INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2012; **68**:403–421
Published online 15 February 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/fld.2510

Variational data assimilation using targetted random walks

S. L. Cotter[1,*,†], M. Dashti[2] and A. M. Stuart[2]

[1]*OCCAM, Mathematical Institute, University of Oxford, Oxford OX1 3LB, U.K.*
[2]*Mathematics Institute, University of Warwick, Coventry CV4 7AL, U.K.*

SUMMARY

The variational approach to data assimilation is a widely used methodology for both online prediction and for reanalysis. In either of these scenarios, it can be important to assess uncertainties in the assimilated state. Ideally, it is desirable to have complete information concerning the Bayesian posterior distribution for unknown state given data. We show that complete computational probing of this posterior distribution is now within the reach in the offline situation. We introduce a Markov chain–Monte Carlo (MCMC) method which enables us to directly sample from the Bayesian posterior distribution on the unknown functions of interest given observations. Since we are aware that these methods are currently too computationally expensive to consider using in an online filtering scenario, we frame this in the context of offline reanalysis. Using a simple random walk-type MCMC method, we are able to characterize the posterior distribution using only evaluations of the forward model of the problem, and of the model and data mismatch. No adjoint model is required for the method we use; however, more sophisticated MCMC methods are available which exploit derivative information. For simplicity of exposition, we consider the problem of assimilating data, either Eulerian or Lagrangian, into a low Reynolds number flow in a two-dimensional periodic geometry. We will show that in many cases it is possible to recover the initial condition and model error (which we describe as unknown forcing to the model) from data, and that with increasing amounts of informative data, the uncertainty in our estimations reduces. Copyright © 2011 John Wiley & Sons, Ltd.

Received 31 August 2010; Revised 26 October 2010; Accepted 28 October 2010

KEY WORDS: uncertainty quantification; probabilistic methods; stochastic problems; transport; incompressible flow; partial differential equations

1. INTRODUCTION

Data assimilation, using either filtering [1, 2] or variational approaches [3–5], is now a standard component of computational modelling in the geophysical sciences. It presents a major challenge to computational science as the space-time distributed nature of the data makes it an intrinsically four-dimensional problem. Filtering methods break this dimensionality by seeking sequential updates of three-dimensional problems and are hence very desirable when online prediction is required [6–8]. However, when reanalysis (hindcasting) is needed, for example to facilitate parameter estimation in sub-grid scale models, it is natural to confront the fully four-dimensional nature of the problem and not to impose a direction of time; variational methods are then natural [9–11].

However, for both filtering and variational methods an outstanding computational challenge is the incorporation of uncertainty into the state estimation. Many of the practical methods used to confront this issue involve *ad hoc* and uncontrolled approximations, such as the ensemble Kalman filter [12]. While this has led to a very efficient methodology, and will doubtless continue to have

*Correspondence to: S. L. Cotter, OCCAM, Mathematical Institute, University of Oxford, Oxford OX1 3LB, U.K.
†E-mail: cotter@maths.ox.ac.uk

significant impact for some time to come, there remains scope for developing new methods which, while more computationally demanding, attack the problem of quantifying statistical uncertainty without making *ad hoc* or uncontrolled approximations. The natural setting in which to undertake the development of such new methods is the Bayesian framework in which one seeks to characterize the posterior distribution of state given data [13]. The Markov chain-Monte Carlo (MCMC) methodology provides a well-founded approach to fully probing this posterior distribution. The purpose of this paper is to show that sampling of this posterior distribution is now starting to fall within reach, via MCMC methods, in the offline situation. We also emphasize the possibility of, in future work, using such MCMC based studies of the posterior distribution to benchmark more practical algorithms such as filtering and variational approaches, both of which may be viewed as providing approximations to the posterior.

The Bayesian approach to inverse problems [14] is a conceptually attractive approach to regularization which simultaneously forces consideration of a number of important modelling issues such as the precise specification of prior assumptions and the description of observational noise [15]. However, in the context of PDE inverse problems for functions the approach leads to a significant computational difficulty, namely probing a probability measure on the function space. To be clear: the probabilistic viewpoint adds a further degree of high dimensionality to the problem, over and above that stemming from the need to represent the unknown function itself. Variational methods, which minimize a functional which is a sum of terms measuring both model-data mismatch and prior information, corresponds to finding the state of maximal posterior probability, known as maximum *a posteriori* (MAP) estimator in statistics [15]. This will be a successful approach if the posterior distribution is in some sense 'close' to Gaussian and with small variance. However, since the dynamics inherent within the observation operator can be highly non-linear, and since the observations may be sparse, it is not necessarily the case that a Gaussian approximation is valid; even if it is, then the spread around the MAP estimator (the variance) may be important and detailed information about it is required. Characterizing the whole posterior distribution can be important, then, as it quantifies the uncertainty inherent in state estimation.

Markov chain–Monte Carlo (MCMC) methods are a highly flexible family of algorithms for sampling probability distributions in high dimensions [16, 17]. There exists substantial literature on the use of the MCMC methodology for the solution of inverse problems from fluid mechanics [18–23] and from other application domains [24–28]. A key computational innovation in this paper, which distinguishes the methods from those in the preceding references, is that the algorithms we use estimate the posterior distribution on a function space in a manner which is robust to mesh refinement.

We employ random walk Metropolis-type algorithms, introduced in [29], and generalized to the high-dimensional setting in [30–32]. The method that we implement here does not use any gradient information (adjoint solvers) for the model-data mismatch. It proposes states using a random walk on the function space and employs a random accept/reject mechanism guided by the model-data mismatch on the proposed state. This method is hence straightforward to implement, but not the most efficient method: other methods (Langevin or Hybrid Monte–Carlo), which require implementation of an adjoint model, can explore the state space in a more efficient manner [32]. In this paper, we stick to the simpler Random Walk method as a proof of concept, but the reader should be aware that gradient methods, such as the Langevin Algorithm, or Hybrid Monte–Carlo, could be implemented to increase the efficiency. The basic mathematical framework within which we work is that described in [33] in which we formulate Bayesian inverse problems on the function space, and then study data assimilation problems arising in fluid mechanics from this point of view. The abstract framework of [33] also provides a natural setting for the design of random walk-type MCMC methods, appropriate to the function space setting. The probabilistic approximation theory associated to such algorithms is studied in [34].

In Section 2, we describe the data assimilation problems that we use to illustrate our Bayesian MCMC approach; these problems involve the incorporation of data (Eulerian or Lagrangian) into a model for Stokes' flow. Section 3 briefly describes the mathematical setting for our data assimilation problems, and for the MCMC algorithm that we employ throughout the paper. In

DOI: 10.1002/fld

Sections 4 and 5, we consider Eulerian and Lagrangian observations, respectively, incorporating them into a two-dimensional Stokes' flow model with periodic boundary conditions. In each case we will consider first the problem of recovering the initial condition of the velocity field, given that we assume that we know the forcing present in the system during the window of observation. We will then consider, in each of the two data scenarios, what happens if we assume that we know the forcing inherent in the system and incorporate this into the statistical algorithm, but in fact there is a degree of model error which can only be accounted for by altering the distribution on the initial condition. Finally, we will consider full initial condition and model error inference (which together are equivalent to finding a probability distribution on the state of the velocity field for the duration of the window of observation). In particular, we study the extent to which model error may be recovered from data, and the uncertainty in this process. In Section 6, we present some brief conclusions.

2. BAYESIAN DATA ASSIMILATION IN STOKES' FLOW

We describe three steps required for the Bayesian formulation of any inverse problem involving data: specification of the forward model; specification of the observational model; and specification of prior models on the desired states.

2.1. Forward model

We consider a fluid governed by a forced Stokes' flow on a two-dimensional box with periodic boundary conditions (for shorthand we write $\mathbb{T}^2$, the two-dimensional torus):

$$\partial_t v - \nu\triangle v + \nabla p = f \quad \forall (x,t) \in \mathbb{T}^2 \times (0,\infty), \tag{1}$$

$$\nabla \cdot v = 0 \quad \forall t \in (0,\infty), \tag{2}$$

$$v(x,0) = u(x) \quad x \in \mathbb{T}^2. \tag{3}$$

This equation determines a unique velocity field v, given the initial condition u and the forcing f. Our objective is to find u and/or f, given observations of the fluid. It is sometimes convenient to employ the compact notation which follows from writing Stokes' equation as on ordinary differential equation for a divergence-free velocity field [35]:

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \nu A v = \eta, \quad v(0) = u. \tag{4}$$

Here A is e Stokes' operator, the Laplacian on divergence-free fields; and η is the projection of the forcing f onto divergence-free fields.

