Ocean Science
Discussions

Model-to-model data assimilation method for fine resolution ocean modelling

Georgy I. Shapiro[1], Jose Maria Gonzalez-Ondina[2]

[1] School of Biological and Marine Sciences, University of Plymouth, Plymouth, PL4 8AA, United Kingdom

[2]University of Plymouth Enterprise Ltd, Plymouth, PL4 8AA, United Kingdom

Correspondence to Georgy I. Shapiro (gshapiro@plymouth.ac.uk)

Abstract

An effective and computationally efficient method is presented for data assimilation in a high-resolution (child) ocean model, which is nested into a coarse-resolution good quality data assimilating (parent) model. The method named Data Assimilation with Stochastic-Deterministic Downscaling (SDDA) reduces bias and root mean square errors (RMSE) of the child model and does not allow the child model to drift away from reality. The basic idea is to assimilate data from the parent model instead of actual observations. In this way, the child model is physically aware of observations via the parent model. The method allows to avoid a complex process of assimilating the same observations which were already assimilated into the parent model. The method consists of two stages: (1) downscaling the parent model output onto the child model grid using Stochastic-Deterministic Downscaling, and (2) applying a simplified Kalman gain formula to each of the fine grid nodes. The method is illustrated in a synthetic case where the true solution is known, and the child model forecast (before data assimilation) is simulated by adding various types of errors. The SDDA method reduces the child model bias to the same level as in the parent model and reduces the RMSE typically by a factor of 2 to 5.

Introduction

Fine resolution ocean modelling is becoming a ubiquitous practice to resolve important mesoscale and sub-mesoscale features such as eddies, fronts, boundary currents and localised upwellings which play important roles in ocean dynamics, see e.g. (GFDL, 2021; Dufour et al., 2015; Shapiro et al, 2010; T. Meunier et al, 2012.) Such localised models can be run by relatively small groups due to availability of good quality ocean models such as ROMS or NEMO to the wider oceanographic community (ROMS, 2021; NEMO, 2021). Local fine resolution models require initial and boundary conditions which can be obtained from good quality, but coarser resolution models run by major ocean modelling centres such as Mercator Ocean International (France) or Met Office (UK) via Copernicus Marine Service (CMEMS, 2021).

Due to inevitable approximations in the equations, numerical schemes, parameterisation and uncertainties in input data, ocean models tend to drift from reality. A process called data assimilation (hereafter DA) is often regarded as a way of keeping a model 'on the tracks' by constantly correcting it with fresh observations (DARC, 2021; Lorenc, 1986). DA is considered a cornerstone of all ocean analysis and forecasting efforts, where the rigorous and systematic combination of ocean

Ocean Science
Discussions

Open Access

41   observations and ocean models yields an optimal estimate of the ocean state, see e.g. (Bell et al, 2000;
42   Moore et al, 2019). Numerical models of the ocean are able to assimilate oceanographic observations,
43   creating a dynamically consistent, complete and accurate description of atmosphere and ocean, see
44   e.g. (Ciavatta et al, 2018; Carrassi et al, 2018). Modern data assimilation is a complex process involving
45   statistics, methods of dynamical systems and numerical optimization, with an additional difficulty
46   which arises due to increasing sophistication of the environmental models. While the DA problem can
47   be formulated precisely, the solution is challenging because of the vast number of degrees of freedom
48   of the ocean state simulated by operational ocean models that represent many complex non-linear
49   processes (Moore et al, 2019; Dobricic and Pinardi, 2008; Ghil and Malanotte-Rizzoli, 1991). Thus,
50   despite data assimilation being nowadays ubiquitous in geosciences, it has so far remained a topic
51   mostly reserved to experts (Kubryakov et al, 2012; Carrassi et al, 2018).

52   The purpose of this paper is to develop a simple and computationally efficient method of DA which
53   can be implemented by smaller academic and operational centres which do not have resources of the
54   same scale as the major ocean forecasting institutions. The basic idea suggested in this study is to
55   assimilate data from a good quality but coarser resolution parent model instead of observations. The
56   parent model is assumed to be data-assimilating itself, ensuring that the observations, however
57   indirectly, will not allow the fine resolution (hereafter called child) model to deviate significantly from
58   the true state of the sea. In other words, the coarse model can be regarded as a physically aware
59   method for incorporating a sparse number of observations into a fine resolution regular grid.

60   While in principle a number of existing methods could be used for the model-to-model DA, this study
61   uses the Stochastic-Deterministic Downscaling (Shapiro et al, 2021) as its core element. The model-
62   to-model DA algorithm (hereafter called Data Assimilation with Stochastic-Deterministic Downscaling,
63   or SDDA) is described in the Data and Methods section. This section also describes the synthetic
64   idealised case where the true solution is known. Section Results uses the synthetic case for
65   demonstrating the performance of the SDDA method. The strengths and limitations of the SDDA
66   methods are discussed in the final section of the paper in comparison with the commonly used
67   combination of Hollingworth-Lönnberg and variational DA (VAR, also known as Optimal Interpolation)
68   methods (Hollingsworth and Lönnberg, 1986; Kalnay, 2003).

69

Data and methods

71   This section presents the algorithm used for Data Assimilation with Stochastic-Deterministic
72   Downscaling and describes how the idealised cases are set and processed.

The algorithm

74   The proposed algorithm is designed for assimilating data from an ocean model of coarser resolution
75   parent model, into a finer resolution child model.

76   A common approach to data assimilation into ocean (or atmospheric) model is based on minimising
77   the cost function J given in Eq (1)

78 $$J(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}^b)^T \mathbf{B}^{-1}(\boldsymbol{x} - \boldsymbol{x}^b) + \left(\boldsymbol{y} - H(\boldsymbol{x})\right)^T \mathbf{R}^{-1}\left(\boldsymbol{y} - H(\boldsymbol{x})\right) \qquad (1)$$

79   where $\boldsymbol{x}$ is the (unknown) vector of best estimates of the true values, $\boldsymbol{x}^b$ is the model forecast before
80   data assimilation, $\boldsymbol{y}$ is the vector of observed values, H is the operator which projects data from the
81   model grid onto the locations of observations, $\mathbf{B}$ and $\mathbf{R}$ are error covariance matrices for the model
82   and observations respectively. In this paper we use the notation recommended in (Ide et al, 1997).

Ocean Science

Discussions

Open Access

83   The value of $\boldsymbol{x} = \boldsymbol{x}^a$ which minimises the cost function J is called the analysis, it is closest in an RMS
84   (root-mean-square) sense to the true state $\boldsymbol{x}^t$ (Bouttier and Courtier, 1999) This approach works well
85   for a relatively small (compared to model output) number of data. When the number of observations
86   is large, the matrices used in Eq (1) become very large and their inversion causes significant
87   computational problems.   For example, Bouttier and Courtier (1999) state that 'except in analysis
88   problems of very small dimension (like one-dimensional retrievals), it is impossible to compute exactly
89   the least-squares analysis.'

90   However, this study is related to assimilation of large amounts of data from one model to another.
91   For this purpose, an alternative cost function J_S is proposed, namely

92   $$J_S(\boldsymbol{x}') = \left(\boldsymbol{x}' - \boldsymbol{x}'^b\right)^T \mathbf{B}^{-1}\left(\boldsymbol{x}' - \boldsymbol{x}'^b\right) + (S'(\boldsymbol{y}) - \boldsymbol{x}')^T \mathbf{R}^{-1}(S'(\boldsymbol{y}) - \boldsymbol{x}') \qquad (2)$$

93   where the primed variables denote deviations of the respective variable from their statistical mean,
94   and the statistical mean is designated by angle brackets

95   $$\boldsymbol{x}' = \boldsymbol{x} - <\boldsymbol{x}>$$

96   $$\boldsymbol{x}' = \boldsymbol{x}^b - <\boldsymbol{x}^b>$$

97   $$S'(\boldsymbol{y}) = S(\boldsymbol{y}) - <S(\boldsymbol{y})>$$

98

99   Here $\boldsymbol{y}$ is the vector of data from the parent model, and S is the operator which projects (downscales)
100 the data from the parent model onto the fine grid of the child model. The best estimate for the true
101 value x is obtained by minimising the cost function

102 $$\nabla J_S = \mathbf{0}$$

103 Taking into account that matrices $\mathbf{B}$ and $\mathbf{R}$ are symmetric so that $\mathbf{B}^T = \mathbf{B}$ and $\mathbf{R}^T = \mathbf{R}$ one gets the
104 equation

105 $$\nabla J_S\left(\boldsymbol{x}'^a\right) = \left(\boldsymbol{x}'^a - \boldsymbol{x}'^b\right)^T (\mathbf{B}^{-1} + (\mathbf{B}^{-1})^T) - \left(S'(\boldsymbol{y}) - \boldsymbol{x}'^a\right)^T (\mathbf{R}^{-1} + (\mathbf{R}^{-1})^T) =$$

106 $$2\left(\boldsymbol{x}'^a - \boldsymbol{x}'^b\right)^T \mathbf{B}^{-1} - 2\left(S'(\boldsymbol{y}) - \boldsymbol{x}'^a\right)^T \mathbf{R}^{-1} = \mathbf{0} \qquad (3)$$

107

108 where $\boldsymbol{x}'^a$ is the fluctuation of the analysis vector around its statistical mean.

