



Relevance of conservative numerical schemes for an Ensemble Kalman Filter

Svetlana Dubinkina*

Centrum Wiskunde & Informatica, P.O. Box 94079, 1090 GB Amsterdam, the Netherlands

*Correspondence to: Svetlana Dubinkina, Centrum Wiskunde & Informatica, P.O. Box 94079, 1090 GB Amsterdam, the Netherlands.
E-mail: s.dubinkina@cw.nl

In this paper we study relevance of quantities conserved by a numerical scheme for data assimilation through statistical equilibrium mechanics. We use the Ensemble Kalman Filter with perturbed observations as a data assimilation method. We consider three Arakawa discretisations of the quasigeostrophic model that either preserve energy (Arakawa E), or enstrophy (Arakawa Z), or both (Arakawa EZ). We perform a twin experiment, where observations are generated from the Hamiltonian particle-mesh method (HPM), which preserves energy and infinite number of Casimirs though trivially. Due to the chosen initial conditions and conservation laws of the HPM the true probability density function (PDF) is skewed, while due to the conservation laws of an Arakawa discretisation the modelled PDF is normal. Numerical experiments show that if observations of stream function are assimilated, the choice of a numerical scheme is crucial for a good reconstruction of the time-averaged fields and PDF estimation. Arakawa E completely fails to reproduce the true nonlinear behaviour, Arakawa Z is sensitive to localisation and inflation, and Arakawa EZ provides the best estimate. If observations of potential vorticity are assimilated, a good time-averaged field reconstruction is independent of a numerical scheme. The PDF estimations are comparable for all Arakawa discretisations. For obtaining nonzero skewness localisation has to be applied even for a very large ensemble size of 600. Inflation, on the contrary, deteriorates skewness estimation for both small and large ensemble sizes.

Key Words: data assimilation; Ensemble Kalman Filter; conservative numerical discretisations; statistical mechanics; quasigeostrophic flow

Received ...

1. Introduction

Data assimilation is a methodology that combines available observations together with a model state estimation in order to reduce an error in initial conditions and, as a result, to give a better estimation and prediction of the state (Jazwinski 1970). One obstacle of data assimilation is availability of the data. As we cannot observe all variables, the question arises whether data assimilation performs better when assimilating one variable or another. For example Compo *et al.* (2011) showed that assimilation of pressure considerably improves climate estimation. This could be taken into account when planning to install new measuring devices. However, when it comes to estimating the past such as paleoclimate, we need to work with the limited data available to us. Then the question becomes how to improve data assimilation results using those limited data.

One option is to provide better data assimilation methods and another one better model estimations. A better model estimation could be obtained by including missing physical terms, for example. However, as no analytic solution is generally available for complex dynamical models, we need to introduce a numerical

approximation and solve the numerical model. Therefore the numerical approximation introduces an additional error. In order to decrease this error, we could consider a high-order numerical scheme or a scheme that maintains conservation properties of the mathematical (non-discretised) model, since it has been recognised for a long time that preservation of quantities by a numerical scheme avoids spurious behaviour (e.g. Hairer *et al.* 2006). Data assimilation methods, however, are non-conservative in general, with a few exceptions such as Ensemble Kalman Filter without localisation that preserves linear properties (e.g. Evensen 2009b), particle filtering (e.g. Doucet *et al.* 2001), when the data assimilation method does not alternate the state, and recently developed data assimilation methods that ensure conservation laws (Jacobs and Ngodock 2003; Janjić *et al.* 2014). Therefore a combination of a conservative numerical scheme with a typically non-conservative data assimilation method results in a loss of otherwise preserved quantities. Zeng and Janjić (2016) studied how a data assimilation method—the Local Ensemble Transform Kalman Filter—influences conservation laws in the shallow water model that preserves mass, energy and momentum, and in addition

enstrophy for non-divergent flow; Wang *et al.* (2016) investigated influence of assimilation with an ensemble Kalman Filter on variable with a physical constrain for conservation of mass heat and salt content. In this paper we, however, study the influence of conservation properties of different numerical discretisations on the data assimilation results for quasigeostrophic (QG) model through statistical mechanics.

Pioneering work of Onsager (1949) analysed a mathematical (continuous) model from the statistical equilibrium mechanics perspective. The statistical equilibrium mechanics is derived based on conservation laws and it predicts long-time averages (e.g. Majda and Wang 2006). Therefore, a conservative numerical discretisation should be potentially favourable for predictions and thus for applying data assimilation.

We consider the quasigeostrophic (QG) model. The QG model with forcing and dissipation is relevant for practical geophysical flows. A considerable progress has been made in statistical theories over the past years for a continuous model (e.g. Griffa and Salmon 1989; Montgomery *et al.* 1992; DiBattista *et al.* 2001; Verkley and Lynch 2009), but discrete QG models with different numerical treatments of dissipation and forcing terms have not been studied through statistical mechanics. In contrast to forced-dissipative systems, the QG model without forcing nor dissipation is an idealised setting. However, statistical equilibrium theories were developed for both a continuous (Kraichnan 1975; Salmon *et al.* 1976; Carnevale and Frederiksen 1987; Ellis *et al.* 2002) and a discrete form of such QG (Dubinkina and Frank 2007, 2010). As our goal is to analyse data assimilation from the statistical mechanics point of view, we consider the QG model without forcing nor dissipation.

The continuous QG model possess energy and infinite number of Casimir conservation laws. For the continuous QG model, an energy-enstrophy statistical equilibrium theory was derived based on conservation of energy and enstrophy by Kraichnan (1975); Salmon *et al.* (1976); Carnevale and Frederiksen (1987) and a statistical equilibrium theory was derived based on prior probability density function (PDF), conservation of energy and circulation by Ellis *et al.* (2002). Dubinkina and Frank (2007, 2010) showed that the discrete system obtained by a numerical approximation has a different statistical equilibrium theory than the continuous one. Dubinkina and Frank (2007) showed that if the QG model is discretised by Arakawa schemes, then different statistical equilibrium theories can be derived for the corresponding discrete systems depending on the conserved quantities the discrete system possesses. Dubinkina and Frank (2010) showed that the Hamiltonian particle-mesh method for the QG model provides the statistical equilibrium theory based on energy and circulation preservation and the prior PDF, which coincides with conservation of the infinite number of Casimirs, though they are only trivially conserved by the method.

The paper is organised as following: In section 2, we briefly recall the QG model and its conservation properties. In section 3, we review Arakawa discretisations, the Hamiltonian particle-mesh method, and their conservation properties and in section 5, their statistical equilibrium theories. In section 6, we give a description of Ensemble Kalman Filter with perturbed observations. The numerical experiments are presented in section 7 and conclusions are summarised in section 8.

2. The quasigeostrophic model

The quasigeostrophic potential vorticity (QG) equation (e.g. Pedlosky 1987; Salmon 1998; Majda and Wang 2006) describes

barotropic divergence-free flow over topography

$$\frac{d}{dt}q(\mathbf{x}, t) = 0, \quad (1a)$$

$$\Delta\psi(\mathbf{x}, t) = q(\mathbf{x}, t) - h(\mathbf{x}), \quad (1b)$$

where q is the potential vorticity (PV) field, ψ is the stream function, and h is the orography of the earth. The Laplacian operator is denoted by Δ and the material derivative by $d/dt = \partial/\partial t + \mathbf{u} \cdot \nabla$. Here, the divergence-free velocity field \mathbf{u} is related to the stream function by $\mathbf{u} = \nabla^\perp \psi$, where $\nabla^\perp = (-\partial/\partial y, \partial/\partial x)^T$ and T stands for transpose. In this paper we consider the QG equation on a doubly periodic domain $\mathbf{x} = (x, y) \in \mathcal{D} \equiv [0, 2\pi) \times [0, 2\pi)$.

The QG model describes a Hamiltonian PDE with Lie-Poisson structure (Salmon 1998), implying the conservation of the total kinetic energy

$$\mathcal{E} = -\frac{1}{2} \int_{\mathcal{D}} \psi \cdot (q - h) d\mathbf{x}$$

as well as the infinite class of Casimir functionals

$$\mathcal{C}[f] = \int_{\mathcal{D}} f(q) d\mathbf{x}$$

for any function f for which the integral exists. Of particular interest are the PV moments:

$$\mathcal{C}_\alpha = \int_{\mathcal{D}} q^\alpha d\mathbf{x}, \quad \alpha = 1, 2, \dots,$$

and especially the circulation \mathcal{C}_1 and enstrophy \mathcal{C}_2 .