We have chosen this simple linear model of a fluid for two reasons: (i) the linear nature enables us, when observations are also linear, to construct exact Gaussian posterior distributions which can be used to check MCMC methods and (ii) the linearity of the model means that the FFT can be employed to rapidly evaluate the forward model, which is desirable as this may need to be performed millions of times in order to obtain the complete statistical sampling of the posterior. Regarding (i) note, however, that we will also consider non-linear observations and then the posterior is not Gaussian. Regarding (ii) note that the MCMC method is trivially paralellizable and our choice of the linear forward model is made simply to facilitate computations without recourse to the use of a large number of processors; for more realistic forward models, however, such massive parallelization would be necessary.

2.2. Observations

We consider two kinds of observations: Eulerian, which are direct measurements of the velocity field and Lagrangian which are measurements of passive tracers advected by the velocity field. We label the data vector by y and assume, for simplicity, that it is always subject to mean zero

DOI: 10.1002/fld

Gaussian observational noise σ with covariance Γ. Note, however, that other non-Gaussian models for the observational noise are easily incorporated into the methodology herein.

In the Eulerian case the observations are

$$y_{j,k} = v(x_j, t_k) + \sigma_{j,k}, \quad j = 1, \ldots, J, \;\; k = 1, \ldots, K. \tag{5}$$

In the Lagrangian case, we define the J Lagrangian tracers $\{z_j\}_{j=1}^J$ as solutions of the ODEs

$$\frac{\mathrm{d}z_j}{\mathrm{d}t}(t) = v(z_j(t), t), \quad z_j(0) = z_{j,0}, \tag{6}$$

where we assume that the set of starting positions $\{z_{j,0}\}_{j=1}^J$ is known (although they too could be part of the estimated state if desired). The Lagrangian observations are then

$$y_{j,k} = z_j(t_k) + \sigma_{j,k}, \quad j = 1, \ldots, J, \;\; k = 1, \ldots, K. \tag{7}$$

In both the cases, we may define an *observation operator* $\mathscr{H}$ mapping the unknown function u (or u and η) to the place where the observations are made. Thus, $\mathscr{H}(u)_{j,k} = v(x_j, t_k)$ in the Eulerian case and $\mathscr{H}(u)_{j,k} = z_j(t_k)$ in the Lagrangian case. Then, the data assimilation problem is to find u from y given by

$$y = \mathscr{H}(u) + \sigma. \tag{8}$$

The *observation error* σ is not known to us, but the common assumption is that its statistical properties are known and should be exploited. To this end we introduce the *covariance weighted least-squares function*

$$\Phi(u; y) = \tfrac{1}{2} |\Gamma^{-\frac{1}{2}} (y - \mathscr{H}(u))|^2. \tag{9}$$

Here, $|\cdot|$ denotes the Euclidean norm and hence Φ measures the model-data mismatch, normalized by the scale set by the standard deviations of the observational error.

In the case where the forcing η is to be estimated as well as the initial condition, then we view $v(x_j, t_k)$ (Eulerian case) or $z_j(t_k)$ (Lagrangian case) as functions of u and η and $\mathscr{H}(u)$ is replaced by $\mathscr{H}(u, \eta)$ in (8) and (9).

2.3. Priors

A key to the particular class of MCMC algorithms that we use in this paper is the specification of Gaussian priors on the unknown states. The algorithms we use require us to be able to sample from the prior measure, but do not require explicit construction of the covariance operator, its inverse or its Cholesky factorization.

The prior on the initial condition u will be a mean zero Gaussian with covariance $\mathscr{C}_{\mathrm{u}} = \delta A^{-\alpha}$, where A is the Stokes operator. The numerical experiments use all the values $\alpha = 2.0$ and $\delta = 400$. This means that, under the prior measure, the Fourier coefficients of $\exp(2\pi i k \cdot x)$ in the stream function are independent Gaussians with a standard deviation proportional to $(4\pi^2 |k|^2)^{-1}$. Samples from the prior can thus be constructed in Fourier space, using this specification.

When we also consider estimation of the forcing η, we again use a Gaussian with mean zero and we implicitly define the covariance operator $\mathscr{C}_\eta$ as follows, by describing how to create draws from this prior. These draws are found by solving the following stationary Ornstein–Uhlenbeck (OU) process:

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} + R\eta = \sqrt{\Lambda}\xi, \quad \eta(0) \sim N\left(0, \frac{1}{2} R^{-1} \Lambda\right). \tag{10}$$

Here ξ is space-time white noise and we assume that R and Λ are self-adjoint positive operators diagonalized in the same basis as the Stokes operator A. In practise, this means that draws from the prior can be constructed by solving independent OU processes for each coefficient of the forcing, written in a divergence-free Fourier basis. At each time t, the spatial distribution of η is mean zero

DOI: 10.1002/fld

Gaussian with covariance operator $\frac{1}{2}R^{-1}\Lambda$. The choice of R determines the decorrelation time of different Fourier modes. Thus, by playing with the choices of R and Λ we can match various space-time covariance structures.

For all the experiments in this paper, we choose Λ and R such that the stationary distribution of (10) is the same as for the initial condition: $\mathcal{N}(0, \delta A^{-\alpha})$ with $\alpha = 2.0$ and $\delta = 400$. The operator R is chosen to be proportional to A so that the decorrelation time in the Fourier mode $\exp(2\pi i k \cdot x)$ is inversely proportional to $|k|^2$.

2.4. Bayes theorem and relationship to variational methods

The previous three sections describe a probability model for the joint random variable (u, y) or (u, η, y). Our aim now is to condition this random variable on a fixed instance of the data y. We first describe the situation where only the initial condition u is estimated and then the generalization to estimating (u, η).

The random variable u is Gaussian $N(0, \mathscr{C}_{\mathrm{u}})$. The random variable $y|u$ is Gaussian $N(\mathscr{H}(u), \Gamma)$. Application of the Bayes theorem shows that

$$\frac{\mathbb{P}(u|y)}{\mathbb{P}(u)} \propto \exp(-\Phi(u; y)). \tag{11}$$

The MAP estimator for this problem is simply the minimizer of the functional

$$J(u) := \tfrac{1}{2}\|\mathscr{C}_{\mathrm{u}}^{-\frac{1}{2}} u\|^2 + \Phi(u; y), \tag{12}$$

where $\|\cdot\|$ denotes the $L^2(\mathbb{T}^2)$ norm. Minimizing J is simply the 4DVAR method, formulated in a non-incremental fashion.

When the pair of states (u, η) are to be estimated then the probability model is as follows. The random variable (u, η) is the product of two independent Gaussians, $N(0, \mathscr{C}_{\mathrm{u}})$ for u and $N(0, \mathscr{C}_{\eta})$ for η. The random variable $y|u, \eta$ is Gaussian $N(\mathscr{H}(u, \eta), \Gamma)$. Application of the Bayes theorem shows that

$$\frac{\mathbb{P}(u, \eta|y)}{\mathbb{P}(u, \eta)} \propto \exp(-\Phi(u, \eta; y)). \tag{13}$$

The MAP estimator for this problem is simply the minimizer of the functional

$$J(u) := \tfrac{1}{2}\|\mathscr{C}_{\mathrm{u}}^{-\frac{1}{2}} u\|^2 + \tfrac{1}{2}\|\mathscr{C}_{\eta}^{-\frac{1}{2}} \eta\|^2 + \Phi(u, \eta; y), \tag{14}$$

where $\|\cdot\|$ denotes the $L^2(\mathbb{T}^2)$ norm on the u variable and the $L^2([0, T]; L^2(\mathbb{T}^2))$ norm on η. Note that finding the initial condition u and space-time dependent forcing η is equivalent to finding the velocity field v for the duration of the observation window. Minimizing J is simply a weak constraint 4DVAR method, formulated in a non-incremental fashion.

3. THE RANDOM WALK ALGORITHM

If, instead of maximizing the posterior probability, we try and draw multiple realizations of the posterior probability in order to get information about it, this will lead to a form of statistical 4DVAR. We describe a random walk algorithm which does this. We describe it first for estimation of the initial condition alone where we generate $\{u^{(j)}\}$ from $\mathbb{P}(u|y)$ given by (11), starting from $\{u^{(0)}\}$. A key fact to note about the algorithm is that the normalization constant in (11) is not needed.

(1) Set $j = 0$.
(2) Draw $\xi^{(j)}$ from the Gaussian $N(0, \mathscr{C}_{\mathrm{u}})$.
(3) Set $v^{(j)} = (1 - \beta^2)^{\frac{1}{2}} u^{(j)} + \beta \xi^{(j)}$.

DOI: 10.1002/fld

(4) Define $\alpha^{(j)} = \min\{1, \exp(\Phi(u^{(j)}; y) - \Phi(v^{(j)}; y))\}$.
(5) Draw $\gamma^{(j)}$, a uniform random variable on [0,1].
(6) If $\alpha^{(j)} > \gamma^{(j)}$ set $u^{(j+1)} = v^{(j)}$. Otherwise set $u^{(j+1)} = u^{(j)}$.
(7) Set $j \to j+1$ and return to 2.

