mathbf{h}}\left(\bar{\mathbf{x}}_\mathrm{b}\right), \dots, \hat{\mathbf{h}}\left(\mathbf{X}_{\mathrm{b},n_e}\right) - \hat{\mathbf{h}}\left(\bar{\mathbf{x}}_\mathrm{b}\right)\right) \in \mathbb{R}^{n_o} \times \mathbb{R}^{n_e} \quad (11)$$

approximates $\hat{\mathbf{H}}\mathbf{X}'_\mathrm{b}$ and this allows us to avoid the calculation of the adjoint derivatives. This is done by approximating perturbations in observation space with $\mathbf{Y}'_b$ rather than transporting state space perturbations to observation space with an adjoint. Note that the transpose of $\mathbf{Y}'_b$ in Eq. (10) is done explicitly since $\mathbf{Y}'_b$ is an explicit matrix. In addition to avoiding adjoint calculations, with the use of an ensemble we are able to borrow an EnKF result to update the ensemble and obtain a estimated posterior error covariance matrix:

$$\mathbf{X}'_a\left(\mathbf{X}'_a\right)^T = (\mathbf{I} - \mathbf{K}\hat{\mathbf{H}})\mathbf{X}'_b\left(\mathbf{X}'_b\right)^T \Longrightarrow \mathbf{X}'_a\left(\mathbf{X}'_a\right)^T \approx \mathbf{X}'_b\left(\mathbf{I} + \left(\mathbf{Y}'_b\right)^T\mathbf{R}^{-1}\mathbf{Y}'_b\right)^{-1} \times \left(\mathbf{X}'_b\right)^T \quad (12)$$

where $\mathbf{K}$ is the Kalman gain matrix that can be approximated using $\mathbf{H}\mathbf{X}'_b \approx \mathbf{Y}'_b$. This calculation involves the use of the Sherman–Morrison–Woodberry formula and corrects an error seen in Equation (A16) in Pinnington et al. (2021). It follows that the updated perturbations $\mathbf{X}'_a \approx \mathbf{X}'_b\left(\mathbf{I} + \left(\mathbf{Y}'_b\right)^T\mathbf{R}^{-1}\mathbf{Y}'_b\right)^{-\frac{1}{2}}$ where we can use a Cholesky decomposition to find a square root matrix which once inverted can be used to update the ensemble.

The context of all the above is that of data assimilation for state estimation. The same applies exactly for parameter estimation as is necessary for the purposes of this study, except that it should be noted that there is no model propagation of the parameter vector. The parameter vector to be optimised in the cost function is assumed fixed, that is to say $\mathbf{x}_{t+1} = \mathbf{x}_t$ where $\mathbf{x}$ now represents a vector of parameters. There can be significant simplifications and dimensional reductions in the context of parameter estimation. In these cases it may be possible to avoid computational gradient descent methods and make use of an analytical solution to $\nabla J(\mathbf{w}) = 0$. It follows from Eq. (10), this analytical solution is given by:

$$\mathbf{w} = -\left(\mathbf{I} + \left(\mathbf{Y}'_b\right)^T\hat{\mathbf{R}}^{-1}\mathbf{Y}'_b\right)^{-1}\left(\mathbf{Y}'_b\right)^T\hat{\mathbf{R}}^{-1}\left(\hat{\mathbf{h}}\left(\bar{\mathbf{x}}_\mathrm{b}\right) - \hat{\mathbf{y}}\right).$$

In the results that follow, we used the analytical solution as a check against the analysis obtained using gradient descent.

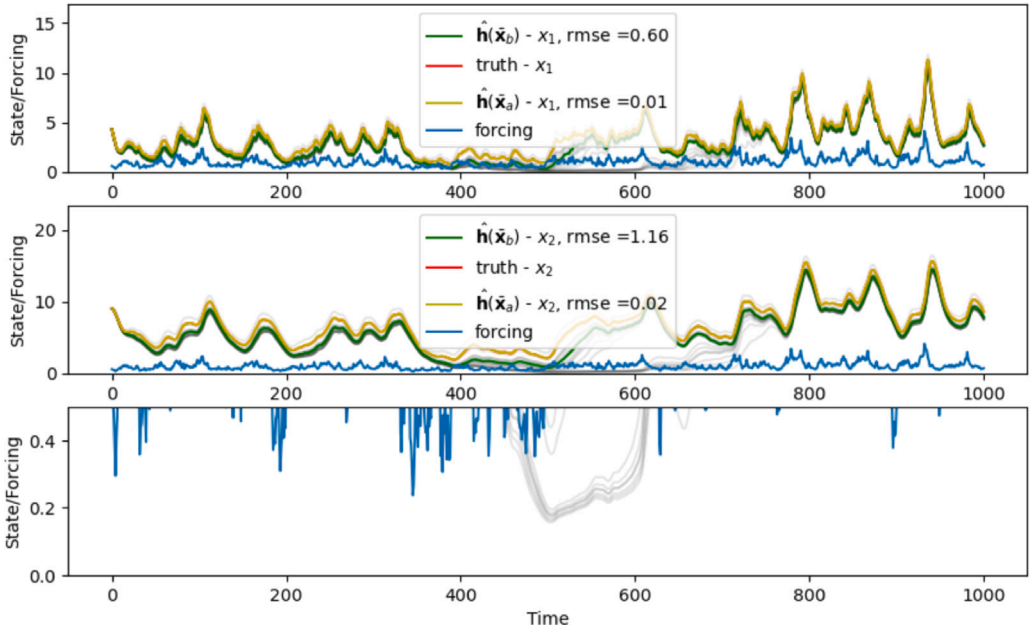


Fig. 2. Results from 4DEnVar for an ensemble size 20 of p_1 parameters. x_1 and x_2 trajectories are in the top and middle plots respectively: ensemble runs (grey), analyses (yellow), ensemble mean (green) and truth (red) unseen due to low RMSE with analyses. The bottom plot zooms in on the lower part of the x_1 trajectory with forcing shown in blue.