3. Spatial semi-discretisations

In the following section we describe several numerical methods to solve the quasigeostrophic model Eq. (1) in space only.

3.1. Poisson equation

Poisson equation Eq. (1b) is a linear elliptic PDE, which we solve by the Fourier spectral method. Let us introduce a uniform $M \times M$ grid on \mathcal{D} with grid spacing $\Delta x = \Delta y = 2\pi/M$, then the potential vorticity q can be approximated on the grid by $\mathbf{q} \in \mathbb{R}^{M \times M}$ with components $q_{i,j} = q(i\Delta x, j\Delta y)$, $i, j = 0, \dots, M-1$, where periodicity is realised by identifying the indices M and 0 . Define the Fourier transform as

$$\hat{q}_{k,\ell} = \frac{1}{M^2} \sum_{i,j=0}^{M-1} q_{i,j} e^{-i(ki + j\ell)}, \quad k, \ell = -M/2 + 1, \dots, M/2,$$

then Eq. (1b) is solved exactly in the Fourier space

$$\hat{\psi}_{k,\ell} = \begin{cases} 0, & k = \ell = 0, \\ -(\hat{q}_{k,\ell} - \hat{h}_{k,\ell})/(k^2 + \ell^2), & \text{otherwise.} \end{cases}$$

Note, that here the inverse Laplacian is restricted to the hyperplane $\hat{\psi}_{0,0} \equiv 0$, which results in stream function field ψ with zero mean.

In this paper, the orography is a function of x only, specifically

$$h(x, y) = 0.2 \cos(x) + 0.4 \cos(2x),$$

since it is predicted from the energy-enstrophy statistical theory that the mean stream function in this case is a function of x variable only, and therefore any departure from that structure illustrates the departure from the energy-enstrophy statistical theory.

3.2. Arakawa method

Arakawa scheme (Arakawa 1966) is a finite difference scheme for the QG model that preserves linear and quadratic invariants. In order to describe the Arakawa scheme let us rewrite Eq. (1a) as

$$q_t = \mathcal{J}(q, \psi),$$

where the operator \mathcal{J} is defined by

$$\mathcal{J}(q, \psi) = q_x \psi_y - q_y \psi_x.$$

Then we can rewrite the continuous \mathcal{J} in three equivalent forms based on the fact that the derivatives with respect to x and y commute, namely

$$\begin{aligned} \mathcal{J}(q, \psi) &= q_x \psi_y - q_y \psi_x, \\ \mathcal{J}(q, \psi) &= \partial_x(q \psi_y) - \partial_y(q \psi_x), \\ \mathcal{J}(q, \psi) &= \partial_y(q x \psi) - \partial_x(q y \psi). \end{aligned}$$

After discretising each $\mathcal{J}(q, \psi)$ on a uniform $M \times M$ grid on \mathcal{D} with grid spacing $\Delta x = \Delta y = 2\pi/M$ by central differences

$$(\mathbf{D}_x \mathbf{q})_{i,j} = \frac{q_{i+1,j} - q_{i-1,j}}{2\Delta x}, \quad (\mathbf{D}_y \mathbf{q})_{i,j} = \frac{q_{i,j+1} - q_{i,j-1}}{2\Delta y},$$

we get three discrete right-hand sides

$$\begin{aligned} \mathbf{J}_0(\mathbf{q}, \psi) &= (\mathbf{D}_x \mathbf{q}) * (\mathbf{D}_y \psi) - (\mathbf{D}_y \mathbf{q}) * (\mathbf{D}_x \psi), \\ \mathbf{J}_E(\mathbf{q}, \psi) &= \mathbf{D}_x(\mathbf{q} * \mathbf{D}_y \psi) - \mathbf{D}_y(\mathbf{q} * \mathbf{D}_x \psi), \\ \mathbf{J}_Z(\mathbf{q}, \psi) &= \mathbf{D}_y(\psi * \mathbf{D}_x \mathbf{q}) - \mathbf{D}_x(\psi * \mathbf{D}_y \mathbf{q}), \end{aligned}$$

where $*$ denotes the element-wise product of two vectors: $(\mathbf{u} * \mathbf{v})_{i,j} = u_{i,j} v_{i,j}$. The three discrete right-hand sides are not equivalent to each other in contrast to the continuous ones due to the fact that the product rule does not hold in the discrete case. It was noticed by Arakawa that by taking an average of the three discrete right-hand sides we obtain a discretisation \mathbf{J}_{EZ} that conserves discrete analogues of both energy

$$E = -\frac{1}{2} \psi^T (\mathbf{q} - \mathbf{h}) \Delta x \Delta y = \frac{1}{2} \sum_{k,\ell=-M/2+1}^{M/2} (k^2 + \ell^2) |\hat{\psi}_{k,\ell}|^2 \Delta x \Delta y$$

and enstrophy

$$Z = \frac{1}{2} \mathbf{q}^T \mathbf{q} \Delta x \Delta y = \frac{1}{2} \sum_{k,\ell=-M/2+1}^{M/2} |\hat{q}_{k,\ell}|^2 \Delta x \Delta y,$$

where $\mathbf{u}^T \mathbf{v}$ denotes the vector inner product of two vectors \mathbf{u} and \mathbf{v} , while the \mathbf{J}_E discretisation preserves only energy, and the \mathbf{J}_Z discretisation preserves only enstrophy. In reference to their conservation properties, we will refer to the discretisations as the Arakawa EZ , E and Z discretisations, respectively. The Arakawa scheme preserves also the phase space volume element by the flow map.

In all experiments in this paper we choose $M = 23$ and initial conditions such that $E = 7$ and $Z = 20$.

3.3. Hamiltonian particle-mesh method

The Hamiltonian particle-mesh (HPM) method is a numerical discretisation of inviscid fluid dynamics that retains Hamiltonian structure. The method makes use of a Lagrangian fluid description to advect fluid particles while conserving energy and an Eulerian grid for evaluating derivatives using finite differences. The method

was adapted for 2D incompressible flow by Cotter and Reich (2004).

The PV field is discretised by introducing a set of A discrete particles with fixed potential vorticity Q_a , $a = 1, \dots, A$ and time-dependent position $\mathbf{X}_a(t) \in \mathcal{D}$. Then the PV field on a uniform $M \times M$ grid on \mathcal{D} with grid spacing $\Delta x = \Delta y = 2\pi/M$ and grid points denoted by (x_i, y_j) is approximated by

$$q_{i,j}(t) = \sum_{a=1}^A Q_a \phi_{i,j} \{\mathbf{X}_a(t)\}, \quad i, j = 0, \dots, M-1,$$

where $\phi_{i,j}(\mathbf{x}) = \phi\{(x - x_i)/\Delta x\} \phi\{(y - y_j)/\Delta y\}$ is a compactly supported basis function satisfying symmetry, normalisation and partition of unity properties, respectively:

$$\phi(\mathbf{x}) = \phi(-\mathbf{x}), \quad \int_{\mathcal{D}} \phi(\mathbf{x}) d\mathbf{x} = 1, \quad \sum_{i,j=0}^{M-1} \phi_{i,j}(\mathbf{x}) = 1, \quad \forall \mathbf{x} \in \mathcal{D}.$$

In our implementation we use the tensor product of normalised cubic B-splines $\phi(\mathbf{x}) = \phi_0(x) \phi_0(y)$, where

$$\phi_0(r) = \begin{cases} \frac{2}{3} - |r|^2 + \frac{1}{2}|r|^3, & |r| \leq 1, \\ \frac{1}{6}(2 - |r|)^3, & 1 < |r| \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

The particles are advected in a divergence-free velocity field according to

$$\frac{d}{dt} \mathbf{X}_a = \nabla^\perp \psi(\mathbf{x}, t) \big|_{\mathbf{x}=\mathbf{X}_a(t)}, \quad a = 1, \dots, A,$$

where the continuous stream function ψ is obtained by

$$\psi(\mathbf{x}, t) = \sum_{i,j=0}^{M-1} \psi_{i,j}(t) \phi_{i,j}(\mathbf{x}),$$

and the discrete stream function $\psi_{i,j}(t)$ is computed by first solving the Poisson equation Eq. (1b) using the Fourier method described in section 3.1 and then employing the Fourier transformation to obtain the grid-based stream function $\psi_{i,j}(t)$.