The parameter $\beta \in [0, 1]$ is a free variable that may be chosen to optimize the rate at which the algorithm explores the space. Note that if the least-squares functional Φ is smaller at the proposed new state $v^{(j)}$ than it is at the current state $u^{(j)}$ then this new state is accepted with probability one. If Φ is larger at the proposed state then the proposed state will be accepted with some probability less than one. Thus, the algorithm tends to favour minimizing the model-data mismatch functional Φ, but in a manner which is statistical, and these statistics are tied to the choice of prior. In this sense, the prior regularizes the problem.

If the target distribution is (13) then the algorithm is the same except that a move in (u, η) space is proposed, from the two Gaussians $N(0, \mathscr{C}_u)$ and $N(0, \mathscr{C}_\eta)$. Then the accept/reject probability $\alpha^{(j)}$ is based on the differences of $\Phi(u, \eta; y)$.

The algorithm draws samples from the desired posterior distribution given by Bayes formula (11) or (13). We refer to this algorithm as Random Walk Metropolis–Hastings (RMWH) but emphasize that this is a non-standard random walk proposal because it is not centred at the current state, but rather at a scaled version of the current state [32]. It is straightforward to implement and involves only repeated solution of the forward model, not an adjoint solver. On the other hand, for probability distributions which are peaked at a single maximum it may be natural to solve the 4DVAR variational problem (12) or (14) first, by adjoint methods, and use this is as a starting point for the above random walk method.

We emphasize that in contrast to variational methods or the extended/ensemble Kalman filter, this method, in principle, computes the entire posterior distribution on the state of the system, given data, with no approximation other than through discretization of the PDE. This distribution is represented through the set of samples of the RWMH method. We reemphasize, however, that this complete information comes at a considerable computational cost, requiring on the order of 10^6 forward model runs. In contrast, variational methods and generalized Kalman filters may only use on the order of 10^2 model runs. Thus, the method we propose can only be used in an offline situation. However, when accurate uncertainty quantification is required, the method we propose constitutes a way of computing the 'ideal solution' and may hence be used to benchmark existing algorithms and to guide the development of new ones.

4. DATA ASSIMILATION OF EULERIAN DATA

Our aim is to sample from the posterior measure (11) using the algorithm from Section 3. In the first two subsections we study the recovery of the initial condition, first with a perfect model and second with an imperfect model. In the third subsection, we study recovery of both the initial condition and the forcing. Note that the posterior measure is Gaussian in this case, because $\mathscr{H}$ is linear, and this has been used to compute explicit solutions to verify the accuracy of the random walk method [36].

In all of the figures we display only the marginal distribution of the Fourier mode $\mathrm{Re}(u_{0,1})$ for the initial condition of the vector field, and $\mathrm{Re}(\eta_{0,1}(0.5))$ for the forcing in the model error case. Other low wave-number Fourier modes behave similarly to this mode. However, high wave-number modes are not greatly influenced by the data, and remain close to their prior distribution. In each case, since we know the value of the Fourier mode that was present in the initial condition of the velocity field that created the data, we plot this value in each example as a vertical black line. In the cases where we are trying to recover an entire time-dependent function, the truth will similarly be given by a black curve.

In each example, since the amount of data we are attempting to assimilate varies, we would expect the optimal value of β in the RWMH method to vary accordingly. Since we do not know this value *a priori*, we can approximate it during the burn-in process. By monitoring the average

DOI: 10.1002/fld

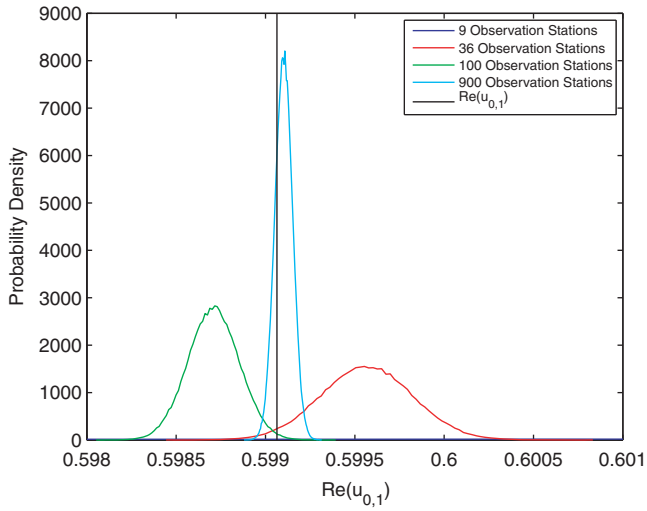

Figure 1. Increasing numbers of observations in space, Eulerian. True value given by black vertical line.

acceptance probability for short bursts of the MCMC method, we can alter β accordingly until we are happy with this acceptance rate. For the purposes of the numerics that follow, β in each case was tuned so that the acceptance rate was approximately 50%. The values of β vary quite widely, but are typically quite small, ranging from 10^{-4} to 10^{-6}.

4.1. Recovering the initial condition

In this section, we assimilate data using a forward model with no forcing: $\eta = 0$. We also use data that is generated by adding noise to output for the forward model, again when $\eta = 0$. In each of the following figures we demonstrate a phenomenon known as *posterior consistency*: as the amount of data is increased the posterior measure becomes more and more sharply peaked on the true value that gave rise to the data. To be precise we demonstrate posterior consistency only for the low wave-number modes: as mentioned above, the data contains little information about high wave numbers and the random walk method returns the prior for these variables. In this subsection, we consider only cases where we do not try to ascertain the model error. In all the examples in this paper, the noise is assumed to be mean zero Gaussian, with covariance matrix $\Gamma = \gamma^2 I$, with $\gamma = 0.01$.

In our first example, the observations are made at 100 evenly spaced times, on an evenly spaced grid with an increasing number of points. Figure 1 shows how the posterior distribution on $\mathrm{Re}(u_{0,1})$ changes as we increase the number of points in space at which we make Eulerian observations of the velocity field. The figure shows converged posterior distributions which resemble Gaussians on $\mathrm{Re}(u_{0,1})$. Note that the curve corresponding to the distribution using data from nine observation stations is so flat on the scale of this graph that the line appears to lie along the x-axis. We also see that we have posterior consistency, since as the number of spatial observations increases, the posterior distribution on $\mathrm{Re}(u_{0,1})$ appears to be converging to an increasingly peaked measure on the value that was present in the true initial condition, denoted by the vertical black line.

We also have results (not included here) showing convergence of the posterior distribution to an increasingly peaked distribution on the correct values as we increase the number of observation times on a fixed time interval, keeping the number of observation stations constant.

4.2. Mismatches in model forcing

Now, we consider the case where we create data with forcing present, but in our algorithm we assume that there is no forcing so that $\eta \equiv 0$. Once again the noise is assumed to be mean zero Gaussian, with covariance matrix $\Gamma = \gamma^2 I$, with $\gamma = 0.01$. We attempt to explain the data arising from

DOI: 10.1002/fld

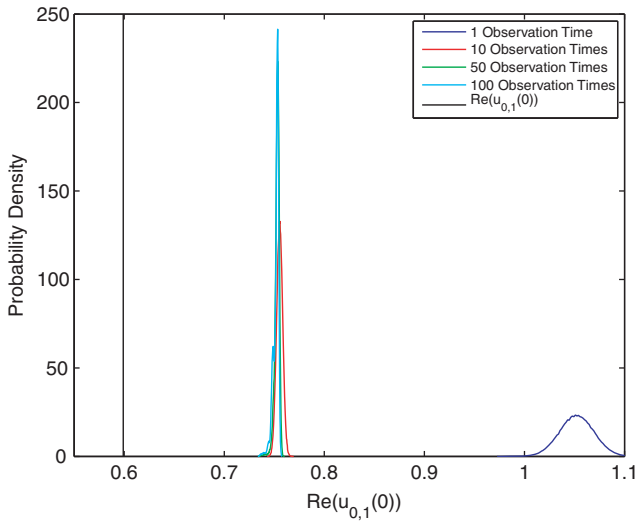

Figure 2. $\mathrm{Re}(u_{0,1}(t))$: increasing number of observation times, unmatched low frequency forcing in data and model, Eulerian data. True value given by black vertical line.

a forced model through the initial condition for an unforced model. Figure 2 shows the marginal distributions of one Fourier mode in such a situation, where the forcing is low frequency in space and constant in time, and where we steadily increase the number of observation times, with 100 observations fixed on a grid. The forcing in this instance is positive for the wave number displayed in the figure. Two things are noteworthy: (i) the posterior tends towards a peaked distribution as the amount of data increases; (ii) this peak is not located at the true initial condition (marked with a black line). This incorrect estimate of the initial condition is, of course, because of the mismatch between model used for the assimilation and for the data generation. In particular, the energy in the posterior on the initial condition is increased in an attempt to compensate for the model error in the forcing.