109 Taking the transpose of Eq (3) and dividing by two one gets

110

111 $$\mathbf{B}^{-1}(\boldsymbol{x}'^a - \boldsymbol{x}'^b) = \mathbf{R}^{-1}(S'(\boldsymbol{y}) - \boldsymbol{x}'^a) \qquad (4)$$

112 Let us introduce the error correlation matrices $\mathbf{C_B}$ and $\mathbf{C_R}$ for the child and downscaled parent
113 models:

114 $$\mathbf{B} = \mathbf{V_B} \mathbf{C_B}, \qquad \mathbf{R} = \mathbf{V_R} \mathbf{C_R}$$

Ocean Science
Discussions

115   where $\mathbf{V}_B$, $\mathbf{V}_R$ are the diagonal matrices containing respective error variances at each fine grid
116   node. The inverted diagonal matrices are also diagonal. Equation (4) then becomes

117

118    $$\mathbf{C}_B^{-1}\mathbf{V}_B^{-1}(x'^a - x'^b) = \mathbf{C}_R^{-1}\mathbf{V}_R^{-1}(S'(y) - x'^a) \qquad (5)$$

119

120   Correlation matrices both for the child and downscaled parent models relate to the same resolution
121   and the same area of the ocean, hence it is reasonable to use the same correlation function, e.g. a
122   Gaussian of a certain length scale and therefore the same correlation matrix for both models, i.e.
123    $\mathbf{C}_B = \mathbf{C}_R = \mathbf{C}$. After pre-multiplying Eq (5) by $\mathbf{C}$ and re-arranging the terms, one gets

124

125    $$(\mathbf{V}_B^{-1} + \mathbf{V}_R^{-1})x'^a = \mathbf{V}_B^{-1}x'^b + \mathbf{V}_R^{-1}S'(y) \qquad (6)$$

126

127   or $x'^a = (\mathbf{V}_B^{-1} + \mathbf{V}_R^{-1})^{-1}(\mathbf{V}_B^{-1}x'^b + \mathbf{V}_R^{-1}S'(y)) \qquad (7)$

128

129   Using the commutative properties of diagonal matrices and the following identity

130

131    $$(\mathbf{V}_B^{-1} + \mathbf{V}_R^{-1})^{-1} = \mathbf{V}_R(\mathbf{V}_B + \mathbf{V}_R)^{-1}\mathbf{V}_B, \qquad (8)$$

132

133   the solution for the analysis state given by Eq (7) can be re-written as

134    $x'^a = (\mathbf{V}_B^{-1} + \mathbf{V}_R^{-1})^{-1}(\mathbf{V}_B^{-1}x'^b + \mathbf{V}_R^{-1}S'(y)) = \mathbf{V}_R(\mathbf{V}_B + \mathbf{V}_R)^{-1}x'^b +$
135    $\mathbf{V}_B(\mathbf{V}_B + \mathbf{V}_R)^{-1}S'(y) \qquad (9)$

136

137   Eq (9) can be interpreted as a zero-dimensional Kalman gain formula applied to the fluctuations of
138   state variables in the parent and child models at each fine grid node independently. The term 'zero-
139   dimensional' reflects the fact that in this case, the matrices used in the Kalman gain formula have the
140   size of 1x1, i.e. are reduced to a scalar. For a single x'_i element of the state vector x' Eq (9) gives

141    $$x_i'^a = \frac{V_{Rii}}{V_{Rii}+V_{Bii}} x_i'^b + \frac{V_{Bii}}{V_{Rii}+V_{Bii}} S'(y)_i \qquad (10)$$

142   For the S' operator, which downscales the fluctuations of the field variable from the coarse to fine
143   grid, we use the Stochastic-Deterministic Downscaling (SDD) method developed in (Shapiro et al,
144   2021). This method calculates the weighting coefficients required for downscaling from coarse to fine
145   grid by adopting the original algorithm of optimal interpolation (Gandin, 1959, Gandin 1965) and
146   statistical properties of field variable estimated from the parent model outputs. As a result, it
147   generates lower errors during downscaling than if using interpolators with prescribed coefficients

Ocean Science
Discussions

148 such as linear or polynomial. The spatial correlation matrix $\mathbf{C}$ is not explicitly included in expressions
149 (9) and (10), but it is used at the downscaling stage of data assimilation, hence the covariances
150 between data at the parent model grid are implicitly present in the operator S'. The ensemble
151 statistical means at each grid node which are required by the SDD method are estimated using the
152 ergodic hypothesis (Stull, 1998), i.e. by spatial averaging over a small trial area around the node at
153 the same time point, see (Shapiro et al 2021) for details.

154 Equation (10) performs data assimilation on the fluctuations, however, the statistical means for the
155 parent and child model are, in a general case, different. There are at least two options of how to
156 obtain the mean for the analysis. Option one is to apply to the means the same Kalman filter as for
157 fluctuations. The problem here is that variances of statistical means are unknown, and hence the
158 weighting coefficients in Eq (9) or Eq (10). Option two is to assume that the parent model is of good
159 quality, and it has been debiased as much as possible during its own data assimilation cycle. Therefore,
160 it is reasonable to replace the (potentially biased) statistical mean from the child model with the
161 statistical mean of the parent model.

162 This gives the final equation for calculation of the analysis state

163 $$x^a = \mathbf{V_R}(\mathbf{V_B} + \mathbf{V_R})^{-1}x'^b + \mathbf{V_B}(\mathbf{V_B} + \mathbf{V_R})^{-1}S'(\boldsymbol{y}) + <S(\boldsymbol{y})> \quad (11)$$

164 Or for an element of the state vector at a particular fine grid node i $(i = 1 \dots N)$ where N is the
165 number of grid nodes in the child model one gets

166 $$x_i^a = \frac{V_{Rii}}{V_{Rii} + V_{Bii}}x'^b{}_i + \frac{V_{Bii}}{V_{Rii} + V_{Bii}}S'(y)_i + <S(y)>_i \quad (12)$$

167 To summarise, the SDDA model-to-model data assimilation procedure includes two steps. Firstly, to
168 downscale the parent model output from the coarse grid onto the child model fine grid using the SDD
169 method. The result is that at each fine grid node there are two values of the same state variable. The
170 best estimate of the true field is obtained at the second step by combining these two values using a
171 zero-dimensional Kalman filter. This algorithm is computationally efficient as it does not require
172 inversion of large matrices or solving a very large system of algebraic equations at the second step,
173 something that is required if using the variational methods based on Eq (1). The inversion of
174 correlation matrices to obtain the weight coefficients for the SDD step can be done only once at the
175 beginning of the model run, as these coefficients do not depend on time. Another benefit of the
176 described method is that the correlation matrices for the downscaling of the parent model have a
177 relatively small rank and condition number, and their inverse counterparts can be calculated without
178 the need of any type of matrix regularisation, just using double precision for the computer
179 representation of the variables. This is due to the fact that the SDD method assumes local
180 homogeneity and isotropy of statistical properties of the field variable and that the correlation
181 function is set to zero for distances larger than a threshold value. More details on the philosophy and
182 the algorithm of the SDD method can be found in (Shapiro et al, 2021).