The HPM method conserves energy

$$E = -\frac{1}{2} \psi^T (\mathbf{q} - \mathbf{h}) \Delta x \Delta y = \frac{1}{2} \sum_{k,\ell=-M/2+1}^{M/2} (k^2 + \ell^2) |\hat{\psi}_{k,\ell}|^2 \Delta x \Delta y$$

and an infinite number of Casimirs on particles, namely

$$\left(\frac{2\pi}{A}\right)^2 \sum_{a=1}^A Q_a^\alpha, \quad \alpha = 1, 2, \dots$$

However, this property does not transfer to the gridded PV field \mathbf{q} . That is, the grid-based analogues

$$\left(\frac{2\pi}{M}\right)^2 \sum_{i,j=0}^{M-1} q_{i,j}^\alpha, \quad \alpha = 1, 2, \dots$$

are not conserved in general, with the sole exception being the total circulation when $\alpha = 1$. On the other hand, it was shown by Dubinkina and Frank (2010) that the motion of particles is embedded in an area-preserving flow.

Next, let us denote enstrophy on particles as

$$\beta = \frac{1}{2} \left(\frac{2\pi}{A}\right)^2 \sum_{a=1}^A Q_a^2$$

and define a normalised 3rd order Casimir on particles as

$$\gamma = \frac{(2\pi/A)^2 \sum_{a=1}^A Q_a^3}{\beta^{3/2}}. \quad (2)$$

In all experiments in this paper we choose $M = A = 23$ and initial conditions such that $E = 7$, $\beta = 20$ and γ is either 0 or 6.

4. Time discretisations

To integrate the Arakawa and HPM semi-discretisations in time we use the implicit midpoint rule, since this method preserves quadratic quantities and the $2A$ -dimensional phase space of particle positions. The step of the time discretisation $\Delta t_{\text{int}} = 0.1$ for the Arakawa schemes and $\Delta t_{\text{int}} = 0.01$ for the HPM method. In Figure 1 we plot the relative error in energy E and enstrophy Z over time for these discretisations

$$\varepsilon_{\text{rel}}[E](t) = \left| \frac{E(t) - E(0)}{E(0)} \right|, \quad \varepsilon_{\text{rel}}[Z](t) = \left| \frac{Z(t) - Z(0)}{Z(0)} \right|. \quad (3)$$

As it can be observed from Figure 1, energy is well preserved by the Arakawa EZ , Arakawa E and HPM discretisations, and enstrophy is well preserved by the Arakawa EZ and Arakawa Z discretisations.

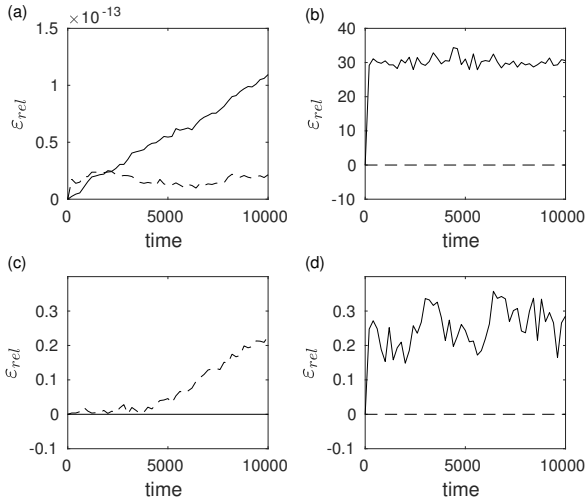


Figure 1. Relative change in energy (dotted line) and enstrophy (solid line) with Arakawa EZ (a), Arakawa E (b), Arakawa Z (c), and HPM (d) discretisations.

5. Statistical mechanics

Given a discrete approximation $\mathbf{q}(t)$ to the solution $q(\mathbf{x}, t)$ of Eq. (1), obtained from a numerical simulation, we would like to analyse the accuracy of computed functions averages of the solution. For example, the long-time average of a function $F(\mathbf{q}(t))$ of the potential vorticity field is denoted

$$\langle F(\mathbf{q}) \rangle = \lim_{T_0 \rightarrow \infty} \frac{1}{T_0} \int_{t_0}^{t_0+T_0} F\{\mathbf{q}(t)\} dt,$$

where t_0 is the time required for decorrelation of the initial condition. If the discrete dynamics is ergodic with respect to a unique invariant measure $\rho(\mathbf{q})$ on the phase space, then the long-time average is equivalent to the ensemble average with respect to ρ ,

$$\langle F(\mathbf{q}) \rangle_\rho = \int F(\mathbf{q}) \rho(\mathbf{q}) d\mathbf{q},$$

where the integral is over the (function) space of potential vorticity fields, and it suffices to derive the invariant measure associated with the numerical method.

Statistical equilibrium theories for explaining the long-time behaviour were derived for Arakawa methods by Dubinkina and Frank (2007) and HPM method by Dubinkina and Frank (2010), where the main findings were that the conserved quantities of a numerical approximation are relevant for the statistical equilibrium. In Figure 2 scatter plots of time-averaged fields for

Arakawa methods are shown for $t_0 = 10^3$ and $T_0 = 10^6$. Both Arakawa E and Z discretisations give rather trivial statistical equilibrium states: $\langle \psi \rangle = 0$ and $\langle \mathbf{q} \rangle = 0$, respectively. Arakawa EZ discretisation, however, gives a linear relation between time-averaged potential vorticity and time-averaged stream function, which is in agreement with the energy-enstrophy statistical theory of Kraichnan (1975); Salmon *et al.* (1976); Carnevale and Frederiksen (1987). Moreover, the energy-enstrophy statistical theory predicts that for the layered topography time-averaged stream function is also layered. This can be observed in Figure 3 for Arakawa EZ discretisation. Arakawa E and Z give $\langle \psi \rangle = 0$ and $\langle \psi \rangle = \mathbf{h}$, respectively.

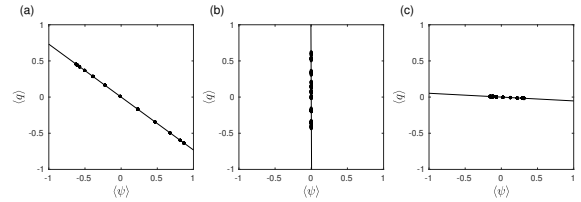


Figure 2. The scatter plot of time-averaged fields $\langle \mathbf{q} \rangle$ vs. $\langle \psi \rangle$ with linear fit for Arakawa EZ (a), Arakawa E (b), and Arakawa Z (c). Time average is computed for $t_0 = 10^3$ and $T_0 = 10^6$.

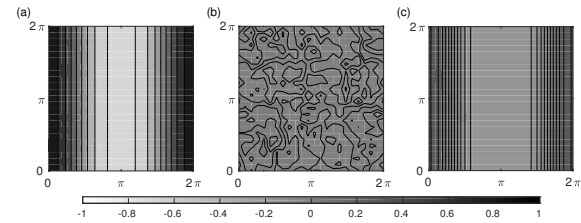


Figure 3. The contour plot of time-averaged stream function $\langle \psi \rangle$ for Arakawa EZ (a), Arakawa E (b), and Arakawa Z (c). Time average is computed for $t_0 = 10^3$ and $T_0 = 10^6$.

Figure 4 shows scatter plot of time-averaged fields for the HPM method for different values of γ from Eq. (2), mainly $\gamma = 0$ and $\gamma = 6$. For $\gamma = 0$ the HPM discretisation exhibits the same long-time statistical behaviour as the Arakawa EZ but for $\gamma = 6$ the behaviour becomes nonlinear instead of linear. This is due to both particular initial conditions—when prior PV on the particles was chosen to be gamma distribution instead of normal distribution—and Casimirs trivially conserved by the particles. Time-averaged stream function shown in Figure 5 is layered for $\gamma = 0$ but is not for $\gamma = 6$, which indicates that the energy-enstrophy statistical theory does not hold for $\gamma > 0$ (Abramov and Majda 2003; Dubinkina and Frank 2010).