Further to this, we consider the similar problem of studying the effect of using a forward model without forcing, when the data is generated with forcing, this time with data created using a stationary solution of the OU process (10). This forcing function is then fluctuating constantly in all Fourier modes.

Figure 3 shows the marginal distribution for one low-frequency Fourier mode with data from one observation time, with an increasing number of observations in space. Notice that once again the distributions are becoming more and more peaked as the number of observations increases, but that the peaks of these distributions are centred away from the value of the Fourier mode that was present in the actual initial condition that created the data.

4.3. Quantifying model error

The previous section shows us that if we assume that our model accurately reflects the real dynamical system generating the data, but in fact it does not do so, then our estimates for the initial condition of the system will be flawed and inaccurate. This motivates us to try to recover not only the initial condition of the system, but also the model error forcing η.

We first observe that, in the Eulerian case, parts of the forcing are unobservable and cannot be determined by the data. To state this precisely, we define

$$F(t_1,t_2)=\int_{t_1}^{t_2}\exp(-\nu A(t_2-t_1))\eta(t)\,\mathrm{d}t.$$

DOI: 10.1002/fld

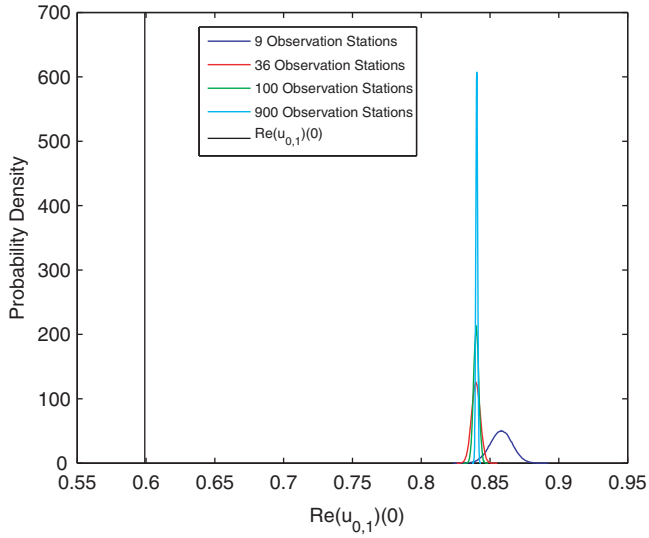

Figure 3. Re($u_{0,1}(t)$): increasing number of observation times, unmatched forcing in data and model, Eulerian data, forcing given by a stationary solution of OU process. True value given by black vertical line.

This is the term in the solution map of the Stokes equations that is dependent on the forcing term η. If the observations are made at times $\{t_i\}_{i=1}^{K}$ then the data is informative about $\tilde{F} := \{F(t_j, t_{j-1})\}_{j=0}^{K-1}$, rather than about η itself. Proof of this is given in [36], by showing that many different (in fact infinitely many) functions η give rise to the same vector $\tilde{F}$, and therefore to the same velocity field at each of the observation times. Because of this, we expect that the posterior measure will give much greater certainty to estimates of $\tilde{F}$ than η.

This basic analytical fact is manifest in the numerical results which we now describe. First, there are infinitely many functions compatible with the observed data so that obtaining a converged posterior on η is a computationally arduous task; second the prior measure plays a significant role in weighting the many possible forcing functions which can explain the data.

As in the non-model error case, we look at a set of experiments where the observations are made at 100 evenly spaced times, with the observation stations evenly spaced on a grid with an increasing number of points. Once again, as we increase the number of observation stations, we are able to recover the value of any given Fourier mode in the initial condition with increasing accuracy and certainty as we increase the amount of data. We omit the majority of the graphical representations of this fact and concentrate instead mainly on the posterior distribution on the forcing function η. Once again the noise is assumed to be mean zero Gaussian, with covariance matrix $\Gamma = \gamma^2 I$, with $\gamma = 0.01$.

Figures 4(a) and (b) show the marginal posterior distributions on Re($\eta_{0,1}(0.5)$) (the real part of the $(0, 1)$ Fourier coefficient of the forcing at time $t = 0.5$) and on Re($F_{0,1}(0.5)$), respectively, given an increasing number of observation stations. The first figure shows that even with a large amount of data the standard deviation about the posterior mean for Re($\eta_{0,1}(0.5)$) is comparable in magnitude to the posterior mean itself. In contrast, for Re($F_{0,1}(0.5)$) and for a large amount of data, the standard deviation around the posterior mean is an order of magnitude smaller than the mean value itself. The data is hence much more informative about $\tilde{F}$ than it is about η, as predicted.

We pursue this further by comparing samples from the Markov chain used to explore the posterior distribution, with the empirical average of that chain. Figure 5(a) shows an example trace of the value of Re($\eta_{0,1}(0.5)$) in the Markov chain (fluctuating), together with its empirical average (smoother). The slow random walk of the sample path, around the empirical mean, is the hallmark of an unconverged Markov chain. In contrast, Figure 5(b) shows how well the value

DOI: 10.1002/fld

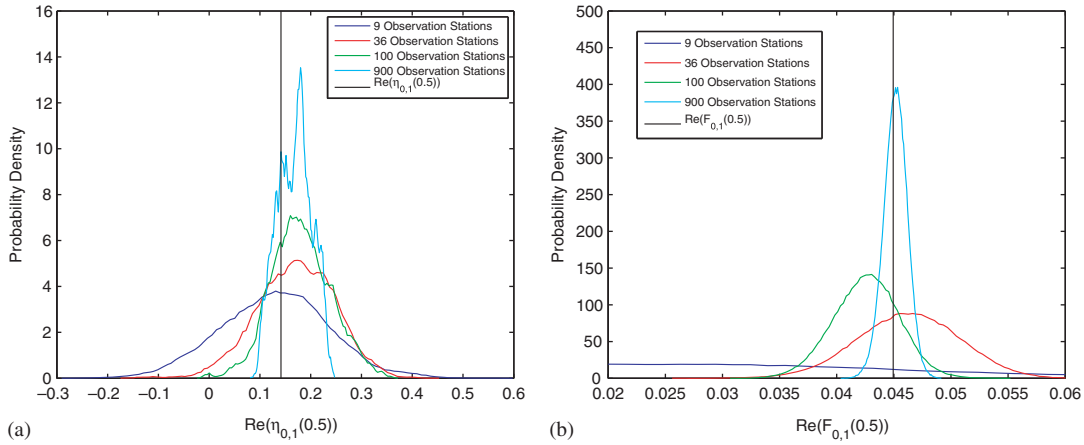

Figure 4. Increasing numbers of observations in space, Eulerian model error case. (a) $\mathrm{Re}(\eta_{0,1}(0.5))$, (b) $\mathrm{Re}(F_{0,1}(0.5))$. Note much smaller variances in the distributions in (b) in comparison with those in (a). True values given by black vertical lines.

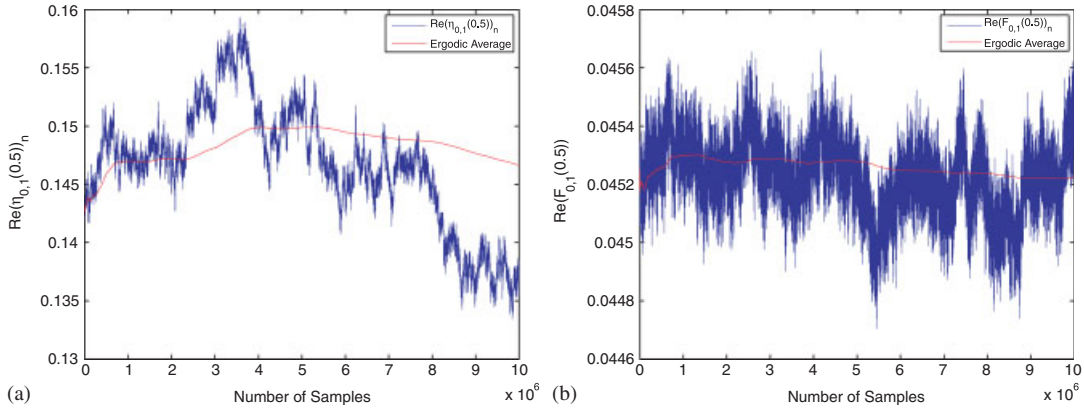

Figure 5. Value of (a) $\mathrm{Re}(\eta_{0,1}(0.5))$ and (b) $\mathrm{Re}(F_{0,1}(0.5))$ in the Markov chain.

of $\mathrm{Re}(F_{0,1}(0.5))$ converges in distribution in the Markov chain. The whole trace stays within a reasonably tight band around the empirical mean, fluctuating in a white fashion; it does not display the slow random walk behaviour of Figure 5(a).