183 Synthetic Idealised case
184 In this section the SDDA algorithm is illustrated in a synthetic idealised case where the true solution is
185 known. The task is to generate an analysis state using the fine-resolution (child) model forecast and
186 the output from a good quality data assimilating parent coarse model which will be used instead of
187 actual observations. It should be noted that, even if both coarse and fine model were perfect, there
188 would be some unavoidable discrepancies, or *representativeness errors*, between the two models due
189 to different meshes they use (Bouttier and P. Courtier, 1999).

190 The parent model output is simulated by sampling the true solution (a prescribed function) on the
191 parent model coarse grid, to which random noise might be added. The fine-resolution forecast is
192 simulated by sampling the same true solution on the fine grid, adding random noise, bias, and some
193 spatial shift. The latter is to simulate the 'double penalty effect' which is common to fine resolution
194 models, see e.g. (Zingerlea and Nurmib, 2008).

195 Three distinct examples are considered, all of them relate to data assimilation of one field variable at
196 a single computational surface, whether it be a geopotential ('horizontal') level, such as used in Bell
197 et al. (2000), or a curved level such as used in sigma (Mellor and Blumberg, 1985) or multi-envelope
198 vertical coordinate system (Bruciaferri et al, 2018). The properties of the SDDA method are analysed
199 for: (i) an ocean front, (ii) an isolated isotropic eddy, (iii) a system of densely packed anisotropic
200 mesoscale eddies. In all examples the parent model has resolution of $\Delta x_p = \Delta y_p = 10$ km, and the
201 child model has resolution of $\Delta x_p = \Delta y_p = 10$ km. For all the examples, the correlation function
202 which is used to calculate the covariance matrix $\boldsymbol{C}$ was set to zero for distances greater than twice the
203 correlation length of $L = 17$ km, and the trial area used to calculate local spatial averages was a square
204 of 68 x 68 km² centred at each node.

205

206

207   a) Ocean front

208

209 An ocean front is a narrow area separating two water masses, and it has significant impact on
210 horizontal and vertical exchanges, see e.g. (Fedorov,1986). This example is set in a square domain of
211 200x200km, and the front extends in the meridional (y) direction at the centre of the domain.

212 The true solution is set to be in the form

213 $$F = \text{A} \tanh\left(\frac{x}{L_x}\right) \tag{13}$$

214 Where A is the amplitude of the front and L_x is its half-width. For this exercise $= 1$, and L_x ranges
215 between 6 and 40 km.

216   b) Isolated mesoscale eddy
217 Mesoscale eddies are a ubiquitous feature of the World Ocean. Originally, they were thought to exist
218 only next to jet currents such as the Gulf Stream, however since the 1960s-1970s it became clear that
219 mesoscale eddies exist nearly anywhere in the ocean. Most of the kinetic energy of the ocean is
220 contained in mesoscale eddies (Robinson, 1983). In this example an isotropic eddy is placed in the
221 centre of a domain of 200×200 km and the true solution is set to be in the form

222

223 $$F = \text{A} \exp\left\{-\left[\left(\frac{x}{L_e}\right)^2 + \left(\frac{y}{L_e}\right)^2\right]\right\} \tag{14}$$

224 Where A is the amplitude of the eddy and L_e is its radius. For this exercise $A = 1$, and L_e ranges
225 between 6 and 46 km.

226

Ocean Science
Discussions

Open Access

EGU

227   c)   Multiple mesoscale eddies
228   For this example, let us consider a square domain 1000×1000 km and take the true solution for a
229   variable F in the form of multiple anisotropic mesoscale eddies

230

231    $$F = \sin\left(\frac{\pi x}{L_x}\right) \sin\left(\frac{\pi y}{L_y}\right)$$ (15)

232   where L_x is the eddy 'radius' in the x-direction, and $L_y = 105$ km is the eddy 'radius' size in the y-
233   direction. Let us consider the range of eddy sizes L_x between 12 and 24 km. At L_x =24 km the parent
234   model can be classed as eddy-permitting as it has 2.4 grid point over the smaller 'diameter' of the
235   eddy. At L_x =12 km the parent model is not even eddy-permitting but only showing an 'embryonic'
236   representation of the eddy. The child fine-resolution model is eddy resolving for any L_x within the
237   chosen range of eddy sizes.

238

239   Results
240   The SDDA data assimilation method detailed in the section Algorithm above was applied to the
241   simulated child model forecast in order to create analysis for the next forecasting cycle. The results
242   for four examples are presented in this section: an ocean front, an isolated single eddy, a set of
243   multiple eddies.

244   A)   Ocean front
245   This example uses Eq (13) for the representation of an ocean front in the meridional direction. Fig.1
246   shows a map of the true solution F for a sharp ocean front at L_x=6 km (see Fig.1(a)) , its representation
247   by the parent model (Fig.1(b)) , by child model forecast before data assimilation (Fig. 1(c)) as well as
248   transects across the front. The following parameters of added errors are used to simulate the child
249   model forecast before data assimilation: normally distributed noise with standard deviation 0.15,
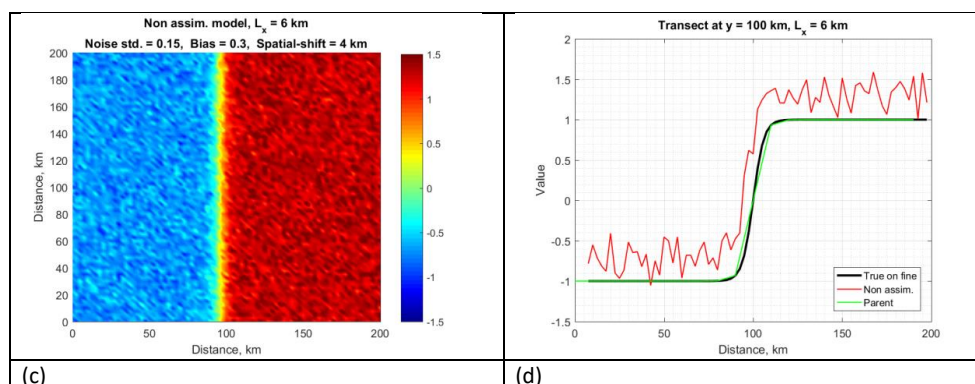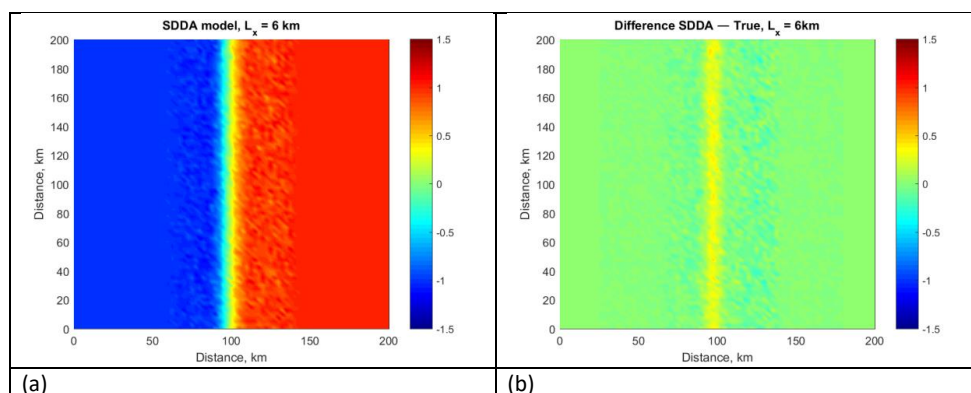250   positive bias of 0.3, and shift of the field by 4 km to the west.

251

(a)

(b)

Ocean Science
Discussions

Open Access

(c)

(d)

Fig. 1: *Ocean front of width $L_x = 6$ km. (a) True field on the fine mesh, (b) parent model output, (c) simulated child model before data assimilation, (d) transects of the previous fields along the line at $y = 100$ km.*
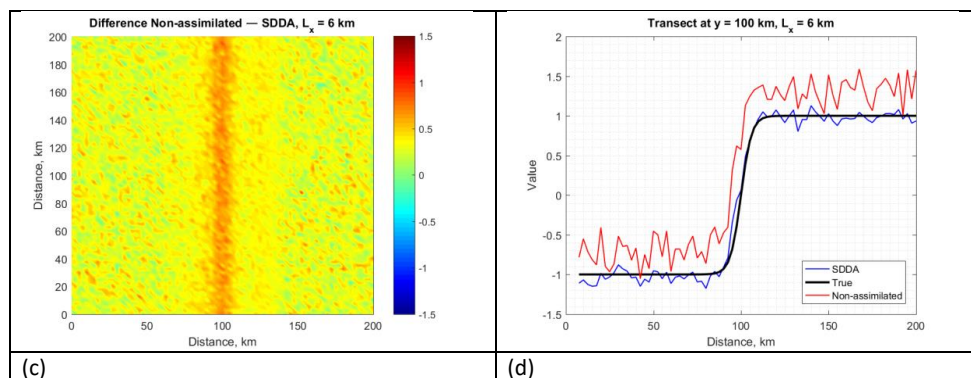
Even at a resolution of 10 km that does not resolve the structure of the front (half-width of 6 km) the parent model gives a reasonable representation of areas outside the front where the changes in the state variable are smooth- see Fig. 1(b). However, as expected, the width of the front is exaggerated due to insufficient resolution. The child model forecast shown in Fig. 1(c) is noisy and clearly shifted westward relative to the true state.