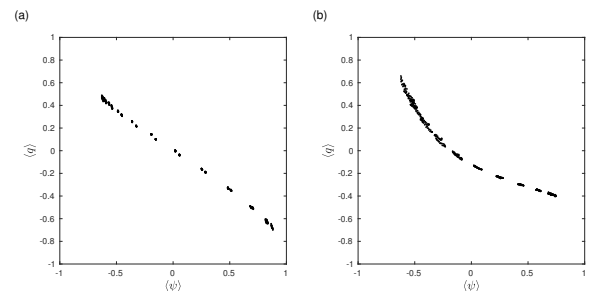


Figure 4. The scatter plot of time-averaged fields $\langle \mathbf{q} \rangle$ vs. $\langle \psi \rangle$ for the HPM method with $\gamma = 0$ (a) and $\gamma = 6$ (b). Time average is computed for $t_0 = 10^3$ and $T_0 = 10^6$.

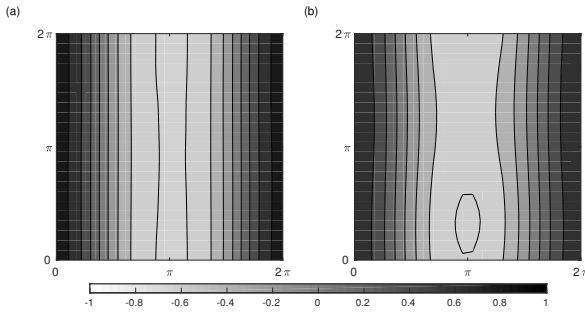


Figure 5. The contour plot of time-averaged stream function $\langle \psi \rangle$ for the HPM method with $\gamma = 0$ (a) and $\gamma = 6$ (b). Time average is computed for $t_0 = 10^3$ and $T_0 = 10^6$.

Note that since any Arakawa scheme preserves only quantities up to second order, choosing initial PV on the grid distributed according to gamma PDF *does not* result in the nonlinear statistical behaviour of long-time averages (not shown as it is identical to Figures 2–3). However, if time averaging is performed on a much shorter time scale, for example $t_0 = 0$ and $T_0 = 10^2$, then the initial gamma-distributed PV leads to nonlinear effects as it can be seen in Figures 6–7. This means that by replacing initial

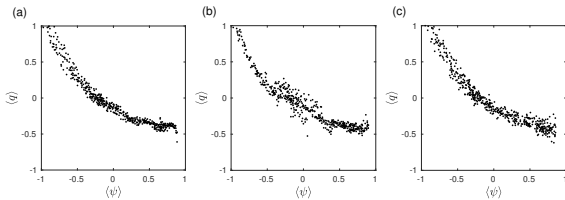


Figure 6. The scatter plot of time-averaged fields $\langle \mathbf{q} \rangle$ vs. $\langle \psi \rangle$ for Arakawa EZ (a), Arakawa E (b), and Arakawa Z (c). Time average is computed for $t_0 = 0$ and $T_0 = 10^2$.

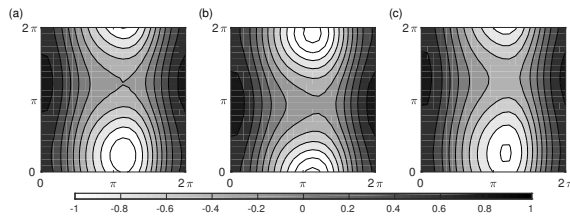


Figure 7. The contour plot of time-averaged stream function $\langle \psi \rangle$ for Arakawa EZ (a), Arakawa E (b), and Arakawa Z (c). Time average is computed for $t_0 = 0$ and $T_0 = 10^2$.

conditions of PV often enough leads to nonlinear behaviour. The goal of data assimilation is to decrease an error in initial conditions by using noisy observations from time to time. However, the results of data assimilation depend on which dynamical variable is observed, observation noise, assimilation interval, etc. and the question that we ask is does the choice of a numerical scheme influence the results of data assimilation?

6. Ensemble Kalman Filter

The Ensemble Kalman filter is a Monte Carlo approximation of the Kalman filter for nonlinear models, which avoids evolving the covariance matrix. The Ensemble Kalman Filter is a widely used data assimilation method (e.g. Evensen 2009a). There are different types of Ensemble Kalman Filter, here we consider an Ensemble Kalman filter with perturbed observations of Burgers *et al.* (1998) which we denote as EnKF.

Denoting a state variable of a nonlinear model at discrete times $n = 0, 1, \dots$ by \mathbf{z}_n we aim at finding an estimate $\tilde{\mathbf{z}}_n$ of \mathbf{z}_n for $n = 1, 2, \dots$ from available observations

$$\mathbf{d}_n = \mathbf{H}\mathbf{z}_n + \boldsymbol{\eta}_n, \quad n = 1, 2, \dots,$$

where \mathbf{H} is the observation operator that projects the model state \mathbf{z}_n into the observation space and $\boldsymbol{\eta}_n$ is an observation error here assumed to be from a normal distribution with mean zero and variance \mathbf{R} , by solving the following set of equations

$$\tilde{\mathbf{z}}_n^e = \mathbf{z}_n^e + \mathbf{K}(\mathbf{d}_n^e - \mathbf{H}\mathbf{z}_n^e), \quad e = 1, \dots, m, \quad (4)$$

$$\mathbf{d}_n^e = \mathbf{d}_n + \boldsymbol{\nu}_n^e, \quad e = 1, \dots, m,$$

$$\mathbf{K} = \mathbf{B}_n \mathbf{H}^T (\mathbf{H} \mathbf{B}_n \mathbf{H}^T + \mathbf{R})^{-1}, \quad (5)$$

$$\mathbf{B}_n = \frac{1}{m-1} (\mathbf{Z}_n - \bar{\mathbf{z}}_n \mathbf{1}_m^T) (\mathbf{Z}_n - \bar{\mathbf{z}}_n \mathbf{1}_m^T)^T.$$

Here $\{\mathbf{z}_n^e\}_{e=1}^m$ is an ensemble of model states obtained by solving the nonlinear model with m different initial conditions, $\{\mathbf{d}_n^e\}_{e=1}^m$ are perturbed observations with $\{\boldsymbol{\nu}_n^e\}_{e=1}^m$ drawn from normal distribution with mean zero and variance \mathbf{R} , \mathbf{Z}_n is a matrix containing each ensemble member \mathbf{z}_n^e as a column, and $\bar{\mathbf{z}}_n$ is the ensemble mean defined as $\bar{\mathbf{z}}_n = 1/m \mathbf{Z}_n \mathbf{1}_m$ with $\mathbf{1}_m$ being an m -dimensional vector with all elements equal to 1. In this paper, the so-called assimilation period, which is the time interval between two consecutive observations, $\Delta t_{\text{assim}} := t_{n+1} - t_n$, is kept constant and the nonlinear model is a discretised QG model with $\mathbf{z} = (\mathbf{q}, \psi)^T$. Since potential vorticity and stream function are obtained on a 23×23 grid, the size of vector \mathbf{z} is equal to 1058.

In an EnKF localisation is often employed in order to remove spurious correlations due to a small ensemble size (Hamill *et al.* 2001). This is done by replacing the covariance matrix \mathbf{B}_n in Eq. (5) by Schur product with a distance-based correlation matrix \mathbf{L}

$$\mathbf{K} = (\mathbf{L} \circ \mathbf{B}_n) \mathbf{H}^T \left\{ \mathbf{H} (\mathbf{L} \circ \mathbf{B}_n) \mathbf{H}^T + \mathbf{R} \right\}^{-1}.$$

Here \mathbf{L} is a matrix of four identical blocks and an element of a block is defined by

$$\mathbf{L}_{iM+j, i'M+j'} = \exp \left(-\frac{1}{2} \left[\frac{\text{dist} \{ (x_i, y_j), (x_{i'}, y_{j'}) \}}{r_0} \right]^2 \right),$$

where $\text{dist} \{ (x_i, y_j), (x_{i'}, y_{j'}) \}$ is the coordinate distance between two grid points taking into account periodic boundary conditions and scaled by Δx , and r_0 is a localisation radius, also scaled by Δx . Other choices for \mathbf{L} exist, such as for example Gaspari-Cohn function (Gaspari and Cohn 1999).