Figure 6 shows the expectation of the entire function $\mathrm{Re}(F_{0,1}(t))$, given an increasing number of points in the vector field to be observed. As the number of observations increases, the expectation of $\mathrm{Re}(F_{0,1}(t))$ nears the true value, given by the black curve.

If we now look at the posterior mean of the initial condition with a varying number of observation stations, and compare this to the true initial condition, we get an error curve as shown in Figure 7. This shows that as the number of observation stations increases in a sensible way (for example using a space-filling algorithm), the posterior mean of the initial condition converges to the true initial condition.

Similarly, if we look at the $L^2([0,T];L^2(\mathbb{T}^2))$-norm of the difference between the posterior mean of the time-dependent forcing, and the true forcing that created the data, we get an error curve as shown in Figure 8(a). This shows that as the number of observation stations increases in a sensible way that the posterior mean of the time-dependent forcing converges to the true forcing. Notice however, that the convergence is slower in this quantity than in that given in Figure 8(b), which shows the norm of the difference between the posterior mean of F and the true forcing F. Again, this shows that as we increase the number of observation stations, the posterior mean

DOI: 10.1002/fld

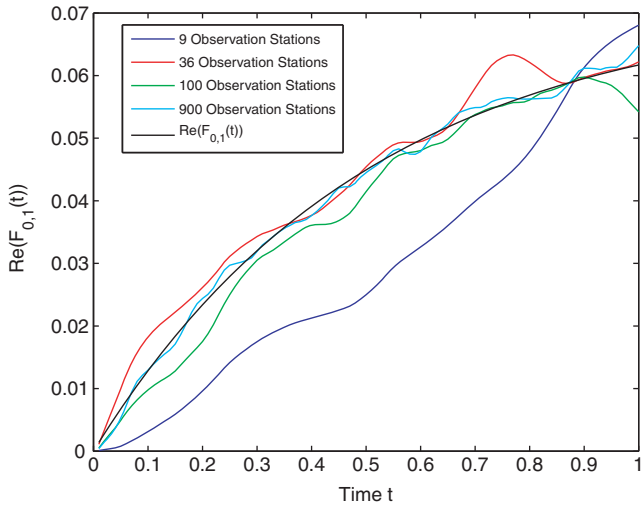

Figure 6. $\mathrm{Re}(F_{0,1}(t))$: increasing numbers of observations in space, Eulerian model error case. True value given by black curve.

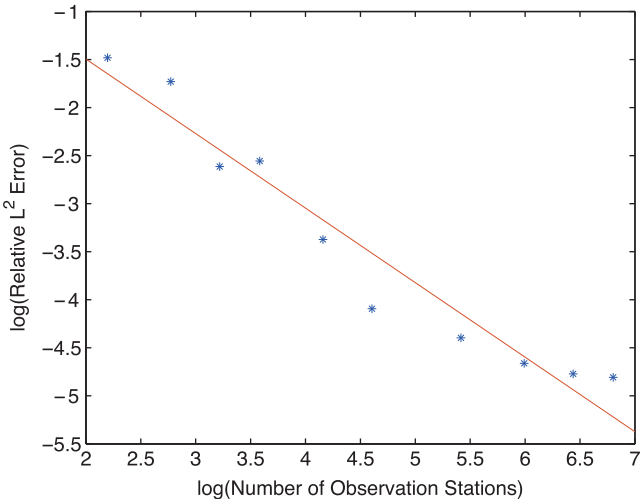

Figure 7. $\|\mathbb{E}(u)-u_{\mathrm{Act}}\|_{L^2}$: increasing numbers of observations in space, Eulerian model error case. u_{Act} is actual initial condition that created the data.

converges to the true answer. The convergence of this quantity (gradient ≈ -0.323) is much quicker than that of η (gradient ≈ -0.216), but slower than that of u (gradient ≈ -0.778). Note that these gradients are reliant on the way in which the amount of data is increased, where in this case the observation stations were placed on an increasingly refined grid.

We also have analogous results (given in [36]) for the case where we increase the number of observations in time, but keep the number of observation stations the same.

We may also be interested in understanding how well we are able to characterize high-frequency (in space) forcing from Eulerian data. In the following experiment, all Fourier modes in the forcing that created the data were set to zero, apart from two high-frequency modes for $k=(5,5)$ and $k=(4,5)$. An increasing number of observation stations were placed on a grid, with observations made at 100 evenly spaced times up to $T=1$. Figure 9 shows the mean forcing function (an average over all the realizations in the Markov chain) for the Fourier mode $\mathrm{Re}(\eta_{5,5}(t))$, as a function of time. The actual value that was present in the forcing that created the data is indicated by the solid

DOI: 10.1002/fld

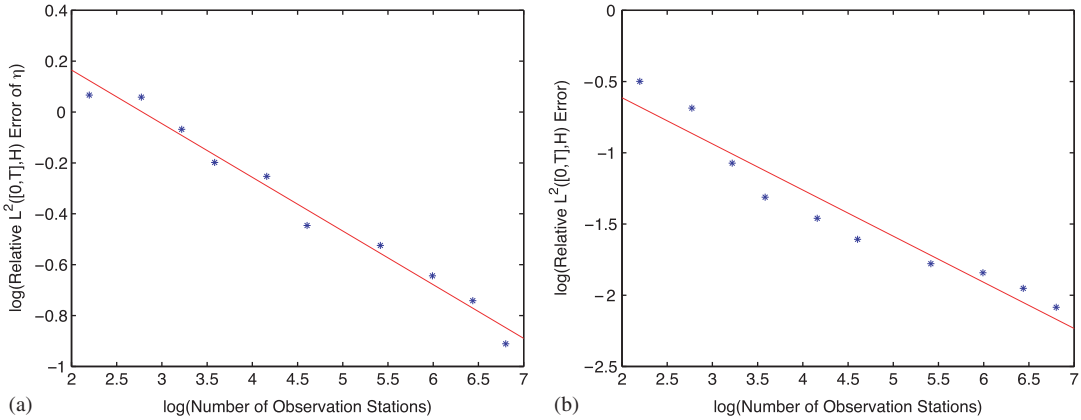

Figure 8. Increasing numbers of observations in space, Eulerian model error case, (a) $\|\mathbb{E}(\eta)-\eta_{\mathrm{Act}}\|_{L^2(0,T;H)}$ (b) $\|\mathbb{E}(F)-F_{\mathrm{Act}}\|_{L^2(0,T;H)}$. η_{Act} and F_{Act} are the actual forcing functions that created the data.

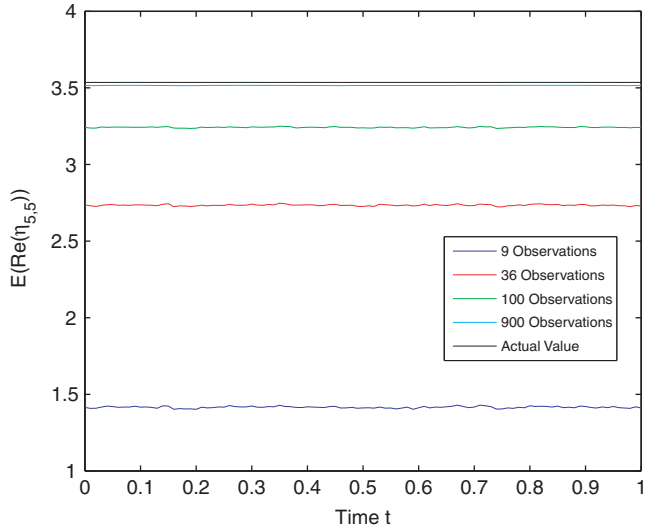

Figure 9. $\mathrm{Re}(\eta_{5,5}(t))$: increasing numbers of observations in space, Eulerian model error case, high frequency forcing. True value given by black line.

line. Figure 10 shows the absolute value of the difference between the mean function and the true value that created the data.

In each Fourier mode, the more observations in space that are assimilated, the better the estimate. Moreover, the variance in these estimates (omitted here) reduces as the amount of information increases, leading to peaked distributions on the forcing function that created the data.

Note that high-frequency modes require more spatial observations to be able to make accurate estimates than the low-frequency modes. This is simply due to the fact that with few spatial observations, no matter how many observations we have in time, our information about the high-frequency Fourier modes is under-determined, and aliasing leads to a great deal of uncertainty about these modes.

So far we have only considered examples where the model error forcing that created the data is constant in time. In the following experiment, we take a draw from the model error prior (10) as the forcing function that is used in the creation of the data. This means that we have non-zero forcing in all of the Fourier modes, and this value is constantly changing at each time step also.