Fig.2 shows the results of assimilation of output from the parent model using the SDDA method.

(a)

(b)

Ocean Science
Discussions

Open Access

EGU

(c)   (d)

263

Fig. 2. *Assimilation results for the ocean front shown in **Figure 1**: (a) child model after SDDA*
assimilation, (b) difference between the assimilated model and the true field, (c) difference between
the non-assimilated and assimilated child models, (d) zonal transects of the previous fields along a
line at $y = 100\ km$.

The analysis state (after assimilation) removes the spatial shift and bias and reduces the noise even
for such a sharp front where the resolution of the parent model is inadequate. For comparison, Fig.3
shows the improvement provided by the SDDA assimilation for fronts of different sharpness, with L_x
= 10, 20, 30 and 40 km. In all examples the errors in non-assimilating child model forecasts are
simulated by adding bias = 0.3, random normally distributed noise with standard deviation (STD) =
0.15 and spatial shift of 4 km to the west.
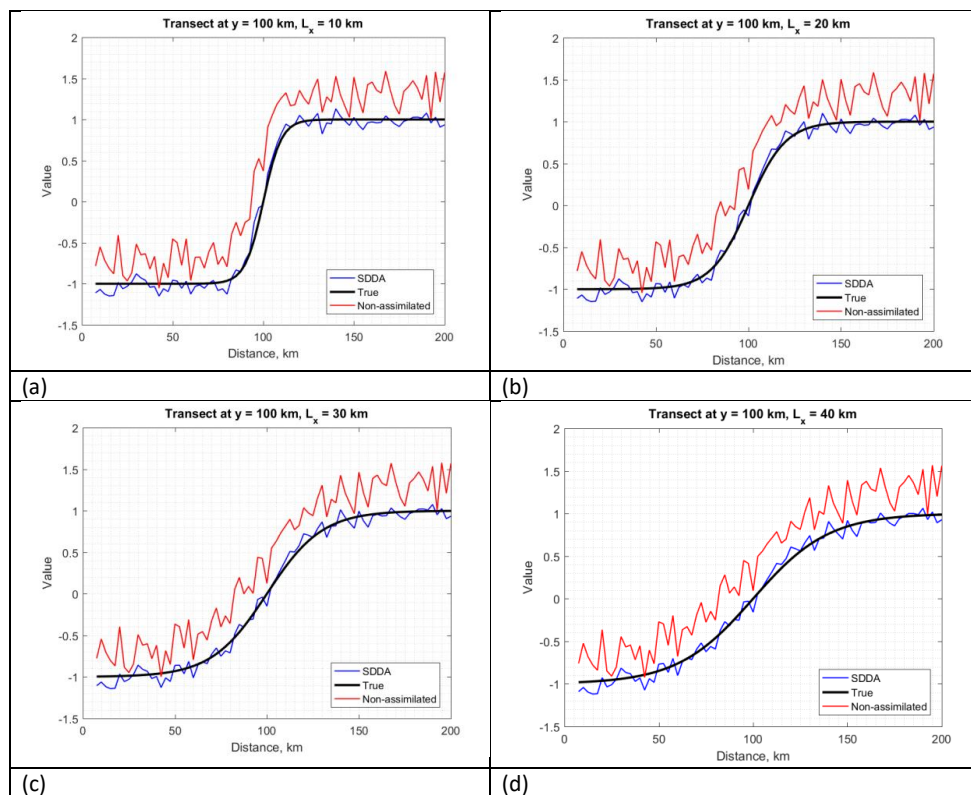
274

275

276

277

278

279

280

281

282 **Fig.3.** *Results of SDDA assimilation for fronts of different sharpness: (a) $L_x = 10$ km, (b) $L_x = 20$*
283 *km, (c) $L_x = 30$ km, (d) $L_x = 40$ km. The curves present data for the true field (black), non-*
284 *assimilated noisy child model (red) and child model after SDDA assimilation (blue).*

285 Comparison of panels (a)-(d) in Fig. 3 and panel (d) in Fig. 2 shows that the improvement due to data
286 assimilation from the parent model is achieved both for sharp fronts not resolved by the parent model
287 and for smooth fronts.

288 Fig.4 shows how the bias and RMSE of the child model against the true state change with the
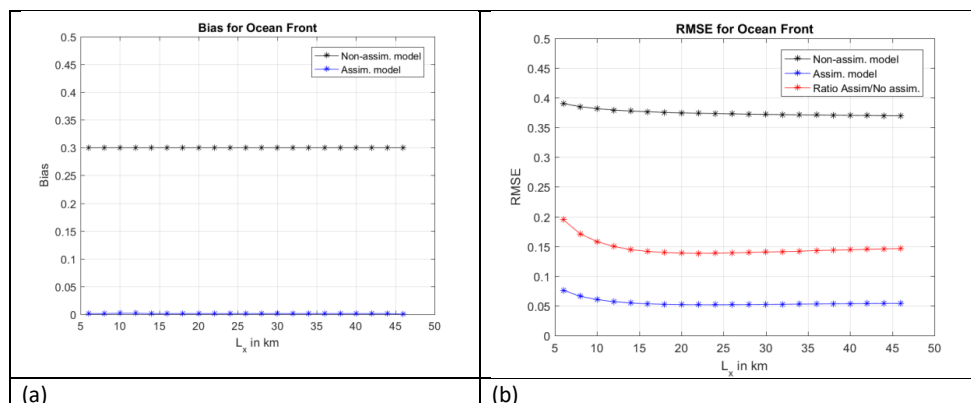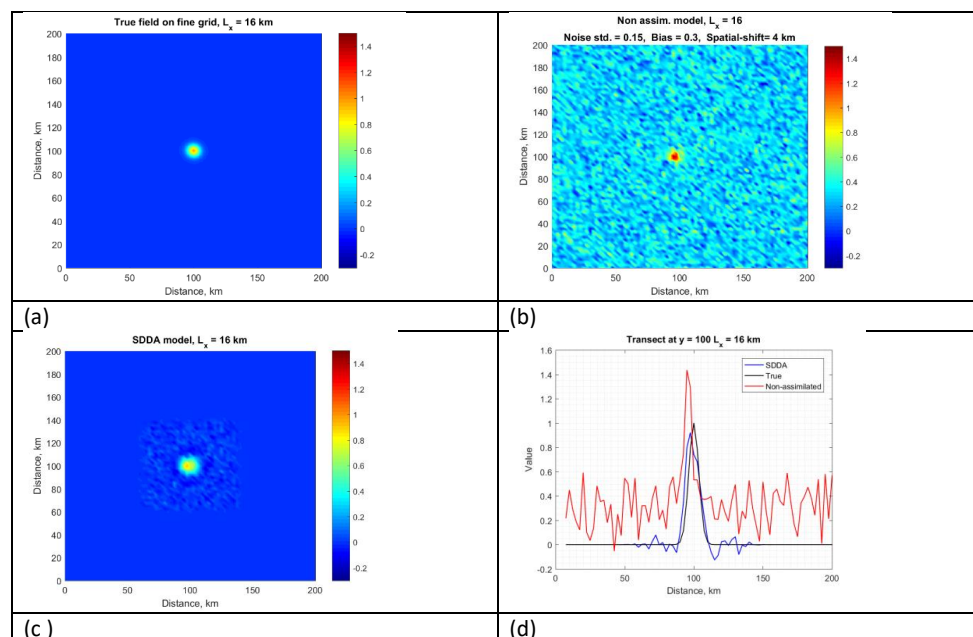289 sharpness of the front before and after data assimilation.