In order to define r_0 we compute empirical variogram of potential vorticity and analogously of stream function by

$$\text{variogram}(\mathbf{q}) = \frac{1}{2|N(d)|} \sum_{\{(i,j), (i',j')\} \in N(d)} |\mathbf{q}_{i,j} - \mathbf{q}_{i',j'}|^2,$$

where $N(d)$ denotes the set of pairs of $\{(i,j), (i',j')\}$ such that $\text{dist} \{ (x_i, y_j), (x_{i'}, y_{j'}) \} = d$ and $N(d)$ is the number of pairs in the set, and fit a theoretical variogram to it. The theoretical fit is needed in order to compute the range—the distance beyond which the correlation is negligible—or in other words the localisation radius. Variogram of potential vorticity is a bounded variogram (not shown) and thus spherical theoretical variogram is used to define the range. Variogram of potential vorticity is unbounded and thus Gaussian theoretical variogram is used to define the range. In Figures 8–9 we plot range of variogram for potential vorticity and stream function over time for Arakawa EZ (a),

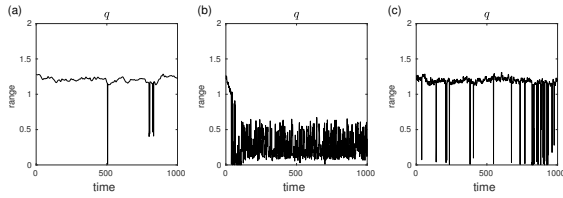


Figure 8. Range of variogram for potential vorticity over time with Arakawa *EZ* (a), Arakawa *E* (b), and Arakawa *Z* (c).

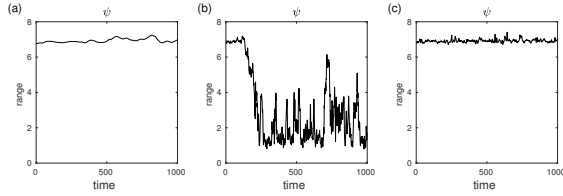


Figure 9. Range of variogram for stream function over time with Arakawa *EZ* (a), Arakawa *E* (b), and Arakawa *Z* (c).

Arakawa *E* (b), and Arakawa *Z* (c), from which it can be observed that the localisation radius for potential vorticity should be 2 and for stream function 8.

In addition to localisation, ensemble inflation is often used in an EnKF in order to increase the spread of an ensemble (Anderson and Anderson 1999). This is done by multiplying anomalies with a scalar p (usually between 1 and 1.2) prior to the application of the EnKF

$$\mathbf{z}_n^e = \bar{\mathbf{z}}_n + p(\mathbf{z}_n^e - \bar{\mathbf{z}}_n), \quad e = 1, \dots, m.$$

7. Numerical experiments

We perform so-called twin experiments, where observations are obtained by perturbing an output of a “true” mathematical model rather than from a measuring device. Such a numerical setup allows to have an access to the true solution and analyse the results of data assimilation with respect to the truth. The true model here is the QG model discretised by the HPM method from section 3.3 with $\gamma = 6$, since it preserves an infinite number of Casimirs by particles and is capable of producing the nonlinear statistical behaviour. The analysed model here is the QG model discretised by an Arakawa scheme from section 3.2.

In order to assess quality of an estimation we use the root mean square error defined as

$$\text{RMSE}(\mathbf{q}) = \sqrt{\frac{1}{M^2} (\bar{\mathbf{q}}_n - \mathbf{q}_n^{\text{truth}})^T (\bar{\mathbf{q}}_n - \mathbf{q}_n^{\text{truth}})}$$

and ratio between spread of the ensemble and RMSE, where

$$\text{spread}(\mathbf{q}) = \frac{1}{M^2} \text{std}(\mathbf{q})^T \mathbf{1}_{M^2}$$

with $\mathbf{1}_{M^2}$ being an M^2 -dimensional vector with all elements equal to 1 and

$$\text{std}(\mathbf{q}) = \sqrt{\frac{1}{m-1} \sum_{e=1}^m (\mathbf{q}_n^e - \bar{\mathbf{q}}_n) * (\mathbf{q}_n^e - \bar{\mathbf{q}}_n)}.$$

The ratio should be near one indicating that the system is reliable. The metrics are averaged over time with a transient of 500 assimilation time steps. The same metrics are defined for stream function.

In all experiments, otherwise stated, the ensemble size is $m = 25$. The assimilation period Δt_{assim} is taken to be greater than the time step of numerical discretisation Δt_{int} , namely $\Delta t_{\text{assim}} = 2$, in order to investigate the properties of a numerical discretisation in terms of error propagation. The total integration is performed on the interval $[0, t_0 + T_0]$. In all experiments we use $t_0 = 10^3$ and $T_0 = 10^4$, this makes 5500 assimilation cycles. To compare time-averaged fields, the solutions are averaged over the time interval $t \in [t_0, t_0 + T_0]$.

7.1. Stream function observations

In this section, we observe stream function ψ from the HPM method with $\gamma = 6$ at every grid point, thus the operator \mathbf{H} is equal to $[\mathbf{O} \ \mathbf{I}]$, where \mathbf{O} is the zero matrix of $M^2 \times M^2$ dimension and \mathbf{I} is the identical matrix of the same dimension. The standard deviation of the observation error is taken to be 1, thus $\mathbf{R} = \mathbf{I}$. It should be noted that neither localisation nor inflation is used.

In Figure 10 we plot relative error in energy (dashed line) and enstrophy (solid line) from Eq. (3) of analysis for different Arakawa schemes. Though none of the analyses preserve neither energy nor enstrophy, relative error in energy fluctuates around small values for both Arakawa *EZ* (0.1) and Arakawa *Z* (0.2).

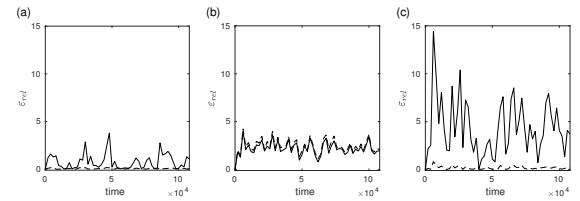


Figure 10. Relative change in energy (dotted line) and enstrophy (solid line) with Arakawa *EZ* (a), Arakawa *E* (b), and Arakawa *Z* (c) analyses.

We investigate the performance of the Arakawa schemes with respect to the time-averaged field reconstruction. Figure 11 shows scatter plots of time-averaged fields for Arakawa methods. It can be seen that the original behaviour of the Arakawa schemes—linear fit for Arakawa *EZ*, $\langle \psi \rangle = 0$ for Arakawa *E*, and $\langle \mathbf{q} \rangle = 0$ for Arakawa *Z*—is strongly influenced by data assimilation which results in nonlinear statistical behaviour. Arakawa *EZ* reproduces the nonlinear fit with better accuracy than any other Arakawa scheme. As it was noted by Dubinkina and Frank (2007), Arakawa *Z* statistically behaves as Arakawa *EZ* on long-time scales, while Arakawa *E* deviates quickly to its own statistical behaviour. Thus assimilation period has a large influence on the data assimilation results as expected.

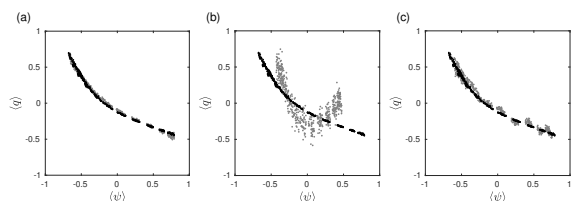


Figure 11. The scatter plot of time-averaged fields $\langle \mathbf{q} \rangle$ vs. $\langle \psi \rangle$ obtained by assimilating observations of stream function without localisation nor inflation in Arakawa *EZ* (a), Arakawa *E* (b), and Arakawa *Z* (c) shown in grey and obtained by the true HPM method shown in black.

Time-averaged stream function obtained by the Arakawa schemes shown in Figure 12 exhibits closed streamlines, which is in agreement with the HPM method shown in Figure 5(b).

In order to investigate the small-scale behaviour, we display moments of potential vorticity PDF at the grid point (3, 11) in

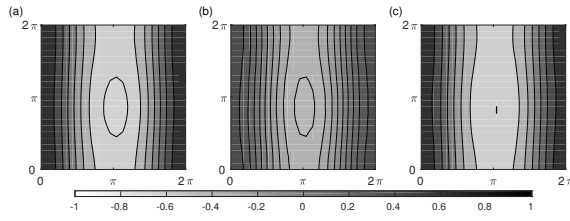


Figure 12. The contour plot of time-averaged stream function $\langle \psi \rangle$ obtained by assimilating observations of stream function without localisation nor inflation in Arakawa *EZ* (a), Arakawa *E* (b), and Arakawa *Z* (c).

the first column of metrics of Table 1. We see that the true PDF is skewed, while approximated PDFs are not for any Arakawa scheme. This means that even though Arakawa *EZ* does not estimate well small-scale behaviour, it does estimate well the large-scale behaviour.