DOI: 10.1002/fld

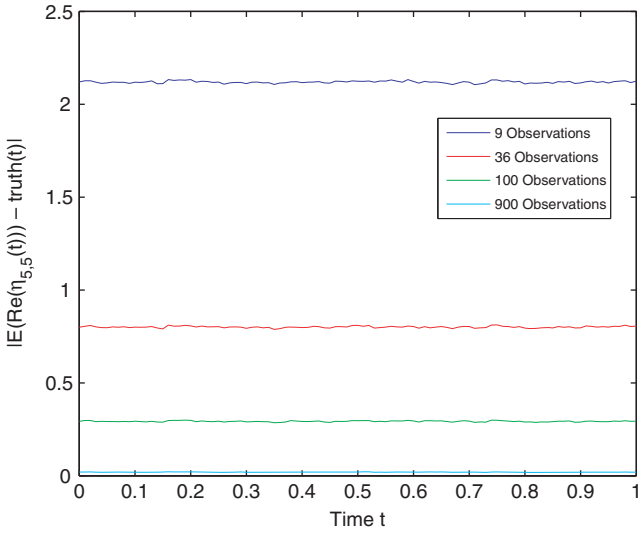

Figure 10. $\mathrm{Re}(\eta_{5,5}(t))$: absolute value of the difference between the mean and truth, Eulerian model error case, high frequency forcing.

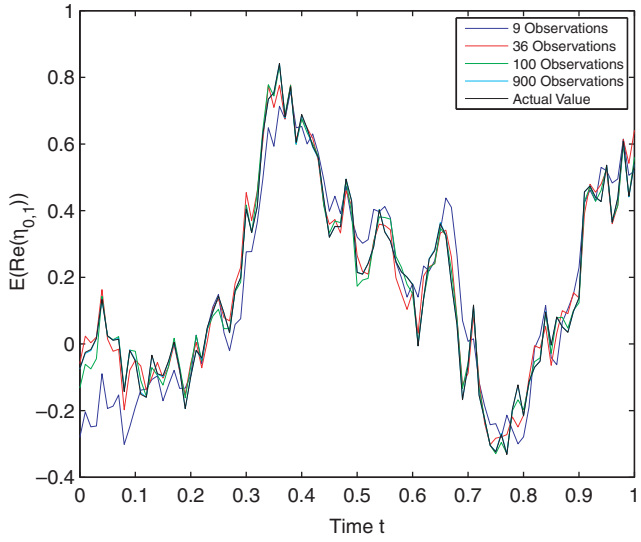

Figure 11. $\mathrm{Re}(\eta_{0,1}(t))$: increasing numbers of observations in space, Eulerian model error case, forcing function taken from prior. True value given by black curve.

Figure 11 shows the estimates (with varying numbers of spatial observations with a fixed number of 100 observations in time) of the forcing function for a particular Fourier mode, along with the actual forcing function that created the data. In Figure 12 the absolute value of the difference between the estimates and the function that created the data are plotted.

These graphs show that as we increase the number of spatial observations, the estimates of the forcing functions converge to the function which created the data.

5. LAGRANGIAN DATA ASSIMILATION

Now our aim is to sample from the posterior measure (13) using the algorithm from Section 3. Unlike the Eulerian case, the posterior measure is not Gaussian, because $\mathscr{H}$ is non-linear, and so

DOI: 10.1002/fld

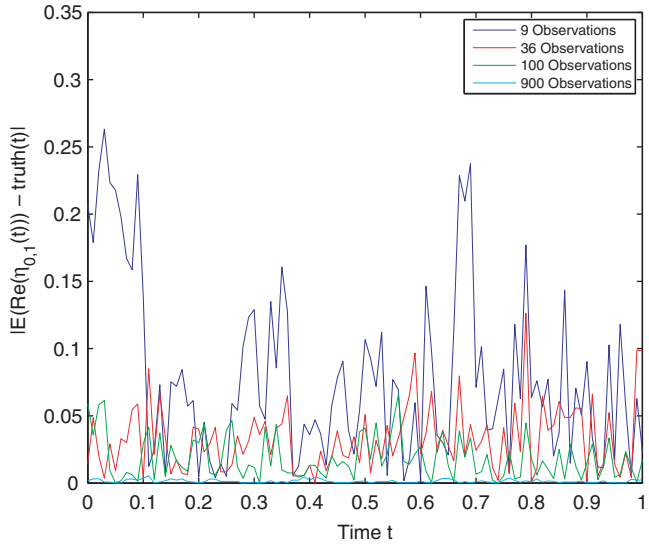

Figure 12. Re($\eta_{0,1}(t)$): absolute value of the difference between the mean and truth, Eulerian model error case, forcing function taken from prior.

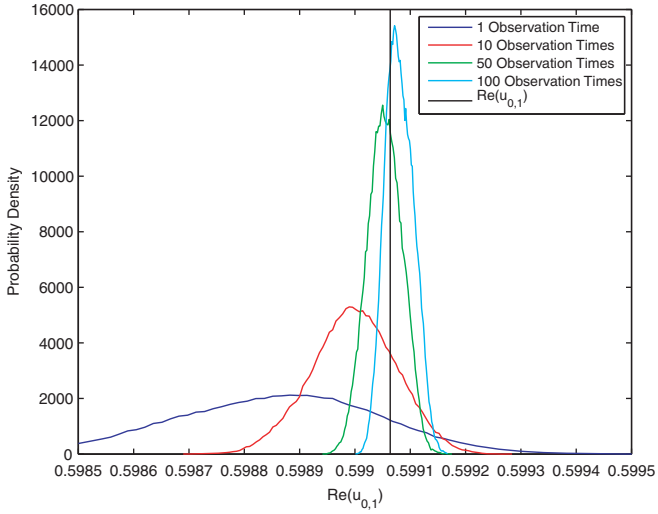

Figure 13. Increasing numbers of observations in time, Lagrangian. True value given by black vertical line.

the posterior cannot be computed simply by means of linear algebra. In the first two subsections, we study recovery of the initial condition, first with a perfect model and second with an imperfect model. In the third subsection, we study recovery of both the initial condition and the forcing.

5.1. Recovering the initial condition

We consider an example where we have a fixed number of 25 tracers, whose initial positions are evenly spaced on a grid, and assumed to be known. We observe each tracer at an increasing number of times evenly spaced on the unit interval, as we have previously done in the Eulerian case, and attempt to recover the initial condition of the fluid. Once again the noise is assumed to be mean zero Gaussian, with covariance matrix $\Gamma = \gamma^2 I$, with $\gamma = 0.01$. Figure 13 shows that the marginal distributions for Re($u_{0,1}$) have converged to approximate Gaussian distributions, and

DOI: 10.1002/fld

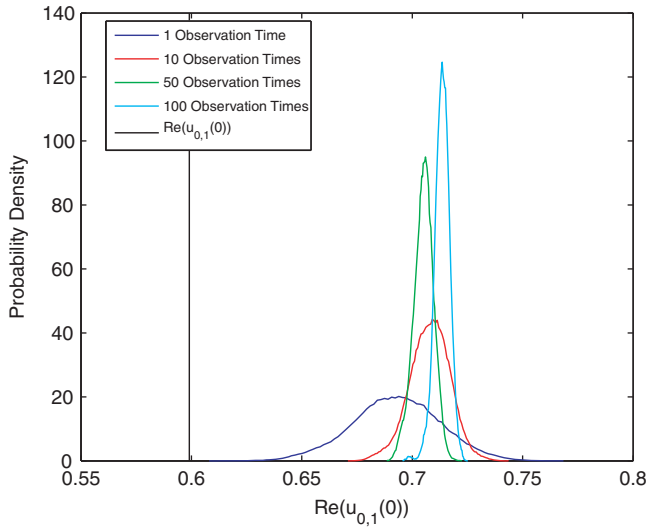

Figure 14. $\mathrm{Re}(u_{0,1}(t))$: increasing number of observation times, unmatched forcing in data and model, Lagrangian data, constant low frequency forcing used in the data creation. True value given by black vertical line.

exhibits posterior consistency as the number of temporal observations increases. We also have similar results showing posterior consistency in the case that we have a fixed observation time and an increasing number of tracers whose initial positions are evenly spaced on a grid (given in [36]).

5.2. *Mismatches in model forcing*

As in Section 4.2, we now consider the case where the model in our statistical algorithm does not reflect the dynamical system from which we are making our observations. We attempt to explain the data arising from a forced model through the initial condition for an unforced model. Once again the noise is assumed to be mean zero Gaussian, with covariance matrix $\Gamma = \gamma^2 I$, with $\gamma = 0.01$. Figure 14 shows the marginal distributions of one Fourier mode in such a situation where we introduce low-frequency constant forcing in the data creation process, but set $\eta \equiv 0$ in the statistical algorithm. Again, we consider Lagrangian data in the case where we steadily increase the number of observation times. As in the Eulerian example, two things are noteworthy: (i) the posterior tends towards a peaked distribution as the amount of data increases; (ii) this peak is not located at the true initial condition (marked with a black line). This incorrect estimate of the initial condition is due to the mismatch between model used for the assimilation and for the data generation. In particular the energy in the posterior on the initial condition is increased in an attempt to compensate for the model error in the forcing.