3.3. 4DEnVar tests

There are multiple tests that should be implemented to ensure that the 4DEnVar method is being applied correctly and any assumptions that are made are valid:

1. There is a linearity assumption on the derivation of the 4DEnVar cost function. It must be checked that the following linear approximation is a good one for the entire interval of its use: $\hat{\mathbf{h}}(\bar{\mathbf{x}}_b + \mathbf{X}'_b\mathbf{w}) \approx \hat{\mathbf{h}}(\bar{\mathbf{x}}_b) + \mathbf{Y}'_b\mathbf{w}$
2. Correct implementation of the cost function and its gradient can be checked via the expression:
$$\Phi(\alpha) = \frac{J(\mathbf{w} + \alpha\delta\mathbf{w}) - J(\mathbf{w})}{\alpha\delta\mathbf{w}^T\nabla J(\mathbf{w})} - 1 = O(\alpha) \quad (13)$$
where $\delta\mathbf{w}$ should be of unit length, e.g. $\delta\mathbf{w} = \frac{\nabla J}{|\nabla J|}$. Evaluating the expression for values of α closer and closer to zero should result in the expression tending to zero.
3. If one wishes to update the ensemble (derivation seen in Eq. (12)), it requires a Cholesky decomposition to find the square root matrix. By checking that the $n_e \times n_e$ matrix $\mathbf{I} + (\mathbf{Y}'_b)^T\mathbf{R}^{-1}\mathbf{Y}'_b$ has positive eigenvalues, and since it is symmetric by construction, then this matrix is positive definite and a unique Cholesky decomposition is ensured. This test should be implemented in any code used to apply the 4DEnVar method.

Fig. 3 shows the appropriate tests for the 4DEnVar method as outlined above. Panel (a) shows the expression in Eq. (13) tending to zero as α tends to zero illustrating the correct implementation of the cost function and its gradient. Panel (b) shows a comparison of $\mathbf{Y}'_b\mathbf{w}$ against $\hat{\mathbf{h}}\left(\mathbf{x}_b + \mathbf{X}'_b\mathbf{w}\right) - \hat{\mathbf{h}}\left(\mathbf{x}_b\right)$ for an optimised $\mathbf{w}$ in the reference case to ensure that the approximations hold well.

4. Results

As a basic check, the 4DEnVar method was employed in a preliminary study to test the method's efficacy in estimating the parameters individually (with all other parameters set to reference values — see Table 1) and highlighted the need for use of a restricted ensemble; the method was then applied to the full parameter set using a restricted ensemble in all experiments. The first assessment of applying 4DEnVar for all four parameters used unperturbed synthetic observations generated from the reference case set of parameters and includes an exploration of (1) the unrestricted and restricted parameter distributions and (2) the effect of RMSE with increasing ensemble size. The remaining experiments involved noisy synthetic observations using OptIC parameter sets (see Table 2) and in the first case we performed a principal component analysis. In all experiments a synthetic truth is obtained from a model run for different parameter sets for 1200 daily time steps of size $\Delta t = 1$. The first 200 time steps are discarded to disregard any transient effects of the initial state values — these are set to 1.0 for synthetic truths. Outputs of the synthetic model run for the remaining 1000 time steps are treated as observations. Observation error covariance matrices are diagonal with variances set arbitrarily at 10% of the corresponding observation values. The forcing is generated as detailed in Section 2.2 and the same forcing is used throughout. Parameters for each ensemble member (the ensemble size is 50 as motivated in Section 4.2, notably Fig. 5, unless otherwise stated) are drawn from a normal distribution with mean set to the midpoint of the bound interval as detailed in Table 1 and diagonal covariance matrix with standard deviations set arbitrarily to 25% of the mean values. Ensemble model runs are completed for 1000 time steps with initial values set to be equal to the first observations. All model runs were performed using Runge–Kutta (RK4) integration (step size $\Delta t = 1$). In the data assimilation step, the L-BFGS gradient descent method is implemented and python is used to generate all outputs (see Code Availability section).

4.1. Single parameter estimation

For the preliminary study the synthetic data was obtained from model runs with parameters set at the reference values. We took an ensemble of each parameter in turn, while fixing the other parameters to the reference value, in order to see if the 4DEnVar technique is able to retrieve each of the parameters individually without adding noise to the synthetic observations. This preliminary study illuminated the fact that many of the ensemble members represent mortality events – prolonged depletion in both carbon pools – in these cases the 4DEnVar technique fails with the ensemble size kept at 50. The technique is able to estimate parameters more closely if the ensemble size is increased (since more ensemble members exhibit the correct behaviour type 1 or