Next we study how localisation influences the EnKF performance from statistical mechanics point of view. It was observed in Figures 8–9 that correlations are negligible for distances larger than 1 for potential vorticity and larger than 7 for stream function in Δx units. In Figure 13 we show the time-averaged RMSE of potential vorticity (a) and stream function (b) as a function of localisation radius. For Arakawa *EZ* and

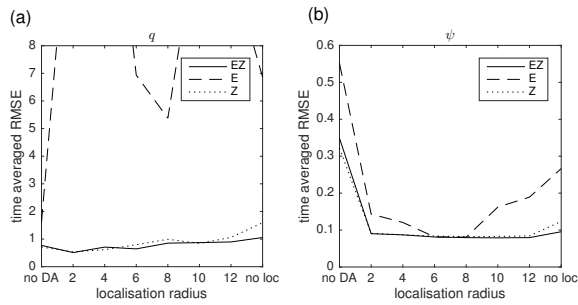


Figure 13. The time-averaged RMSE of potential vorticity (a) and stream function (b) obtained by assimilating observations of stream function using different localisation radii in Arakawa *EZ* (solid line), Arakawa *E* (dashed line), and Arakawa *Z* (dotted line).

Z optimal localisation radius in terms of the smallest RMSE of potential vorticity is 2 and in terms of stream function is 10 and 8, respectively. For Arakawa *E* it is 8 in both cases. However, the optimal localisation radius in terms of the smallest stream function RMSE gives very poor estimation of the PDF, namely mean is -0.88 for Arakawa *EZ* and -1.16 for Arakawa *Z* compared to the true mean -0.32 . Therefore we use the optimal localisation radius in terms of the smallest PV RMSE. In second column of metrics in Table 1 we give mean, standard deviation and skewness of PDF for Arakawa discretisations with optimal localisation. With or without localisation Arakawa *E* estimates poorly the PDF. For Arakawa *EZ* mean is already well approximated by the EnKF without localisation and localisation makes the approximation of the mean only worse, which also happens for Arakawa *Z*. The standard deviation, on the other hand, becomes underestimated instead of overestimated for both Arakawa *EZ* and *Z*, which highlights the need of inflation. Moreover it is interesting to note that the skewness is improved with localisation for both Arakawa *EZ* and *Z*. In Figures 14–15 we display scatter plot of time-averaged fields and time-averaged stream function, respectively, when localisation is applied. From Figure 14 we conclude that even though the RMSE decreases when localisation is used for all Arakawa discretisations, the true nonlinear statistical behaviour can be reproduced by Arakawa *EZ* only.

Table 1. Mean, standard deviation and skewness of PDF of potential vorticity for the true HPM method and for the Arakawa schemes obtained by assimilating observations of stream function without localisation (first column of metrics), and using optimal localisation radius (second column of metrics).

scheme	mean	std	skewness
HPM	-0.32	0.30	0.34
EZ	-0.33 -0.36	0.99 0.18	-0.05 0.17
E	-0.13 -5.27	6.18 1.57	0.06 2.26
Z	-0.25 -0.47	1.53 0.19	0.08 0.47

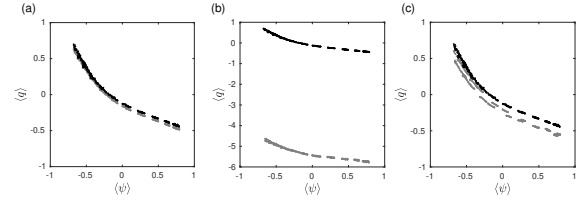


Figure 14. Same as Figure 11, but assimilating observations of stream function using optimal localisation radius.

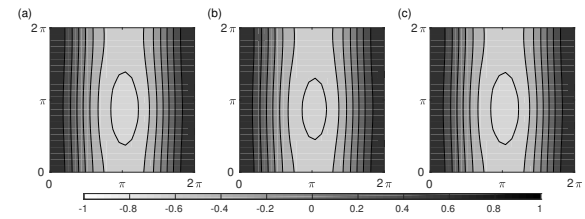


Figure 15. Same as Figure 12, but assimilating observations of stream function using optimal localisation radius.

In addition to localisation we employ inflation. The optimal inflation factor $p = 1.001$ provides the smallest RMSE in terms of potential vorticity for all Arakawa discretisations (not shown). In Figure 16 we plot the time-averaged spread to RMSE ratio for different inflation factors. It can be observed that while Arakawa *EZ* and *Z* are reliable for the optimal inflation, Arakawa *E* is not as it gives the ratio for potential vorticity close to zero. Moreover, for $p > 1.003$ Arakawa *E* is unable to converge during implicit midpoint integration. In Table 2 we show moments of PDF for the

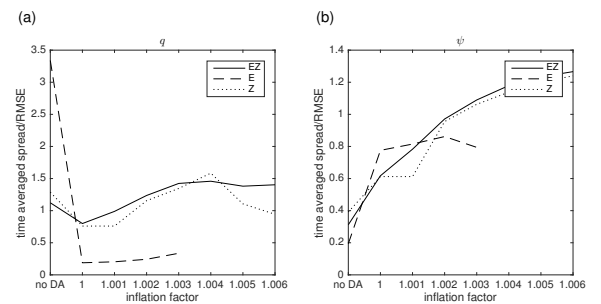


Figure 16. The time-averaged spread to RMSE ratio for potential vorticity (a) and stream function (b) obtained by assimilating observations of stream function with optimal localisation radius and different inflation factors in Arakawa *EZ* (solid line), Arakawa *E* (dashed line), and Arakawa *Z* (dotted line).

optimal inflation factor. We conclude that Arakawa *EZ* provides the best estimation of the true PDF in terms of not only the mean and standard deviation but also skewness.

When observations of stream function are assimilated, localisation and inflation improve estimation of the true PDF while keeping the nonlinear statistical behaviour intact only for Arakawa *EZ*, while for Arakawa *Z* it comes with a price of a shift

Table 2. Mean, standard deviation and skewness of PDF of potential vorticity for the true HPM method and for the Arakawa schemes obtained by assimilating observations of stream function using localisation and inflation.

scheme	mean	std	skewness
HPM	-0.32	0.30	0.34
EZ	-0.37	0.29	0.48
E	-5.27	1.70	1.99
Z	-0.51	0.25	0.45

in nonlinear relation between time-averaged fields, though time-averaged stream function is better approximated. For Arakawa *E* neither localisation nor inflation improve estimation of the true PDF nor of the true nonlinear statistical behaviour. This means that the model error of Arakawa *E* is too large and assimilation of observations of stream function is unable to compensate for this error.

7.2. PV observations

In this section, we observe potential vorticity q from the HPM method with $\gamma = 6$ at every grid point, thus the operator \mathbf{H} is equal to $[\mathbf{I} \ \mathbf{O}]$, where \mathbf{O} is again the zero matrix of $M^2 \times M^2$ dimension and \mathbf{I} is the identical matrix of the same dimension. The standard deviation of the observation error is again 1, thus $\mathbf{R} = \mathbf{I}$.

It should be noted that we were unable to carry out numerical experiments without localisation as the filter diverges, thus in Figure 17 we show the time-averaged RMSE of potential vorticity (a) and stream function (b) as a function of localisation radius. The optimal localisation radius in terms of the smallest RMSE of

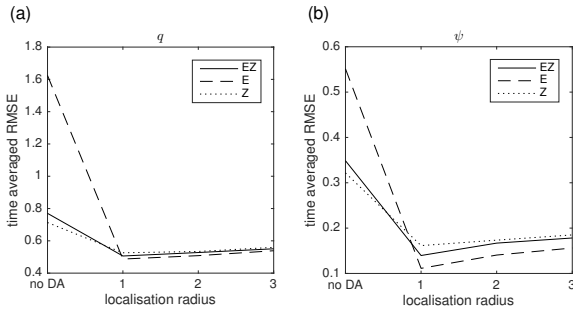


Figure 17. The time-averaged RMSE of potential vorticity (a) and stream function (b) obtained by assimilating observations of potential vorticity using different localisation radii in Arakawa *EZ* (solid line), Arakawa *E* (dashed line), and Arakawa *Z* (dotted line).

both potential vorticity and stream function is 1 for all Arakawa schemes, and all Arakawa schemes are now indistinguishable in terms of the RMSE.