Lagrangian data is much more sensitive to small changes in the model error forcing than the equivalent Eulerian data, due to particles' positions being dependent on the entire history of the velocity field from the initial time. Therefore, creating more complex incongruities between the data creation schemes and the model used within the statistical algorithm, as we did in Section 4.2, can cause more serious problems. To this end, the RWMH algorithm simply failed to converge in several experiments of this type. A random walk method is clearly inadequate for exploring these types of highly complex probability densities. The assumed amount of observational noise can be altered to allow freer exploration of the state space, but finding a value which simultaneously does not simply return a trivial solution (the prior distribution) or does not converge in a reasonable amount of time due to the complexity of the posterior distribution can be challenging. It might be instructive to implement a gradient-type method, or other MCMC methods, in these more challenging scenarios to better understand these types of problems.

DOI: 10.1002/fld

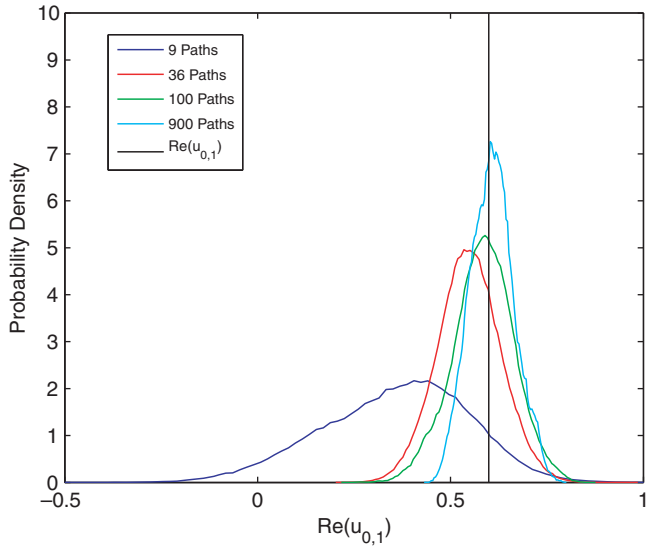

Figure 15. Re($u_{0,1}$): increasing numbers of observations in space, Lagrangian model error case. True value given by black vertical line.

5.3. *Quantifying model error*

The problems we experienced in Section 5.2 serve to highlight the need to estimate not only the initial condition, but also the model error forcing. The Lagrangian equivalent of the model error problem is more harder to sample from in comparison with the Eulerian case. This is due to the fact that the position of a passive tracer is dependent upon the entire history of the velocity vector field up to the observation time. Moreover, a small change in the forcing function can result in the small displacement of hyperbolic points in the flow, which in turn can drastically alter the trajectory of one or more tracers. This makes it hard to move in the Lagrangian model error state space within the MCMC method. As a consequence a simple random walk proposal is highly unlikely to be accepted. Indeed, this was borne out in our initial results, which showed that even after very large amounts of samples, the Markov chains were far from converged, as the size of jump in the proposals of the next state that is required to give reasonable acceptance probabilities was simply too small to allow efficient exploration of the state space in a reasonable time.

One way to tackle this is to alter the likelihood functional (which calculates the relative likelihood that a given choice of functions created the observations) to allow freer exploration of the state space. The data was created with observational noise with variance $\Gamma = \gamma_1^2 I$, but the assimilation is performed with $\Gamma = \gamma_2^2 I$ and $\gamma_2 \gg \gamma_1$. Then, the acceptance probabilities increase, allowing larger steps in the proposal to be accepted more of the time. The results that follow in this section use $\gamma_1^2 = 10^{-4}$ and $\gamma_2^2 = 25$. We will show that, despite this large disparity, it is possible to obtain reasonable estimates of the true forcing and initial condition. In particular, for large data sets, the posterior mean of these functions is close to the true values that generated the data. However, because γ_2/γ_1 is large, the variance around the mean is much larger than in the previous sections where $\gamma_1 = \gamma_2$.

Equivalently to the Eulerian case, we first consider the scenario where we have 100 equally spaced observation times up to time $T = 1$, at which we observe the passive tracers, whose initial positions are given on a grid with an increasing number of points. Figure 15 shows how the marginal distributions on Re($u_{0,1}(0)$) change as we increase the number of tracers to be observed. This figure indicates that as the number of spatial observations is increased in a sensible way, the marginal distribution on this particular Fourier mode is converging to an increasingly peaked distribution on the true value that created the data.

DOI: 10.1002/fld

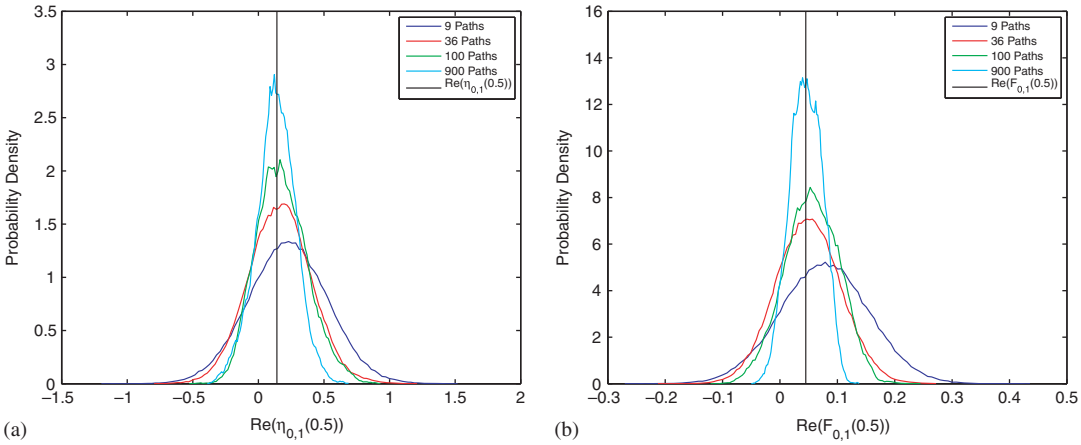

Figure 16. Increasing numbers of observations in space, Lagrangian model error case, (a) Re($\eta_{0,1}(0.5)$) (b) Re($F_{0,1}(0.5)$). True values given by black vertical lines.

Figure 16(a) shows how the marginal distribution on Re($\eta_{0,1}(0.5)$) changes as we increase the number of paths to be observed. In comparison to the Eulerian case, we are able to determine much more about the pointwise value of the forcing function, here using the idea of inflated observational noise variance in the statistical model. Notice that, as in the Eulerian case, the uncertainty in the pointwise value of the forcing is far greater than that for the initial condition of the dynamical system.

Figure 16(b) shows distributions of Re($F_{0,1}(0.5)$) and demonstrates convergence to a sharply peaked distribution on the true value that created the data in the limit of large data sets. The uncertainty in this quantity is less than for the pointwise value of the forcing, as in the Eulerian case, but the discrepancy is considerably less than in the Eulerian case.

Figure 17 shows the expectation of the entire function Re($F_{0,1}(t)$), given an increasing number of points in the vector field to be observed. As the number of paths to be observed increases, the approximation does improve in places. However, since we altered the likelihood to make it possible to sample from this distribution, this also vastly increased the variance in each of the marginal distributions as there is relatively more influence from the prior. Therefore, the expectation of this function does not tell the whole story. This picture does show us however that it is certainly possible to get ballpark estimates for these functions given Lagrangian data.

We also have similar results (given in [36]) for when a fixed number of tracers are observed at an increasing number of observation times on a fixed interval.