290 **Fig. 4.** *Plots of bias (a) and RMSE (b) for the Ocean Front case with different front sharpness for non-*
291 *assimilated (black) and assimilated (blue) child models.*

292

293 After applying the SDDA method, the bias is practically removed with the remaining values being of
294 the order 5×10^{-3} or less. The RMSE calculated against the true state is more than four times lower than
295 before data assimilation as shown by the red line in Fig.4(b).

296 Similar properties are demonstrated in the example of a single eddy.

297 B) Single eddy
298 Fig.5 shows a map of the true solution F for an axisymmetric mesoscale eddy with the radius of $L_e =$
299 16 km, its representation by the parent coarse model and by the child model before data assimilation.
300 It also shows transects across the eddy centre. The parameters of added errors are the same as in the
301 previous example: normally distributed noise with a standard deviation of 0.15, positive bias of 0.3,
302 and shift of the field by 4 km to the west.

303

304

305

306

307

308
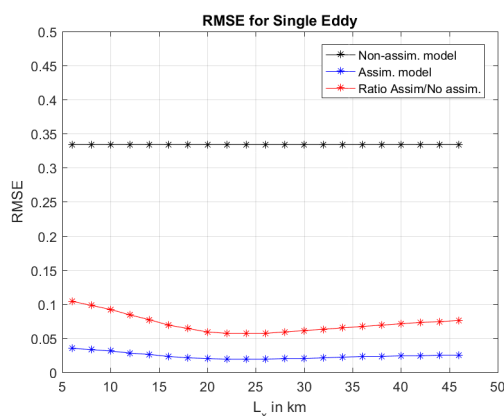
309

Ocean Science
Discussions

310 *Fig.5.* *Results of SDDA assimilation for a single isolated eddy: (a) true field sampled on the fine grid,*
311 *(b) child model before data assimilation, (c) child model after SDDA data assimilation, (d) transects*
312 *showing the same fields as in (a-c).*

313

314 The SDDA method reduces errors nearly to zero outside of the eddy and greatly reduces them inside
315 the eddy (Fig 5 (b,d)). The bias is eliminated from 0.3 before data assimilation to 0.002 or less after
316 assimilation. The reduction of RMSE for various values of eddy radius in the range of 6 to 46 km is
317 shown in Fig 6.

318

319 *Fig. 6* *Dependence of RMSE for different eddy sizes: before data assimilation (black), afte SDDA*
320 *assimilation (blue), the ratio of RMSE after and before data assimilation (red)*
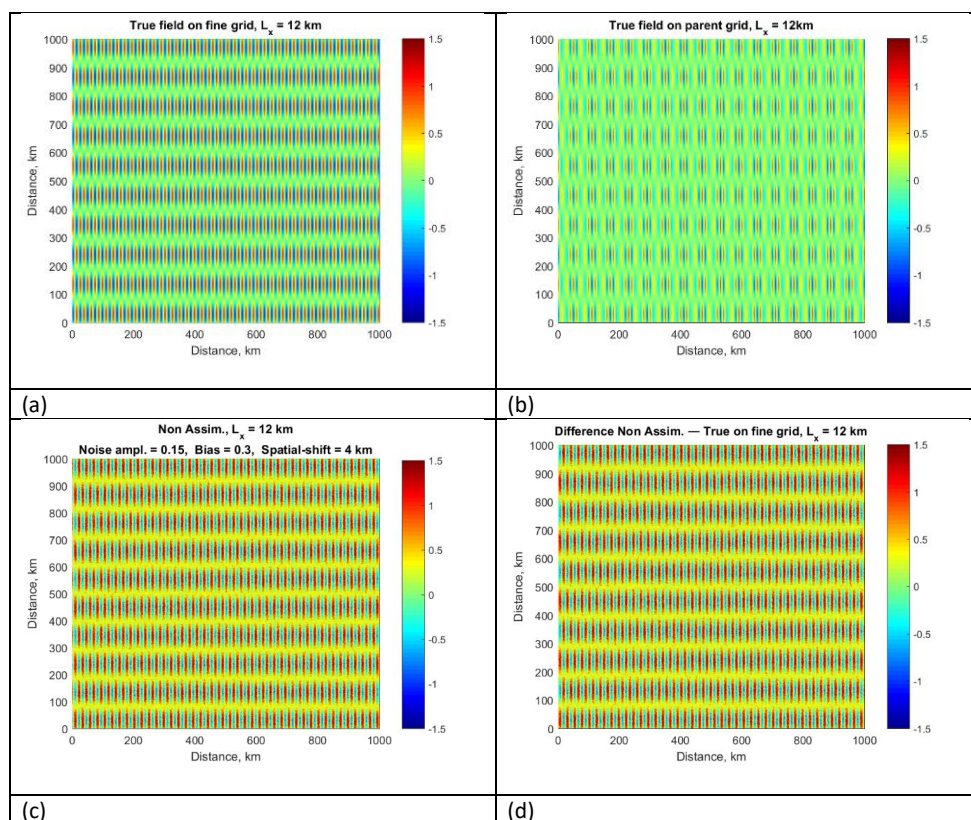
321

322 The next example illustrates the qualities of the SDDA method for a domain of 1000×1000 km
323 packed with anisotropic eddies.

324

325   C) Multiple eddies
326 Fig.7 shows maps of the field variable F from the parent and child models with the following
327 parameters of added errors to simulate child model forecast before data assimilation: normally
328 distributed noise with a standard deviation of 0.15, positive bias of 0.3, and spatial shift of the field by
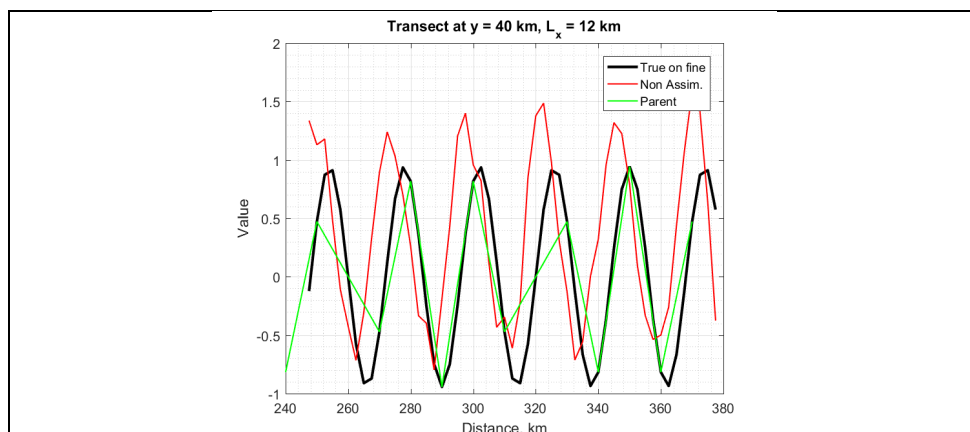329 4 km to the west.

330

Ocean Science
Discussions

Open Access

EGU

Fig. 7. *Maps of the field variable before data assimilation for multiple eddy case with L_x =12km: (a) True field on fine grid, (b) True field on parent grid, (c) Non-assimilated model, (d) Difference between non-assimilated model and true field.*

Fig.7(b) shows that the parent model generally underestimates the true field shown in Fig.7(a) due to representativity errors. The difference between the non-assimilated child model forecast and true state shown in Fig. 7(d) is substantial showing the RMSE =0.61 which is about 30% of the range of the true field. The anatomy of the differences between the true state, and outputs from parent (coarse) and non-assimilated child models are shown on a zoomed-in section of the zonal transect in Fig.8. The largest errors produced by non-assimilated model are due to spatial shift and bias. The errors produced by the parent model are exclusively due its insufficient resolution as we assume that otherwise the parent model is perfect.
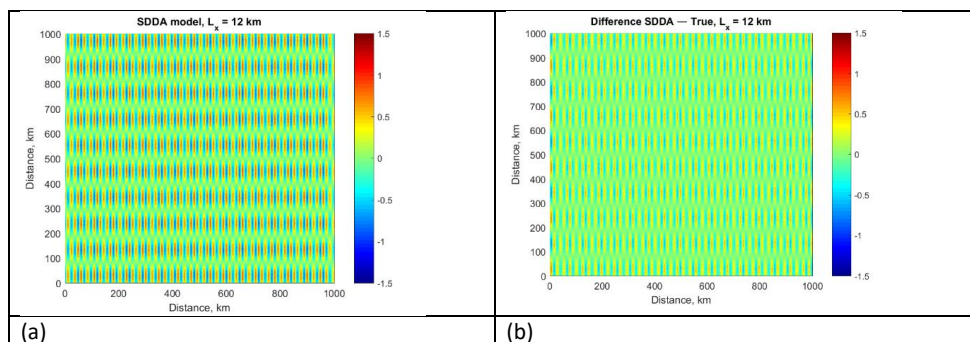
343   **Fig. 8.** *Transects of true field on the fine grid (black) and outputs from the parent (green) and non-*
344   *assimilated child (red) models*

345

346   The SDDA method presented in the previous section is then applied to assimilate the data from the
347   parent model into child model forecast.