In Figure 18 we show scatter plots of time-averaged fields for Arakawa methods with optimal localisation. As it could be expected from the RMSE shown in Figure 17, the approximated time-averaged field relation (shown in grey) coincides with the true one (shown in black) for any Arakawa scheme. Time-averaged stream function obtained by the Arakawa schemes shown in Figure 19 exhibit closed streamlines, which is in agreement with the HPM method shown in Figure 5(b).

The moments of potential vorticity PDF at the grid point (3, 11) are displayed in Table 3 in first column of metrics. While the mean is well approximated by all Arakawa schemes, standard deviation is underestimated, which indicates the necessity of ensemble inflation. We obtained the optimal inflation factor in terms of the smallest RMSE of potential vorticity being 1.04 for all Arakawa

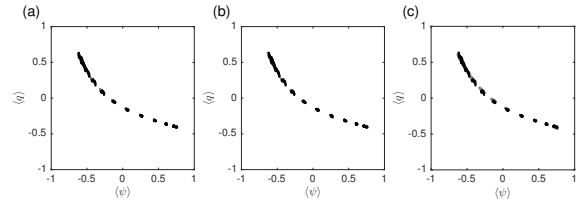


Figure 18. Same as Figure 11, but assimilating observations of potential vorticity with localisation.

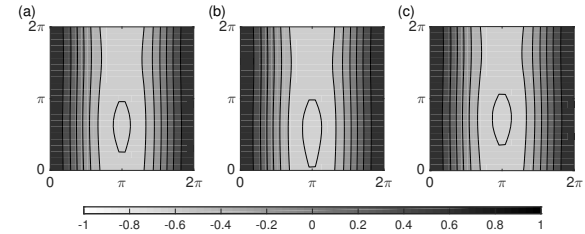


Figure 19. Same as Figure 12, but assimilating observations of potential vorticity with localisation.

discretisations (now shown). In Figure 20 we show the time-averaged spread to RMSE ratio for potential vorticity (a) and stream function (b) as a function of inflation factor. We see that all Arakawa discretisations underestimate the spread for the optimal inflation factor and they become reliable when the inflation factor goes to 1.1.

Table 3. Mean, standard deviation and skewness of PDF of potential vorticity for the true HPM method and for the Arakawa schemes obtained by assimilating observations of potential vorticity using Gauss localisation function without inflation (first column of metrics), and using optimal inflation factor (second column of metrics).

scheme	mean	std	skewness
HPM	-0.32	0.30	0.34
EZ	-0.32 -0.32	0.11 0.23	0.27 0.08
E	-0.32 -0.32	0.15 0.27	0.18 0.00
Z	-0.32 -0.32	0.10 0.23	0.27 -0.06

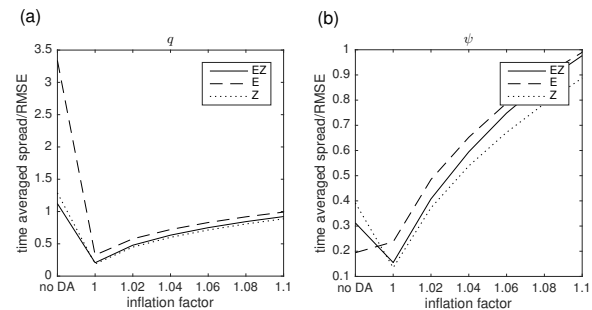


Figure 20. The time-averaged spread to RMSE ratio for potential vorticity (a) and stream function (b) obtained by assimilating observations of potential vorticity with optimal localisation radius using different inflation factors in Arakawa *EZ* (solid line), Arakawa *E* (dashed line), and Arakawa *Z* (dotted line).

In second column of metrics in Table 3 we give moments of PDF for Arakawa discretisations using optimal inflation. We observe that now both mean and standard deviation are well estimated but skewness becomes almost zero unlike in the experiments without inflation, where skewness is estimated well

by Arakawa *EZ* and *Z*. It should be noted that even better estimations of standard deviation can be obtained by tuning the inflation factor. This tuning, however, does not improve skewness as it remains close to zero (not shown). By comparing Table 2 and 3, we conclude that Arakawa *EZ* with localisation and inflation (which are necessary for counteracting sampling error) estimates better the PDF when observations of stream function rather than of PV are assimilated.

Next we investigate how localisation function influences PDF estimation by using Dirac function instead of Gauss function. We find an optimal localisation radius and optimal inflation factor for each Arakawa discretisation when Dirac localisation function is used. In Table 4 we give moments of PDF for the true HPM method and different Arakawa discretisations with and without inflation, where first column of metrics is for experiments without inflation, and second column is for the optimal inflation factor. Comparing Table 3 and Table 4 we observe that without inflation Gauss localisation function provides better estimations of the skewness than Dirac function. When inflation is used the skewness is almost zero independent of the localisation function.

Table 4. Mean, standard deviation and skewness of PDF of potential vorticity for the true HPM method and for the Arakawa schemes obtained by assimilating observations of potential vorticity using Dirac localisation function without inflation (first column of metrics), and using optimal inflation factor (second column of metrics).

scheme	mean	std	skewness
HPM	-0.32	0.30	0.34
EZ	-0.32 -0.32	0.14 0.21	0.13 0.04
E	-0.32 -0.32	0.19 0.27	0.08 0.00
Z	-0.32 -0.32	0.14 0.21	0.14 0.04

Finally we investigate Arakawa performance when a large ensemble size, namely $m = 600$, is used. With such a large ensemble it was possible to perform data assimilation without localisation nor inflation. In Table 5 we display moments of PDF, where first column of metrics is for experiments without localisation nor inflation, second column is for the optimal localisation but no inflation, and third column is for both localisation and inflation. Without localisation the skewness is poorly estimated, while with localisation the skewness becomes better. The same was observed in the experiments of assimilating observations of stream function (Table 1). When both localisation and inflation are applied, the skewness becomes worse. This holds for a small ensemble size (Table 3) as well as a large ensemble size (Table 5). This, on the contrary, was not observed in the experiments of assimilating observations of stream function where inflation improved estimations of not only standard deviation but also skewness (Table 2).

Table 5. Mean, standard deviation and skewness of PDF of potential vorticity for the true HPM method and for the Arakawa schemes obtained by assimilating observations of potential vorticity with 600 ensemble members without localisation (first column of metrics), using Gauss localisation function without inflation (second column of metrics), and using both localisation and inflation (third column of metrics).

scheme	mean	std	skewness
HPM	-0.32	0.30	0.34
EZ	-0.36 -0.32 -0.32	0.18 0.11 0.25	-0.07 0.13 0.06
E	-0.27 -0.32 -0.33	0.15 0.18 0.25	0.15 0.11 0.01
Z	-0.38 -0.32 -0.32	0.20 0.11 0.24	-0.04 0.18 -0.04

8. Conclusions

When performing data assimilation one relies on limited noisy observations, model predictions contaminated with errors and a data assimilation method with its own assumptions and errors. Therefore in order to improve data assimilation results one needs to decrease these errors. Data assimilation and statistical equilibrium mechanics share the same goal of defining a probability measure, though data assimilation combines both a model and observations that typically come from phase spaces with different probability measures, while statistical equilibrium mechanics relies solely on the model (either continuous or discrete). The discrete system obtained by a numerical approximation has its own probability measure depending on the conserved quantities that discrete system possesses. In this paper we studied relevance of quantities conserved by a numerical scheme for data assimilation, namely when observations come from a phase space with one probability measure and the model state comes from a phase space with another probability measure.

We considered three Arakawa discretisations: Arakawa *EZ* which preserves both energy and enstrophy, Arakawa *E* which preserves only energy, and Arakawa *Z* which preserves only enstrophy. The initial condition of the true solution obtained by the HPM method was such that the scaled third moment of potential vorticity is equal to 6, which corresponds to the gamma PDF. In this case the long-time average of the fields exhibit nonlinear relation, according to the statistical mechanical theory of the HPM method, and none of the Arakawa schemes is able to reproduce such behaviour alone. As the data assimilation method we chose an Ensemble Kalman Filter with perturbed observations due to its popularity.