6. CONCLUSIONS

We have argued that, in situations where uncertainty quantification is important, the use of an MCMC method to fully explore the posterior distribution of state given data is both worthwhile and increasingly viable. We have investigated in some details the properties of the posterior measure for simplified models of data assimilation in fluid mechanics, for both Eulerian and Lagrangian data types, with and without model error. Through this work, we were able to say more about what kind of information we can garner from these different data types in different scenarios. Although only indicative, being a markedly simpler model than any model that is currently used in practical meteorological and oceanographical scenarios, it still gives us a better idea of how to attack these problems, and of how much information is contained within different types of noisy observation in them. Furthermore, in the current computational practise various simplifications are used—primarily filtering or variational methods [33]—and the posterior measure that we compute in this paper could be viewed as an 'ideal solution' against which these simpler and more practical

DOI: 10.1002/fld

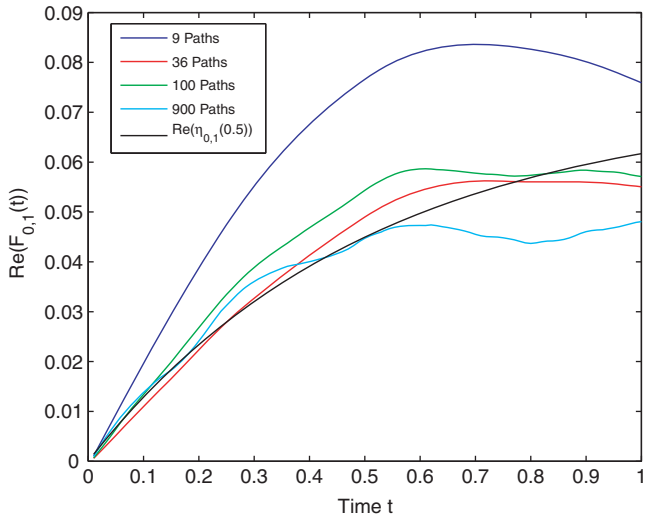

Figure 17. Re($F_{0,1}(t)$): increasing numbers of observations in space, Lagrangian model error case. True value given by black curve.

methods can be compared. To that end it would be of interest to apply the ideas in this paper to a number of simple model problems from data assimilation in meteorology or oceanography and to compare the results from filtering and variational methods with the ideal solution found from MCMC methods.

We have also shown that in the limit of a large amount of informative data, whether that be created by increasing the number of observation times, or by increasing the number of tracers being observed, or the number of observation stations in the Eulerian case that the posterior distribution converges to an increasingly sharply peaked measure on the field/forcing function that created the data, with and without model error. Proving such a result presents an interesting mathematical challenge that will give insight into the large data scenario.

We have also seen certain scenarios in which the random walk method was pushed up to, and beyond its boundaries of efficacy. Implementation of gradient-based methods such as the (Metropolis-Adjusted Langevin) MALA algorithm, as shown in [33], is also of great interest, and may be useful in tackling these more complex posterior densities.

There are also many other application domains to which the algorithmic methodology developed here could be applied and another future direction will be to undertake similar studies to those here for problems arising in other application areas such as subsurface geophysics and applications in image processing such as shape registration.

ACKNOWLEDGEMENTS

A. M. S. is grateful to EPSRC, ERC and ONR for financial support. M. D. was supported by the Warwick Postgraduate Fellowship fund and, as a postdoc, by the ERC. The research of S. L. C. has received a postgraduate funding from EPSRC, and a postdoctoral funding from the ERC (under the *European Community's* Seventh Framework Programme (FP7/2007-2013)/ ERC *grant agreement* No. 239870) and from Award No KUK-C1-013-04, made by the King Abdullah University of Science and Technology (KAUST).

REFERENCES

1. Kalnay E. *Atmospheric Modeling*, *Data Assimilation and Predictability*. Cambridge University Press: Cambridge, 2003.
2. Courtier P, Anderson E, Heckley W, Pailleux J, Vasiljevic D, Hamrud M, Hollingworth A, Rabier F, Fisher M. The ECMWF implementation of three-dimensional variational assimilation (3d-var). *Quarterly Journal of the Royal Meteorological Society* 1998; **124**:1783–1808.

DOI: 10.1002/fld

3. Bennett AF. *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press: Cambridge, 2002.
4. Courtier P. Dual formulation of variational assimilation. *Quarterly Journal of the Royal Meteorological Society* 1997; **123**:2449–2461.
5. Courtier P, Talagrand O. Variational assimilation of meteorological observations with the adjoint vorticity equation. ii: numerical results. *Quarterly Journal of the Royal Meteorological Society* 1987; **113**:1329–1347.
6. Ide K, Kuznetsov L, Jones CKRT. Lagrangian data assimilation for point vortex systems. *Journal of Turbulence* 2002; **3**:53.
7. Kuznetsov L, Ide K, Jones CKRT. A method for assimilation of Lagrangian data. *Monthly Weather Review* 2003; **131**:2247.
8. Jardak M, Navon IM, Zupanski M. Comparison of sequential data assimilation methods for the Kuramoto-Sivashinsky equation. *International Journal for Numerical Methods in Fluids* 2010; **62**:374–402.
9. Derber JC. A variational continuous assimilation technique. *Monthly Weather Review* 1989; **117**:2437–2446.
10. Fang F, Pain CC, Navon IM, Piggott MD, Gorman GJ, Farrell PE, Allison AJH, Goddard PA. A POD reduced-order 4D-Var adaptive mesh ocean modelling approach. *International Journal for Numerical Methods in Fluids* 2009; **60**:709–732.
11. Nodet M. Variational assimilation of lagrangian data in oceanography. *Inverse Problems* 2006; **22**:245–263.
12. Evensen G. *Data Assimilation*: *the Ensemble Kalman Filter*. Springer: New York, 2007.
13. Apte A, Jones CKRT, Stuart AM, Voss J. Data assimilation: mathematical and statistical perspectives. *International Journal for Numerical Methods in Fluids* 2008; **56**:1033–1046.
14. Stuart AM. Inverse problems: a Bayesian approach. *Acta Numerica* 2010; **19**:451–559.
15. Kaipio J, Somersalo E. *Statistical and Computational Inverse Problems*. Springer: Berlin, 2004.
16. Liu J. Monte Carlo strategies in scientific computing. *Springer Texts in Statistics*. Springer: New York, 2001.
17. Robert CP, Casella GC. Monte Carlo statistical methods. *Springer Texts in Statistics*. Springer: Berlin, 1999.
18. Apte A, Jones CKRT, Stuart AM. A Bayesian approach to Lagrangian data assimilation. *Tellus* 2007; **60**(2008): 336–347.
19. Herbei R, McKeague IW. Hybrid samplers for ill-posed inverse problems. *Scandinavian Journal of Statistics* 2009; **36**(4):839–853.
20. Herbei R, McKeague IW, Speer K. Gyres and jets: inversion of tracer data for ocean circulation structure. *Journal of Physical Oceanography* 2008; **38**:1180–1202.
21. McLaughlin D, Townley LR. A reassessment of the groundwater inverse problem. *Water Resources Research* 1996; **32**(5):1131–1161.
22. McKeague IW, Nicholls G, Speer K, Herbei R. Statistical inversion of south atlantic circulation in an abyssal neutral density layer. *Journal of Marine Research* 2005; **63**:683–704.
23. Michalak AM, Kitanidis PK. A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification. *Water Resources Research* 2003; **39**(2):1033.
24. Calvetti D, Somersalo E. Large-scale statistical parameter estimation in complex systems with an application to metabolic models. *Multiscale Modeling and Simulation* 2008; **5**:1333–1366.
25. Dostert P, Efendiev Y, Hou TY, Luo W. Coarse-grain Langevin algorithms for dynamic data integration. *Journal of Computational Physics* 2006; **217**:123–142.
26. Kaipio JP, Somersalo E. Statistical inversion and monte carlo sampling methods in electrical impedance tomography. *Inverse Problems* 2000; **16**:1487–1522.
27. Kaipio JP, Somersalo E. Statistical inverse problems: discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics* 2007; **198**:493–504.
28. Mosegaard K, Tarantola A. Monte carlo sampling of solutions to inverse problems. *Journal of Geophysical Research* 1995; **100**:431–447.
29. Metropolis N, Rosenbluth RW, Teller MN, Teller E. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 1953; **21**:1087–1092.
30. Beskos A, Stuart A. MCMC methods for sampling function space. *ICIAM 07—6th International Congress on Industrial and Applied Mathematics*. European Mathematical Society: Zürich, 2009; 337–364.
31. Beskos A, Stuart A. Computational complexity of Metropolis–Hastings methods in high dimensions. In *Monte Carlo and Quasi-Monte Carlo Methods 2008*, L'Ecuyer P, Owen AB (eds). Springer: Berlin, Heidelberg, 2009; 61–71.
32. Cotter SL, Roberts GO, Stuart AM, White D. MCMC methods for functions: modifying old algorithms to make them faster. In preparation, 2010.
33. Cotter SL, Dashti M, Robinson JC, Stuart AM. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems* 2009; **25**:115008.
34. Cotter SL, Dashti M, Stuart AM. Approximation of Bayesian inverse problems for PDEs. *SIAM Journal on Numerical Analysis* 2010; **48**(1):322–345.
35. Temam R. Navier–Stokes equations and nonlinear functional analysis. *Regional Conference Series in Applied Mathematics*. SIAM: Philadelphia, 1983.
36. Cotter SL. Applications of MCMC Sampling on Function Spaces. *Ph.D. Thesis*, Warwick University, 2010.

DOI: 10.1002/fld