348   Fig.9 shows the results of data assimilation (analysis) and a map of differences between the
349   assimilated child model and true field.

350   **Fig.9.** *Maps for the multiple eddies case with width* $L_x = 12$ *km: (a) SDDA assimilated model, (b)*
351   *Difference between SDDA assimilated model and true field. Parameters of the non-assimilated model*
352   *are the same as in Fig.7 and Fig.8.*

353

354   The details of the improvement achieved by SDDA data assimilation are shown on a zoomed-in zonal
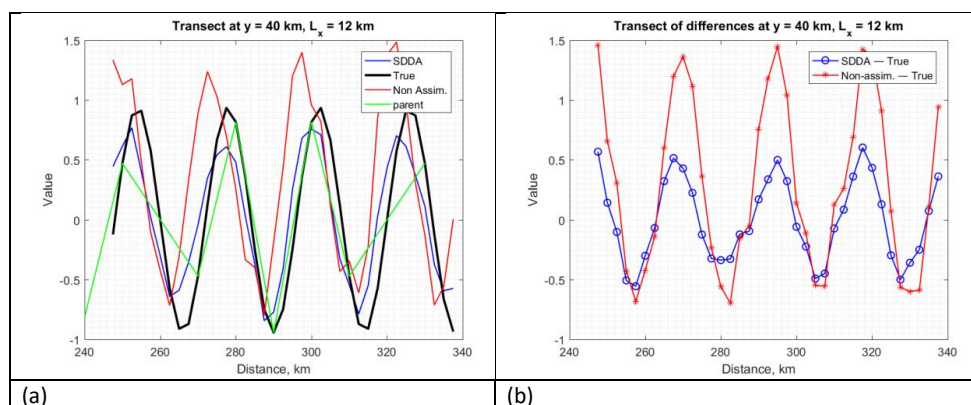355   transect in Fig. 10**.**

(a)   (b)

356   ***Fig.10.*** *Transects at $y = 40$ km for the multiple eddies case with $L_x = 12$ km. (a) Transects of SDDA*
357   *model output (blue), true solution (black), parent model (green) and the non-assimilated child*
358   *model(red); (b) differences between assimilated (blue) and non-assimilated (red) models and the true*
359   *solution.*

360

361   Data assimilation partially reduces the spatial shift which simulated the double penalty effect common
362   to fine-resolution models. The reduction of spatial shift is due to the properties of the SDD component
363   of the SDDA, see (Shapiro et al, 2021) for details. The significant reduction of the bias is due to the
364   assimilation part of the SDDA algorithm, however it also causes some reduction in the amplitudes of
365   the eddies – see Fig. 10(a). The remaining errors shown in Fig. 10(b) are mainly due to incompletely
366   corrected spatial shift and random noise in the non-assimilated model. The RMSE of the assimilated
367   model is lower at 0.25 vs. a non-assimilated value of 0.61, and the bias is reduced by orders of
368   magnitude from 0.3 to - 1.3 10^{-5}.

369   Fig.11 shows the results of the sensitivity analysis with different sizes of eddies and levels of various
370   sources of errors in the non-assimilated child model forecast- noise, bias, and shift.
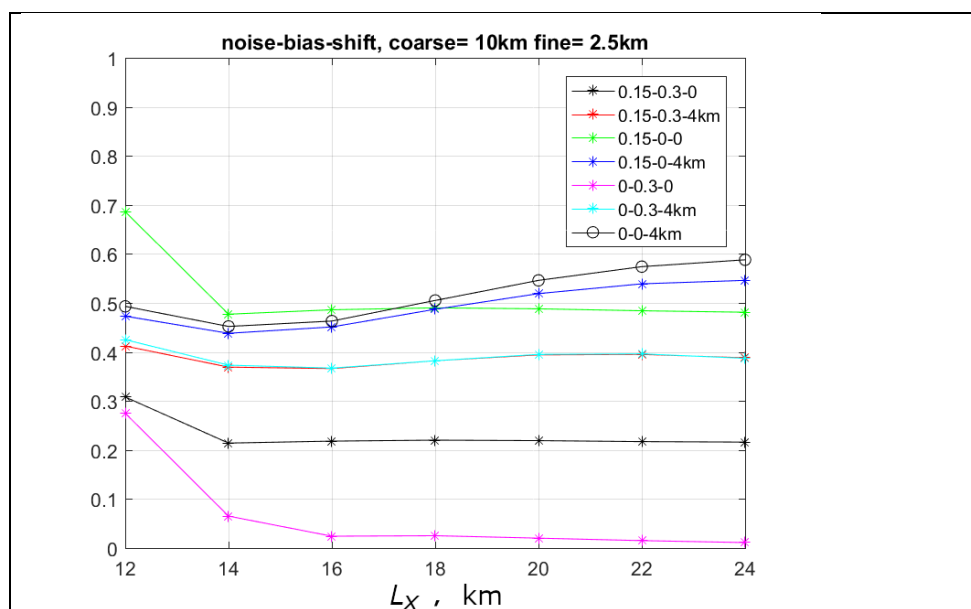
Ocean Science
Discussions

Fig. 11 *Ratio of RSME values of the child model relative to the true solution after and before SDDA*
data assimilation as a function of eddy radius in the zonal direction x. Ratios are calculated for
different combinations of errors generated by the child model – random noise (STD equal to either
zero or 0.15), spatial shift (zero or 4 km), and bias (zero or 0.3). Percentages are shown with respect
to the amplitude of the true signal. The first number in the legend shows the level of noise, the
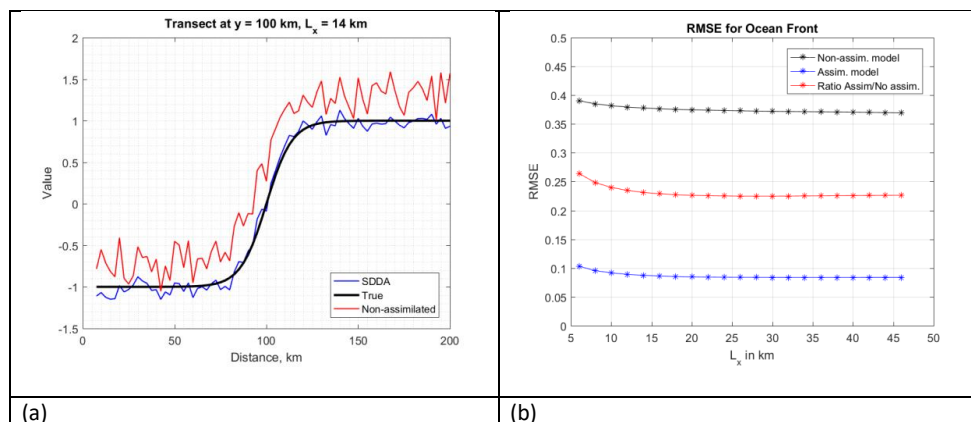*second shows the bias, and the third one shoes the spatial shift.*

In this example, the SDDA method works best when the only source of error is the bias which is
removed nearly completely. The second-best results are achieved when only random noise and bias
are present but not the spatial shift. The RMSE ratio for cases containing errors due to spatial shift
grows slightly at larger eddy sizes. This is due to the fact that a relatively small shift of 4 km does not
distort large eddies to the same extent as small eddies even in the non-assimilated model. Generally,
the curves in Fig.11 show that after application of SDDA data assimilation algorithm the RMSE is
reduced by half or better. Bias between the child and parent models is reduced by orders of
magnitude. Assuming that the parent data assimilating model is unbiased, it means that the fine
resolution model becomes unbiased after application of SDDA data assimilation process.