We showed that when observations of stream function are assimilated, the choice of a numerical scheme is crucial for a good reconstruction of time-averaged fields: Arakawa *EZ* provides the best estimate, Arakawa *Z* gives a slightly worse but still good estimate, and Arakawa *E* completely fails. Moreover, without localisation nor inflation both Arakawa *EZ* and *Z* do not estimate well the potential vorticity PDF as they provide normal PDF while the true PDF is skewed. This indicates the ability of the Ensemble Kalman Filter to reproduce nonlinear large-scale behaviour without reproducing non-Gaussian small-scale behaviour. For better reconstruction of the PDF, when assimilating observations of stream function, we need to use both localisation and inflation. With optimally chosen localisation radius and inflation factor we are able to improve estimations of mean, standard deviation and skewness of the PDF keeping the nonlinear statistical behaviour of time-averaged fields intact only for Arakawa *EZ*. Arakawa *Z* gives better estimation of the standard deviation and skewness but not mean of the PDF. It improves time-averaged stream function but introduces a shift in the nonlinear relation between time-averaged fields. Arakawa *E* does not improve the PDF estimation nor relation between time-averaged fields regardless of the localisation radius and inflation factor.

When observations of potential vorticity are assimilated, we can obtain good reconstructions of the long-time averages of the state and the PDF independent of a numerical scheme, though it becomes indispensable to apply localisation. We showed that Gauss localisation function provides better estimations of the PDF skewness than Dirac function, when inflation is not applied. When inflation is applied, the skewness becomes almost zero independent of the localisation function and ensemble size. As our future work we plan to study whether more advanced data assimilation methods that account for sampling error due to the finite size ensemble (e.g. Bocquet *et al.* 2015) deteriorate skewness estimations.

The difference between assimilating stream function and potential vorticity is that one assimilates large-scale observations and another one small-scale observations. Since Arakawa E does not model correctly the inverse energy cascade leading to too much energy on the smallest scales, assimilating stream function does not improve the estimation, while assimilating potential vorticity does. Arakawa EZ , on the contrary, does not have this flaw and thus provides a good estimation of nonlinear statistical behaviour even when assimilating stream function.

It should be noted that apart from the HPM method there exists another numerical method that exhibits nonlinear statistical behaviour—the Zeitlin scheme (Zeitlin 1991) combined with McLachlan splitting for time discretisation (McLachlan 1993). Abramov and Majda (2003) showed that this numerical approximation is capable of producing both normal and skewed PDF, since it preserves energy and $2M$ Casimirs for an M -mode approximation. We have repeated the numerical experiments with the Zeitlin model combined with McLachlan splitting as the true model and confirmed conclusions drawn in the paper. Therefore the results are not sensitive with respect to the true model.

Acknowledgement

We wish to acknowledge the EnKF-Matlab package (version 0.31) developed by P. Sakov (<http://enkf.nersc.no/Code/EnKF-Matlab/>).

References

- Abramov R, Majda AJ. 2003. Statistically relevant conserved quantities for truncated quasi-geostrophic flow. *Proc. Natl. Acad. Sci. U.S.A.* **100**(7): 3841–3846.
- Anderson J, Anderson SL. 1999. A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review* **127**(12): 2741–2758, doi:10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.
- Arakawa A. 1966. Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part I. *J. Comput. Phys.* **1**: 119–143.
- Bocquet M, Raanes P, Hannart A. 2015. Expanding the validity of the ensemble kalman filter without the intrinsic need for inflation. *Nonlinear Processes in Geophysics* **22**(6): 645–662, doi:10.5194/npg-22-645-2015.
- Burgers G, van Leeuwen P, Evensen G. 1998. Analysis scheme in the ensemble kalman filter. *Monthly Weather Review* **126**(6): 1719–1724, doi:10.1175/1520-0493(1998)126.
- Carnevale G, Frederiksen J. 1987. Nonlinear stability and statistical mechanics of flow over topography. *J. Fluid Mech.* **175**: 157–181.
- Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P, Brnnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ø, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ. 2011. The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society* **137**(654): 1–28, doi:10.1002/qj.776.
- Cotter C, Reich S. 2004. Geometric integration of a wave-vortex model. *Appl. Numer. Math.* **48**: 293–305.
- DiBattista M, Majda A, Grote M. 2001. Meta-stability of equilibrium statistical structures for prototype geophysical flows with damping and driving. *Phys. D* **151**(2–4): 271–304.
- Doucet A, de Freitas N, Gordon N. 2001. *Sequential monte-carlo methods in practice*. Springer-Verlag: New York, ISBN 978-0-387-95146-1.
- Dubinkina S, Frank J. 2007. Statistical mechanics of Arakawa’s discretizations. *J. Comput. Phys.* **227**: 1286–1305.
- Dubinkina S, Frank J. 2010. Statistical relevance of vorticity conservation in the Hamiltonian particle-mesh method. *Journal of Computational Physics* **229**(7): 2634–2648, doi:http://dx.doi.org/10.1016/j.jcp.2009.12.012.
- Ellis RS, Haven K, Turkington B. 2002. Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows. *Nonlinearity* **15**(2): 239–255.
- Evensen G. 2009a. *Data assimilation: The ensemble kalman filter*. Springer, 2 edn, ISBN 978-3-642-03710-8.
- Evensen G. 2009b. The ensemble kalman filter for combined state and parameter estimation. *IEEE Control Systems* **29**(3): 83–104, doi:10.1109/MCS.2009.932223.
- Gaspari G, Cohn S. 1999. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society* **125**(554): 723–757, doi:10.1002/qj.49712555417.
- Griffa A, Salmon R. 1989. Free content wind-driven ocean circulation and equilibrium statistical mechanics. *J. Mar. Res.* **47**: 457–492.
- Hairer E, Lubich C, Wanner G. 2006. *Geometric numerical integration, Springer Series in Computational Mathematics*, vol. 31. Springer-Verlag: Berlin, second edn, ISBN 3-540-30663-3; 978-3-540-30663-4. Structure-preserving algorithms for ordinary differential equations.
- Hamill TM, Whitaker JS, Snyder C. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Review* **129**(11): 2776–2790, doi:10.1175/1520-0493(2001)129(2776:DFOBE)2.0.CO;2.
- Jacobs G, Ngodock H. 2003. The maintenance of conservative physical laws within data assimilation systems. *Monthly Weather Review* **131**: 2595–2607.
- Janjić T, McLaughlin D, Cohn S, Verlaan M. 2014. Conservation of mass and preservation of positivity with ensemble-type kalman filter algorithms. *Monthly Weather Review* **142**: 755–773.
- Jazwinski AH. 1970. *Stochastic processes and filtering theory*. Mathematics in science and engineering, Academic press: New York, ISBN 0-12-381550-9. UKM.
- Kraichnan R. 1975. Statistical dynamics of two-dimensional flow. *J. Fluid Mech.* **67**: 155–175.
- Majda AJ, Wang X. 2006. *Non-linear dynamics and statistical theories for basic geophysical flows*. Cambridge University Press: Cambridge, ISBN 978-0-521-83441-4; 0-521-83441-4.
- McLachlan RI. 1993. Explicit Lie-Poisson integration and the Euler equations. *Phys. Rev. Lett.* **71**(19): 3043–3046.
- Montgomery D, Matthaeus W, Martinez D, S O. 1992. Relaxation in two dimensions and sinh-poisson equation. *Phys. Fluids A* **4**: 3–6.
- Onsager L. 1949. Statistical hydrodynamics. *Nuovo Cimento, Suppl.* **6**(2): 279–287.
- Pedlosky J. 1987. *Geophysical fluid dynamics*. Springer: New York, 2nd edn.
- Salmon R. 1998. *Lectures on geophysical fluid dynamics*. Oxford University Press: New York, ISBN 0-19-510808-6.
- Salmon R, Holloway G, Hendershott M. 1976. The equilibrium statistical mechanics of simple quasi-geostrophic models. *J. Fluid Mech.* **75**: 691–703.
- Verkley W, Lynch P. 2009. Energy and enstrophy spectra of geostrophic turbulent flows derived from a maximum entropy principle. *Journal of the Atmospheric Sciences* **66**: 2216–2236.
- Wang Y, Counillon F, Bertino L. 2016. Alleviating the bias induced by the linear analysis update with an isopycnal ocean model. *Quarterly Journal of the Royal Meteorological Society* **142**(695): 1064–1074, doi:10.1002/qj.2709.
- Zeitlin V. 1991. Finite-mode analogues of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure. *Physica D* **49**(3): 353–362.
- Zeng Y, Janjić T. 2016. Study of conservation laws with the local ensemble transform kalman filter. *Quarterly Journal of the Royal Meteorological Society* **142**(699): 2359–2372, doi:10.1002/qj.2829.