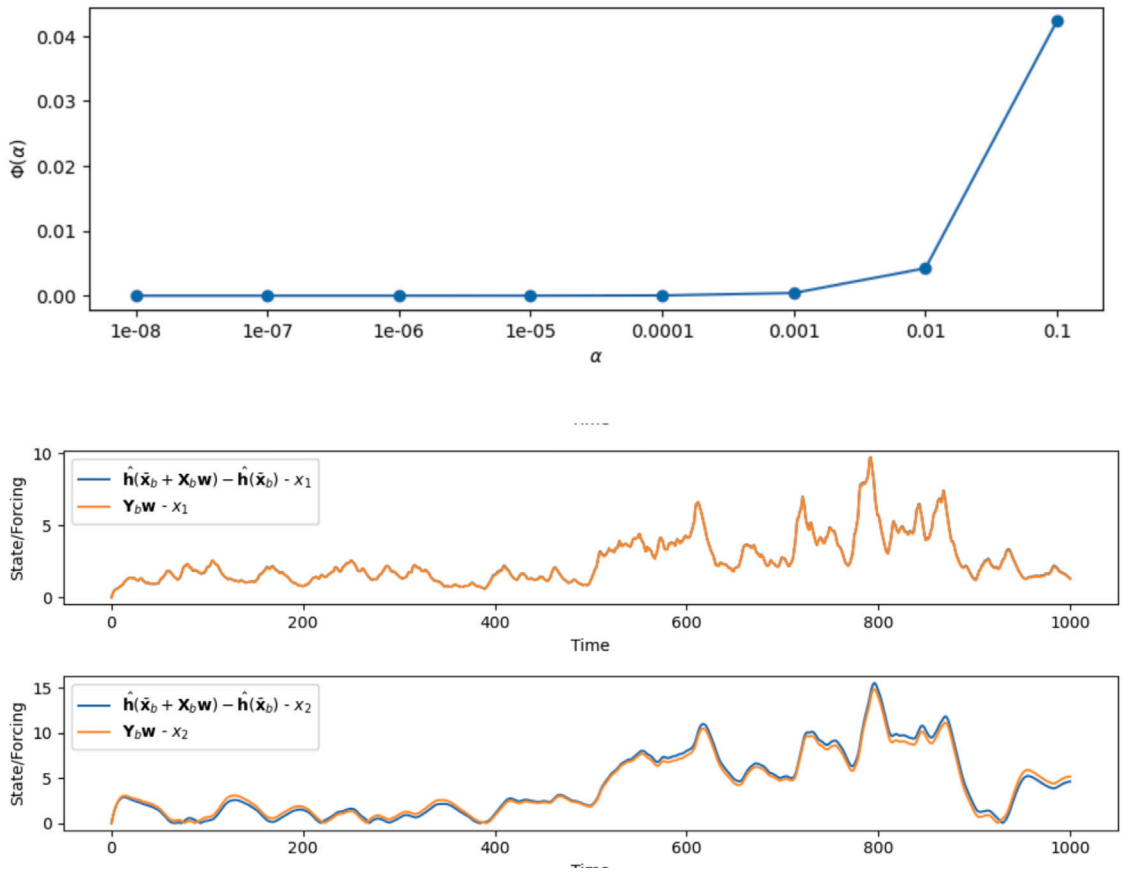
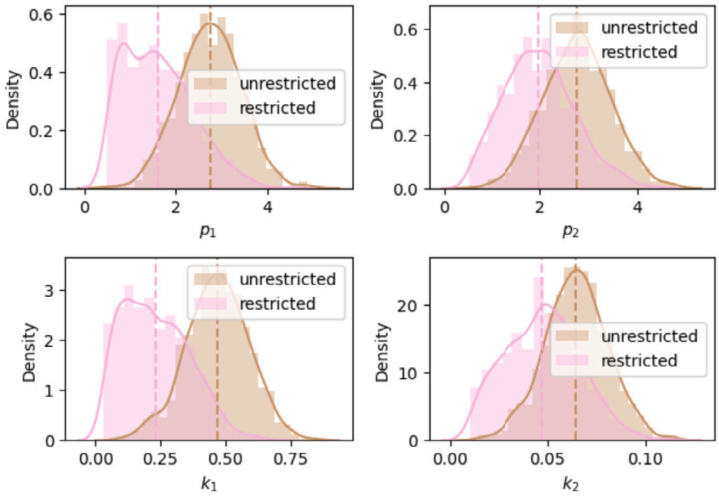


Fig. 3. Relevant tests for 4DEnVar implementation and assumptions.

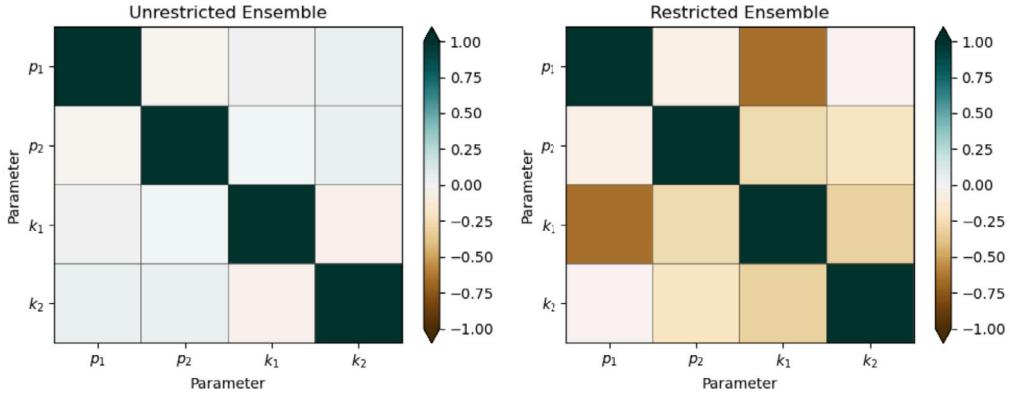
Table 2
True and estimated parameter values and RMSEs in x_1 and x_2.

	Parameters				RMSE in x_1			RMSE in x_2		
	p_1	p_2	k_1	k_2	$\hat{\mathbf{h}}(\bar{\mathbf{x}}_b)$	$\hat{\mathbf{h}}(\mathbf{x}_a)$	$\overline{\hat{\mathbf{h}}(\mathbf{X}_a)}$	$\hat{\mathbf{h}}(\bar{\mathbf{x}}_b)$	$\hat{\mathbf{h}}(\mathbf{x}_a)$	$\overline{\hat{\mathbf{h}}(\mathbf{X}_a)}$
True (Exp A.)	1.04	1.35	0.23	0.08	–	–	–	–	–	–
Estimated	1.089	1.285	0.233	0.079	0.282	0.078	0.141	5.905	0.064	0.276
Est./True	1.047	0.952	1.014	0.986	–	–	–	–	–	–
True (Exp. B)	2.44	2.45	0.11	0.031	–	–	–	–	–	–
Estimated	2.371	2.502	0.097	0.032	5.198	1.250	1.221	16.368	0.864	0.439
Est./True	0.972	1.021	0.878	1.022	–	–	–	–	–	–
True (Exp. C1)	2.44	2.45	0.11	0.011	–	–	–	–	–	–
Estimated	2.264	3.390	0.129	0.016	6.003	1.472	1.286	57.652	20.864	20.064
Est./True	0.927	1.384	1.173	1.427	–	–	–	–	–	–
True (Exp. C2)	0.77	2.73	0.033	0.025	–	–	–	–	–	–
Estimated	1.635	1.432	0.034	0.030	13.720	1.015	1.530	24.016	5.622	6.548
Est./True	2.123	0.524	1.034	1.187	–	–	–	–	–	–