D) Effect of errors in the parent model

The next example shows the quality of the SDDA method when the parent model is still noisy even
after its own data assimilation cycle. To investigate how such noise impacts on the reduction of RMSE
in the child model when the SDDA data assimilation is applied, a quite large random noise was added
to the parent model output with a standard deviation of 0.1, i.e. 10% of the signal. The results for the
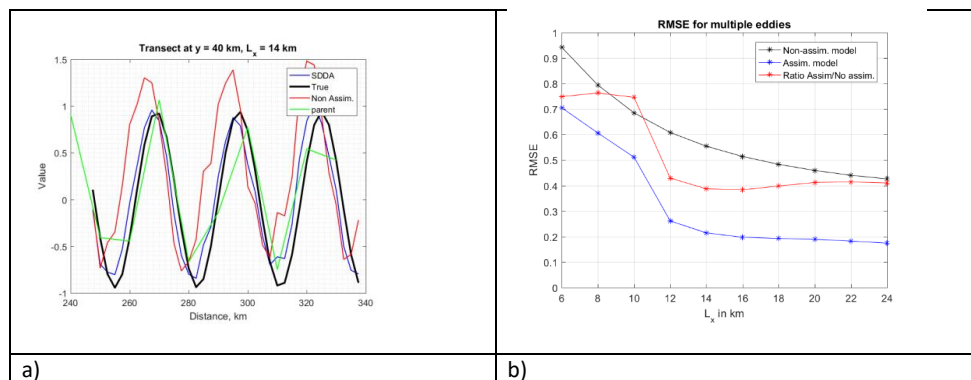ocean front and multiple eddies examples are shown in Fig.12 and Fig.13.

Ocean Science
Discussions

Open Access

394

Fig. 12. *Ocean front case of $L_x = 14$ km, with noisy parent model (added noise STD = 0.1) and noisy non-assimilated child model (added noise STD = 0.15, bias = 0.3 and shift of 4km to the west). (a) Transects at $y = 100$ km of assimilated (blue), non-assimilated (red) child models and the true field (black); (b) RMSE calculated for non-assimilated (black) and assimilated (blue) child models and the ratio between them at different front stepnesses.*

The transect in Fig.12 shows that after SDDA, the front is well represented with small random noise even when the parent model is quite noisy. The improvement is five to ten-fold and is consistent across the range of front widths from 6 to 46 km.

403

404   **Fig.13.** *Multiple eddy case of $L_x = 14$ km with noisy parent model (added noise STD= 0.1) and noisy*
405   *non-assimilated child model (added noise STD = 0.15, bias = 0.3 and shift of 4 km to the west). (a)*
406   *Transects at $y = 100$ km of true field (black), parent (green), assimilated (blue) and non-assimilated*
407   *(red) child models, (b) RMSE for different eddy sizes and the ratio between them.*

408

In case of multiple eddies, the improvement is less dramatic, however for eddy diameters larger than 12 km the RMSE is improved approximately by half. For eddy sizes 10 km and less the improvement due to assimilation of the parent model is limited, which is expected as such eddies are strongly distorted by the parent model of only 10 km resolution.

413

Discussion

414

415 The presented examples show the data assimilation cycle using the SDDA method when the data used
416 for assimilation are not coming from observation as in the common assimilation methods but from a
417 coarser parent model. The examples relate to a synthetic situation when the true values of the state
418 variable are known. The forecast by fine-resolution child model is simulated by sampling the true field
419 on the model's fine grid and adding various sources of errors-random noise, bias, and spatial shift. The
420 output from the coarser parent model is regarded as computer generated 'observations' and is
421 simulated by sampling the true field on the coarser grid to which random noise can be added. In
422 practice, the parent model is assumed to be a good quality, dynamic ocean model assimilating
423 observational data. Therefore, the actual field measurements are used in the SDDA method indirectly
424 – instead of assimilating observation, the child model assimilates data from the coarse model, which
425 in turn assimilates data from observations.

426 The data from the parent model is assimilated by a two-step SDDA process into the forecast produced
427 by the child model to produce the analysis state which is then used as initial condition for the next
428 forecasting cycle. The first step of the SDDA method is the application of Stochastic-deterministic
429 downscaling presented in (Shapiro et al, 2021). As a result, at each fine grid point there are two
430 generally different values of the field variable. The second step is to combine these values separately
431 at each grid point using a suitable data assimilation method. A comprehensive overview of data
432 assimilation methods is given in the paper by Carrassi et al (2018). In this paper, a zero-dimensional
433 Kalman filter is used, similar to what have been done in atmospheric chemistry (Adhikary et al, 2008).
434 For the full 2D or 3D field this is equivalent to have a strictly diagonal background error covariance
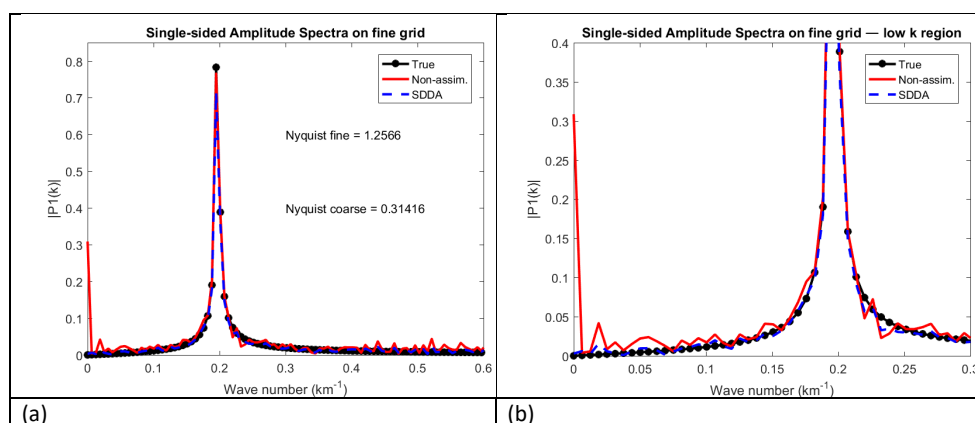435 and 'observation' error covariance matrices.

436 The gain formula used in Kalman filter requires the knowledge of error variances. There are a number
437 of approaches to estimate the variances in an ocean model, e.g. the Canadian, the NMC and the H-L
438 methods. These methods do not separate completely the variances of child model and observations
439 due to model errors and the natural variability of the ocean state among other reasons. In this paper
440 the variances are assessed using a widely used ergodic hypothesis, see e.g. (Stull, 2003), and the
441 ensemble statistical mean at each grid node is estimated by spatial averaging in a small trial area
442 around the node at the same time point. Such scheme is fast and does not consume significant
443 computational resources. If necessary, it can be extended by including data from preceding time
444 points, however these data may not be statistically independent and hence the potential
445 improvement requires further study.

446 The mechanism of improvement of fine model outputs by the SDDA method can be seen from the
447 analysis of spectral characteristics of the downscaled parent model, fine model forecast (before data
448 assimilation) and analysis (after data assimilation) in Fig.14. The bias represented by the peak at zero
449 wavenumber is removed at the data assimilation step of SDDA. Noise in the range below Nyquist
450 wavenumber is reduced by melding the noisy child model forecast with the clean parent model data.
451 It should be noted that the SDD downscaling honours the data of the parent model on the child nodes
452 that coincide with a coarse grid node. A small peak between the coarse and fine grid Nyquist
453 wavenumbers is an artefact created at the SDD step, see (Shapiro et al, 2021) as seen from the
454 spectrum of the downscaled field. Other than this peak, the downscaled field has much lower level of
455 noise than the child model forecast. Reduction of noise at wavenumbers higher than the parent model
456 Nyquist frequency is due to the low level of noise generated at the downscaling step of the process.

Ocean Science
Discussions

Open Access

EGU

457

458

Fig.14. Amplitude spectra on a zonal transect at y=40 km for the example of multiple eddies of $L_x = 16$ km. The STD of noise in the non-assimilated child model is 0.15, bias is 0.3 and the spatial shift is 4 km to the west. The spectra are calculated for the true field, the child model forecast (before assimilation) and analysis (after assimilation). (a) Spectra in the full range of wave numbers, (b) zoomed in area for low wave numbers.