less time at the low steady state for behaviour type 2 as seen in Fig. 1) but this will increase the computational requirements of the method. To keep the ensemble size small we incorporate a criteria on the parameters — reject ensemble members that result in the undesirable extinction behaviour of the carbon pools until an ensemble size of 50 is met. Resulting ensembles satisfying this requirement are henceforth referred to as a 'restricted ensemble'. For the model used here, this is possible by examining the steady states, in a non-forced setting, that exist for the given parameter combinations as detailed in Raupach (2007). Ensembles that give rise to a non-forced steady state in the carbon pools below a value of 0.5 can be excluded and sampling continued until the chosen ensemble size is met. Note that this will not be possible for all models — the steady states are simple to obtain for the model used here, but rejection based on persistent mortality of the observational trajectories could be employed nonetheless. Fig. 2 shows the results obtained when applying 4DEnVar to estimate the p_1 parameter using a restricted ensemble of size 20 (we see the unrestricted and restricted ensemble of size 50 in the full experiment in the next section). The technique is capable of estimating individual parameter values that generate carbon pool trajectories with very low root mean squared error (RMSE) when compared with the true trajectories. Very similar plots and results are obtained when repeating the experiment for the p_2, k_1 and k_2 parameters although some ensemble members only generated type 1 behaviour. It is worth noting that unrestricted large



(a) Unrestricted and restricted prior parameter ensemble distribution size 1000.



(b) Unrestricted and restricted correlation matrices.

Fig. 4. Distributions and correlation matrices for the unrestricted and restricted ensembles (size 1000).

ensembles performed poorly in comparison to restricted ensembles — the technique produced analyses with higher RMSEs.

The results of this preliminary study reveal the necessity to examine the performance of the technique using individual parameters. There should be a careful examination of the prior ensemble when performing any ensemble data assimilation technique for parameter estimation if the user is to have a full understanding of the mechanisms involved.

4.2. Four parameter estimation

Prior to running experiments using parameter sets from the OptIC experiment, 4DEnVar was tested with the reference parameter set with non-perturbed observations in order to check its ability to estimate all 4 parameters simultaneously. Similarly to the single parameter tests, it was found that a significant proportion of the parameter combinations in the ensembles resulted in either mortality events or Behaviour Type 2 — most of the model trajectories either stayed at zero or flipped between steady states or even prolonged fluctuation around the low steady state. Increasing the ensemble size did not work as effectively as it did for the single parameter experiments and so it was essential to generate an ensemble with some desirable behaviour in the trajectories. Ensemble member rejection was employed in the same way as for the single parameter estimation experiment.

Fig. 4 illustrates the parameter distributions before and after rejecting parameter combinations that allow for a second near zero steady state and the subsequent effect on the prior error covariance matrix $\mathbf{B} = \mathbf{X}'_b \left(\mathbf{X}'_b\right)^T$. For a large ensemble size of 1000 we see that restriction of the ensemble reduces the mean parameter values and introduces some left skewing (Fig. 4a) on the parameter distributions (p_1, k_1) which are otherwise symmetric and evidence of bi-modality (p_1). From Fig. 4b (right) we also see some correlations introduced between the parameters when the restricted ensemble is used, compared to the unrestricted ensemble.

In practical applications, ensemble sizes will typically be much smaller than 1000 — the ensemble size is often chosen primarily with regard to computational constraints. We performed an investigation to compare analysis RMSE with ensemble size for noisy synthetic observations obtained from the reference parameter set in order to identify the ideal ensemble size. The RMSE was obtained for ensemble sizes varying from 5 to 45 and the experiments were executed 20 times. Fig. 5 shows the mean RMSE (and 10th/90th percentiles) versus ensemble size. It is clear that RMSE in x_2 is consistently higher than that in x_1 but neither RMSE improves using an ensemble size larger than 30. We proceeded to explore the capabilities of 4DEnVar using the various parameter sets seen in the OptIC study, using noisy data and an ensemble size of 50 to be sure we had selected a large enough ensemble.

In the first non reference case experiment to estimate all four parameters simultaneously, a synthetic truth was obtained by running the model for the first parameter set used in the OptIC experiment — see Experiment A in Table 2. Noise was added to each of the observations y_i in the form $y_i + \sigma z_i$ where $\sigma = 0.1$ and the z_i were drawn from a random Gaussian distribution with zero mean and unit standard deviation. Fig. 6 shows an example of an ensemble where all ensemble

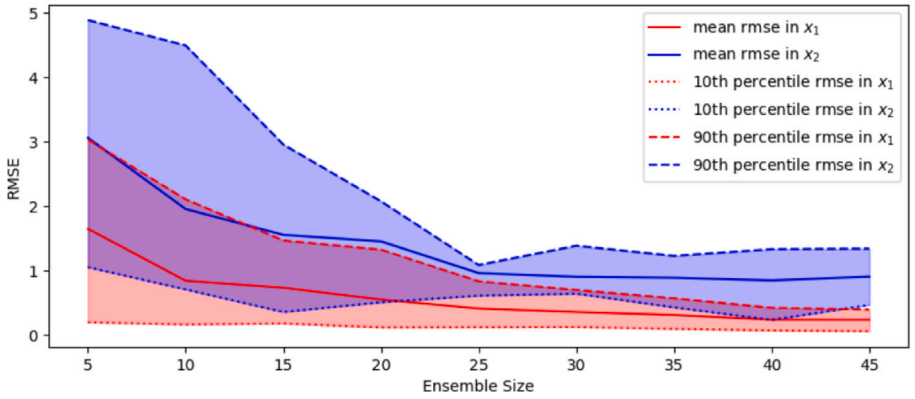


Fig. 5. Mean, 10th and 90th percentiles for analysis versus truth RMSE in x_1 and x_2 versus ensemble size for reference case parameters (experiment size 20).

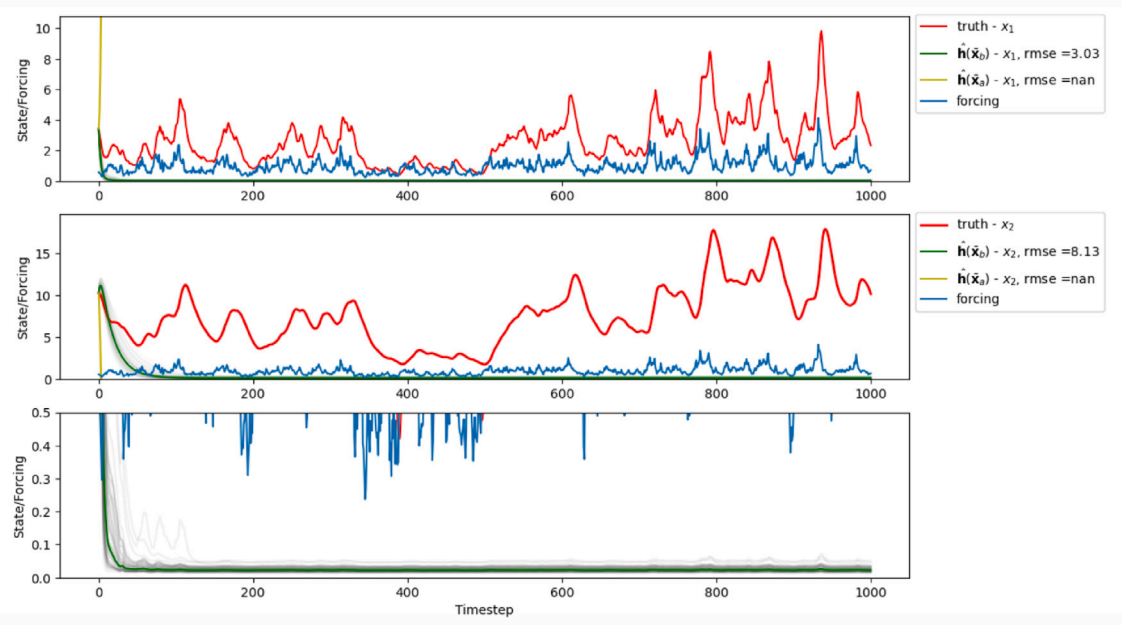


Fig. 6. Results from 4DEnVar for an unrestricted ensemble (size 50) of all parameters. x_1 and x_2 trajectories are in the top and middle plots respectively: ensemble runs (grey), truth (red), ensemble mean (green). A poor analysis (yellow) is obtained. Bottom panel shows a zoom of the top plot to show most trajectories staying in low steady states. Corresponding RMSE is shown in legend (NaN for failed DA) and forcing is plotted in blue.

members exhibit mortality behaviour and the data assimilation fails — the bottom panel shows a close look at the x_1 pool and we see that all of the unrestricted ensemble members converge to a steady state below 0.5. Fig. 7 shows the results of a successful application of 4DEnVar when restricting the ensemble to members that generate nontrivial behaviour and is capable of mimicking the observations. The RMSE in x_1 and x_2 is reduced by 96% and 99% respectively since the prior ensemble mean generates a trajectory with Type 2 behaviour. Fig. 8 shows the prior and posterior distributions of the parameters — all parameters move closer to their true values with significant narrowing of their spread indicating that these new parameter values are estimated with much higher confidence than their priors.

A principal component analysis of the posterior error covariance matrix obtained in the first nontrivial experiment reveals eigenvalues and eigenvectors (matrix columns) as follows:

$$[1.173e^{-01}, 9.406e^{-03}, 2.043e^{-04}, 4.983e^{-06}], \tag{14}$$

$$\begin{pmatrix} -0.379 & 0.925 & -0.0175 & 0.006 \\ 0.925 & 0.379 & -0.002 & 0.004 \\ 0.005 & -0.017 & -0.999 & 0.0215 \\ -0.002 & -0.006 & 0.022 & 0.999 \end{pmatrix} \tag{15}$$

The fourth eigenvector, corresponding to the smallest eigenvalue and therefore the best resolved direction in the optimisation, is dominated by the fourth entry corresponding to the parameter k_2. Similarly, the third eigenvector is dominated by k_1. The larger eigenvalues have eigenvectors with two significant entries corresponding to p_1 and p_2, with p_1 being the better resolved parameter. These findings agree with those of the OptIC experiment in that k_1 and k_2 are more easily observed and there is some correlation between p_1 and p_2 suggesting that it is more difficult to separate their effects.

The experiment was repeated for other parameter sets where Gaussian noise was added in the OptIC experiment as seen in Table 2. The RMSE in the prior ($\hat{\mathbf{h}}(\bar{\mathbf{x}}_b)$) and the analysis ($\hat{\mathbf{h}}(\mathbf{x}_a)$) are also included in this table to see the improvements along with the value of (estimated parameter value)÷(true parameter value) which is 1 when the parameters are perfectly estimated. A value less than (greater than) 1 indicates an underestimation (overestimation) of the parameter value. As an alternative to running the model with the updated ensemble mean, $\hat{\mathbf{h}}(\overline{\mathbf{x}_a})$, it is possible to choose the optimal trajectory as the one obtained from the mean of the updated ensemble of trajectories, $\overline{\hat{\mathbf{h}}(\mathbf{X}_a)}$. The RMSE for the mean updated ensemble trajectories is included for comparison. In all of the experiments, the RMSEs

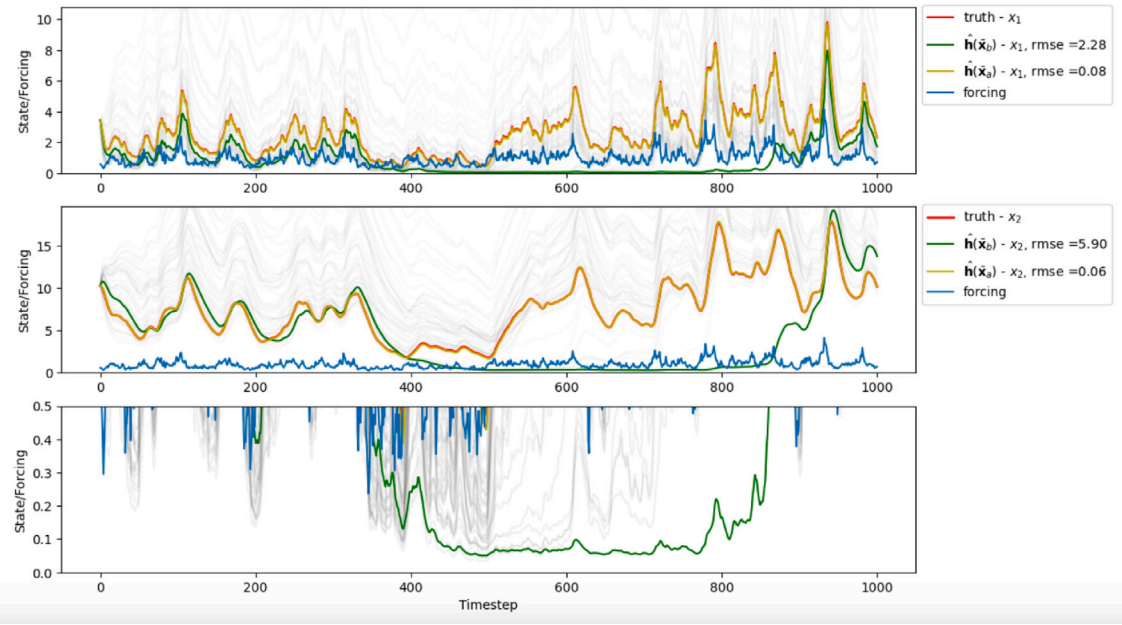


Fig. 7. Results from 4DEnVar for an restricted ensemble (size 50) of all parameters. x_1 and x_2 trajectories are in the top and middle plots respectively: ensemble runs (grey), truth (red), ensemble mean (green). A good analysis (yellow) is obtained. Bottom panel shows a zoom of the top plot to show no trajectories staying in low steady states. Corresponding RMSE is shown in legend and forcing is plotted in blue.

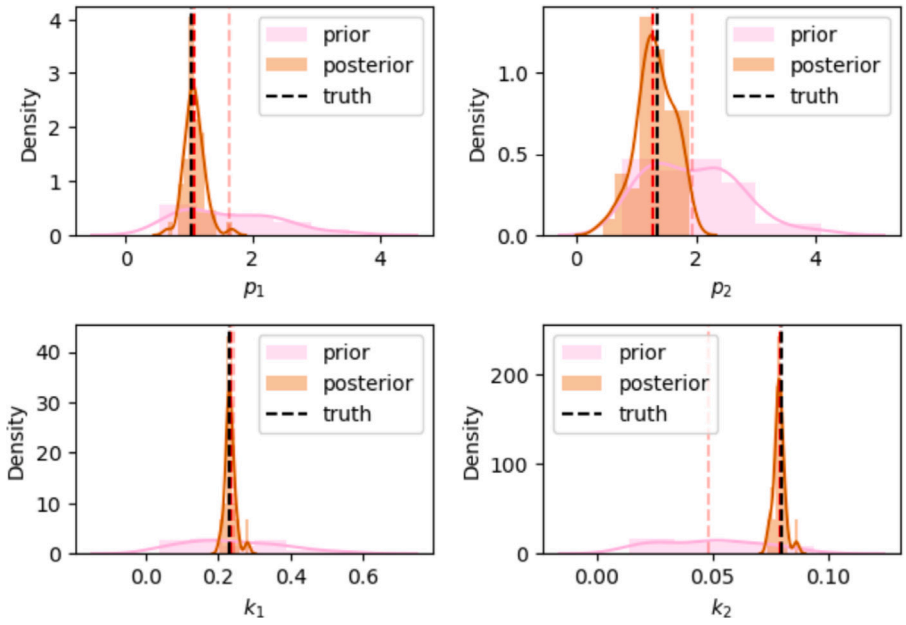


Fig. 8. Prior and posterior parameter distributions from the experiment shown in Fig. 5b using the restricted ensemble.

between the model trajectories obtained from the prior and the estimated parameter values are significantly reduced and the 4DEnVar method with more confidence placed on the observations is able to obtain parameter estimations with increased certainty, i.e. much more narrow posterior parameter distributions. For example, in Table 2, for experiment A, we see that our DA method estimates the parameter set to be $[p_1, p_2, k_1, k_2] = [1.089, 1.285, 0.233, 0.079]$ when the true values $[p_1, p_2, k_1, k_2] = [1.04, 1.35, 0.23, 0.08]$ are used to generate a noisy truth. This gives values for estimated/truth ratios as $[1.047, 0.952, 1.014, 0.986]$. Our method estimates k_1 and k_2 particularly well — within 2% of truth; estimates for p_1 and p_2 are not as accurate as those for k_1 and k_2, however, estimates still fall within 5% of the true values. In experiment B, estimates for p_1, p_2 and k_2 are within 3% of the true values but k_2 is relatively poorly estimated (within 13%). We also find that our method struggles to find close parameter estimates in the C1 and C2 experiments as was the case for the some of the members of the OptIC study. For the C1 experiment this is due to the ensemble (restricted or not) being unlikely to sample members that meet the extreme high model trajectory in the soil pool even for larger ensemble sizes and hence why the RMSE in the analysis, albeit smaller, remains relatively large. Similar behaviour of the trajectories is seen in experiment C2 — the carbon pool trajectory appears to be the upper extreme of what can be obtained by all parameter combinations and our method struggles to approximate the parameters well in these cases.

5. Discussion

Our method is a particularly attractive DA technique to employ since it does not require tangent linear or adjoint code – which can be a significant task for complex models and requires updating with each model release – when minimising the 4DEnVar cost function. Some computational effort is required to run the model on the parameter ensemble, particularly for more complex TBMS, but once an ensemble of model trajectories is obtained, multiple experiments can be performed offsetting the initial cost of set-up. In contrast, the 4DEnVar analysis step is computationally cheap thus fast and is executed separately from any model runs for parameter estimation so there is no need for adaptation of model code. The combined aforementioned features of the technique make for a compelling argument to employ 4DEnVar in parameter estimation studies. Nonetheless, a test to compare the 4DVar and 4DEnVar cost functions should be employed, particularly for models that are highly nonlinear when this method has potential to fail. It should be noted that operational land surface models such as JULES or ORCHIDEE have a lot more than just 4 parameters to optimise for and a parameter sensitivity analysis to the observations used with the 4DEnVar method is critical. Nevertheless, it is possible to extend the parameter set to include additional parameters, such as the seed term, for example, or even to include the initial state. In this investigation we chose our initial state to be the first set of available observations as was applied by some of the participants in the OptIC project but it should be noted that, for some models, not optimising for the initial state can have a detrimental impact on the value of the estimated parameters. It is also possible to optimise for the forcing parameters but there is no scientific value here since the assumptions made about the form of the forcing trajectory is not realistic in nature and driving data are typically assumed exact.

The success of the 4DEnVar method, when paired with the simple toy carbon model in particular, is hinged on a carefully chosen ensemble — parameter sampling from a normal distribution to choose the prior ensemble is insufficient to guarantee an effective ensemble has been selected. Firstly, discarding parameter values that fall outside the parameter bounds is necessary and this has an effect on the definition of the ensemble distribution. In the early stages of our experiment it was evident that even sampling in this way was insufficient for parameter estimation for this particular model — many of the ensemble members were generating trajectories that represent mortality of the carbon pools. It is possible to increase ensemble size in order to capture more of the variability in the model dynamics in the ensemble but in the interest of keeping computational requirements low, especially when applied with models with higher computational requirements, it may be necessary to restrict the ensemble. It should be noted that altering the ensemble in this way will have an effect on the prior parameter distributions and appear to contradict the underlying Gaussian assumptions of the method. Despite this, restricted ensembles with sufficient spread in the parameter distributions were adequate to ensure successful application of the 4DEnVar method. In the repeat experiment to see the effect of ensemble size on the RMSE, it was discovered that some restricted ensembles were incapable of effective parameter estimation when they contained insufficient spread in one or more of the parameters. A parameter sensitivity analysis and twin experiment to assess the efficacy of an ensemble to estimate parameters from synthetic data before using real observations is recommended with any application of 4DEnVar.

6. Conclusion

This study illustrates the success of the 4DEnVar method in its ability to estimate model parameters with increased certainty relative to prior parameter information using a relatively small carefully chosen ensemble of model runs given a set of noisy observations generated from a truth run in non-extreme cases. The 4DEnVar method is easy to implement with fast code to perform the DA algorithm that, in our use case (parameter estimation), is entirely separate from model code and avoids the requirement to compute an adjoint. Once an effective ensemble of model trajectories is obtained, it can be used in multiple experiments without the need for model reruns. For these reasons the 4DEnVar method is an attractive technique to implement in parameter estimation studies. It should be noted that the success of the method hangs on the careful selection of a prior ensemble – there should be sufficient variability in the trajectories generated from the ensemble members for efficient and effective parameter estimation – and the method still may struggle to estimate parameters for extreme cases.

CRediT authorship contribution statement

Natalie Douglas: Writing – original draft, Software, Investigation, Conceptualization. **Tristan Quaife:** Writing – review & editing, Supervision, Funding acquisition. **Ross Bannister:** Writing – review & editing, Supervision.

Software and data availability

All model code and detailed notes on the derivation of the 4DEnVar method can be found at:
https://github.com/NatalieDouglas/4DEnVar
•Name of software: fourDEnVar
•Developer: Natalie Douglas
•Year first available: 2023
•Program language: Python
•Cost: free
•Program size: 425B
•Data availability: https://github.com/NatalieDouglas/4DEnVar

Declaration of competing interest

All authors declare that they have no conflicts of interest.

Acknowledgements

This work was funded by the UKRI National Centre for Earth Observation, under the International Science Programme (NE/X006328/1).

Data availability

I have shared a link to my code.

References

Bannister, R.N., 2017. A review of operational methods of variational and ensemble-variational data assimilation. R. Meteorol. Soc. 143, 607–633.
Best, M., Pryor, M., Clark, D., Rooney, G., Essery, R., Menard, C., J.M., E., Hendry, M., Porson, A., Gedney, N., Mercado, L., Sitch, S., Blyth, E., Boucher, O., Cox, P., Grimmond, C., Harding, R., 2011. The joint UK land environment simulator (JULES), model description - part 1: Energy and water fluxes. Geosci. Model. Dev. 4, 677–699.
Braswell, B., Sacks, W., Linder, E., Schimel, D., 2005. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. Global Change Biol. 11, 335–355.
Clark, D., Mercado, L., Sitch, S., Jones, C., Gedney, N., Best, M., Pryor, M., Rooney, G., Essery, R., Blyth, E., Boucher, O., Harding, R., Huntingford, C., Cox, P., 2011. The joint UK land environment simulator (JULES), model description - part 2: Carbon fluxes and vegetation dynamics. Geosci. Model. Dev. 4, 701–722.
Courtier, P., Thepaut, J.-N., Hollingsworth, A., 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. R. Meteorol. Soc. 120, 1367–1387.
Desroziers, G., Camino, J.-T., Berre, L., 2014. 4Denvar: link with 4D state formulation of variational assimilation and different possible implementations. R. Meteorol. Soc. 140, 2097–2110.
Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean. Dyn. 53, 343–367.

Fox, A., Williams, A.D., Cameron, D., Gove, J.H., Quaife, T., Ricciuto, D., Reichstein, M., Tomelleri, E., Trudinger, C.M., Van Wijk, M.T., 2009. The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. Agricult. Forest. Meterol. 149, 1597–1615.

Kaminski, T., Knorr, W., Rayner, P., Heimann, M., 2002. Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle. Glob. Biogeochem. Cycles 16, http://dx.doi.org/10.1029/2001GB001463.

Knorr, W., Heimann, M., 1995. Impact of drought stress and other factors on seasonal land biosphere CO2 exchange studied through an atmospheric tracer transport model. Tellus B 47, 471–489.

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., Prentice, C., 2005. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. Glob. Biogeochem. Cycles 19, http://dx.doi.org/10.1029/2003GB002199.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., Richardson, A.D., 2012. Constraining a global ecosystem model with multi-site eddy-covariance data. Biogeosciences 9, 3757–3776.

Le Dimet, F.-X., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. Tellus 38A 2, 97–110.

Liu, C., Xiao, Q., Wang, B., 2008. An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formation and preliminary test. Mon. Weather Rev. 136, 3363–3373.

Pinnington, E., Amezcua, J., Cooper, E., Dadson, S., Ellis, R., Peng, J., Robinson, E., Morrison, R., Osborne, S., Quaife, T., 2021. Improving soil moisture prediction of a high-resolution land surface model by parameterising pedotransfer functions through assimilation of SMAP satellite data. Hydrol. Earth Syst. Sci. 25, 1617–1641.

Pinnington, E., Quaife, T., Black, E., 2018. Impact of remotely sensed soil moisture and precipitation on soil moisture prediction in a data assimilation system with the JULES land surface model. Hydrol. Earth Syst. Sci. 22, 2575–2588.

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., Scoby, D., 2020. The land variational ensemble data assimilation framework: LAVENDAR v1.0.0. Geosci. Model. Dev. 9, 55–69.

Post, H., Vrugt, J., Fox, A., Vereecken, H., Franssen, H.-J., 2017. Estimation of community land model parameters for an improved assessment of net carbon fluxes at European sites. Biogeosciences 122, 661–689.

Raoult, M.N., Jupp, T.E., Cox, P.M., Luke, C.M., 2016. Land-surface parameter optimisation using data assimilation techniques: the adJULES system V1.0. Geosci. Model. Dev. 9, 2833–2852.

Raoult, N., Ottle, C., Paylin, P., Bastrikov, V., Maugis, P., 2021. Evaluating and optimizing surface soil moisture drydowns in the ORCHIDEE model at in situ locations. J. Hydrometeorol. 22, 1025–1043.

Raupach, M., 2007. Dynamics of resource production and utilisation in two-component biosphere-human and terrestrial carbon systems. Hydrol. Earth Syst. Sci. 11, 875–889.

Rawlins, F., Ballard, S., Bovis, K.J., Clayton, A.M., Li, D., Inverarity, G.W., Lorenc, A.C., Payne, T.J., 2007. The met office global four-dimensional variational data assimilation scheme. R. Meteorol. Soc. 133, 347–362.

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). Glob. Biogeochem. Cycles 19, http://dx.doi.org/10.1029/2004GB00225.

Sujay, K., Kolassa, J., Reichle, R., Crow, W., de Lannoy, G., de Rosnay, P., Macbean, N., Girotto, M., Fox, A., Quaife, T., Draper, C., Forman, B., Balsamo, G., Steele-Dunne, S., Albergel, C., Bonan, B., Calvet, J.-C., Dong, J., Liddy, H., Ruston, B., 2022. An agenda for land data assimilation priorities: Realizing the promise of terrestrial water, energy, and vegetation observations from space. J. Adv. Model. Earth Syst. 14, http://dx.doi.org/10.1029/2022MS003259.

Trudinger, C.M., Raupach, M.R., Rayner, P.J., Kattge, J., Liu, Q., Pak, B.C., Reichstein, M., Renzullo, L., Richardson, A.E., Roxburgh, S.H., Styles, J.M., Wang, Y.P., Briggs, P.R., Barrett, D.J., Nikolova, S., 2007. The OptIC project: an intercomparison of optimisation techniques for parameter estimation in terrestrial biogeochemical models. J. Geophys. Res. 112, G02027.

Williams, M., Schwarz, P., Law, B., Irvine, J., Kurpius, M., 2005. An improved analysis of forest carbon dynamics using data assimilation. Global Change Biol. 11, 89–105.

Zobitz, J., Moore, D., Quaife, T., Braswell, B., Bergeson, A., Anthony, J., Monson, R., 2014. Joint data assimilation of satellite reflectance and net ecosystem exchange data constrains ecosystem carbon fluxes at a high-elevation subalpine forest. Agricult. Forest. Meterol. 195–196, 73–88.