466

The qualities of the SDDA method can be illustrated by comparison with existing data assimilation techniques applied to mode-to-model data assimilation. Let us consider the combination of H-L (Hollingsworth and Lonnberg, 1986) and variational methods, which is considered to be the most commonly used estimation technique, see e.g. (Stewart et al, 2014; Carrassi et al 2018). There are a number of variants of this method, here we applied the 'practical' algorithm described in the textbook by Kalnay (2003), which will be named hereafter as the 'standard' method. The 'standard' method consists of the following stages: calculate innovations (differences between observations and model) at each observational point, estimate the covariances between innovations at different locations, fit the best-fit curve (usually Gaussian) using all covariances except at zero distance, estimate the model and observation error variances using the value where the best-fit curve intersects the r=0 vertical line on the covariance plot. In the practical implementations of this algorithm the model and observation error variances are assumed spatially homogenous (Kalnay, 2003). The model variances and the best-fit curve are used to create the background error covariance **B**-matrix, and the diagonal observational error covariance **R**-matrix. Due to spatial homogeneity, all diagonal elements in B and R matrices are the same but different between matrices. The H-L method requires combination of innovations into spatial bins, so that all innovations within a certain bin are allocated to the same distance from the central point. In practice the covariances are calculated by averaging individual products of innovations over a period of time instead of statistical averaging assuming the ergodic hypothesis (Stull, 2003). In the model-to-model DA approach the role of observations is played by the output from the parent data assimilating model.

For comparison between the SDDA and 'standard' methods we have selected the 'multiple eddies example' (as in section Results, subsection C). The covariances required to build the **B** and **R** matrices

Ocean Science
Discussions

Open Access

489   were estimated as follows. We generated 150 random realisations of the child model outputs which
490   were the sum of for the true field given by Eq (15) with $L_x = 40$ km and $L_y = 105$ km in a subdomain
491   of 300×300 km, and the following errors: (i) random noise with standard deviation of 0.15; (ii) random
492   spatial shift with a standard deviation of 4 km, (iii) constant positive bias of 0.3. Random components
493   were normally distributed with zero mean. Due to significantly larger computing resources required
494   by the 'standard' method compared to SDDA, the 'standard' DA was done in the reduced domain
495   300×300 km instead of 1000×1000 km for the SDDA method. The bins required for the H-L method
496   were 1 km in size.

497   The innovations were computed as $\boldsymbol{d} = \mathbf{H}\boldsymbol{x^b} - \boldsymbol{y}$, where $\mathbf{H}$ is a subsampling matrix operator
498   composed only of ones and zeroes. Following the 'standard' method we used a simple isotropic
499   Gaussian function for the covariance of innovations at every parent model grid point x:

500

501    $C_x(r) = a\, exp(-r^2/D^2)$ (16)

504
505
502   where a and D are fitting parameters. The binned covariances and the fitting curve are shown in
503   **Fig.15.**

506

507

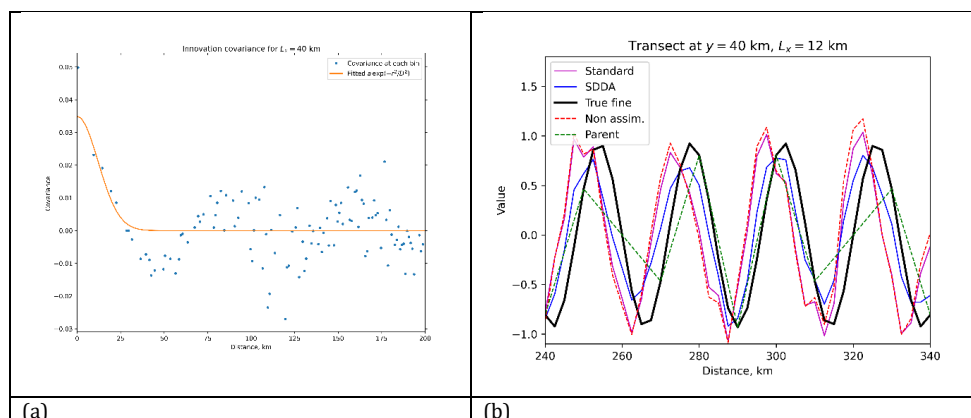

(a)   (b)

508

509   ***Fig. 15.** (a) Estimates of background and observational error covariances for multiple eddy field with*
510   *$L_x = 40$ km and $L_y = 105$ km and the corresponding fitting curves. The error variances estimated at*
511   *r=0 are used as diagonal elements of the background matrix **B** (equal to 0.031) and the*
512   *'observational' matrix **R** (equal to 0.016), the length scale D=17km; (b) the zonal transect at $y =$*
513   *40 km showing the true field (black line), parent model (dashed green), child model noisy forecast*
514   *before DA (dashed red), analysis state after 'standard' DA (magenta), and analysis after SDDA (blue).*

515

516   Despite the parent model is assumed 'perfect', i.e. it has no errors relative to the true field, the H-L
517   method gives the error variance of approximately 50% of the noisy child model value. The **B** matrix is
518   then constructed using a variance (the diagonal) equal to 0.031 (see Fig 15 (a)) and the Gaussian

519 formula Eq (16), and the diagonal matrix $\mathbf{R}$ is build using a value for its variance (diagonal elements)
520 of 0.016.

521 The $\mathbf{B}$ and $\mathbf{R}$ matrices were used to carry out the variational DA cycle to the simulated child model
522 forecast with the following parameters: multi eddy field with $L_x = 12$ km, $L_y = 105$ km, random
523 noise with STD=0.15, bias=0.3, spatial shift=4 km to the west. The analysis state is estimated using the
524 following equations of the 'standard' method (Kalnay, 2003, page 155)

525 $$x^a = x^b + \mathbf{W}[y - H(x^b)] = x^b + \mathbf{W}d$$

526 $$\mathbf{W} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}$$

527 The transect in Fig.15(b) shows the analysis state for eddies with $L_x = 12$ km and $L_y = 105$ km
528 obtained by the 'standard' method together with the true solution, parent model output, noisy child
529 model forecast (before DA) and, for comparison, the analysis state obtained by the SDDA method.
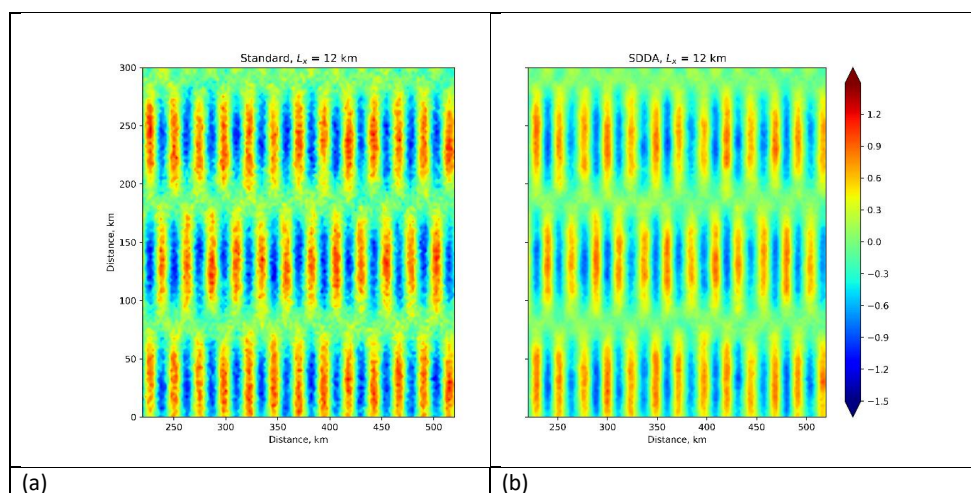
530 The bias and RMSE relative to the true solution for the child model forecast and the analysis state
531 after 'standard' and SDDA data assimilation process is shown in **Table 1**

532 **Table 1.** Errors in the child model outputs before and after Data Assimilation

	Forecast (before DA)	'standard' DA	SDDA
Bias	0.300	-0.0066	0.000
RMSE	0.608	0.5268	0.250

533

534 Fig. 16 shows the map of the analysis state produced by the SDDA and the standard methods. The
535 errors are removed more efficiently in the areas of low values of the field variable F.

(a)   (b)

536 *Fig. 16. The child model output after DA using (a) the 'standard' method, (b) the data assimilation*
537 *with Stochastic-Deterministic Downscaling (SDDA). The SDDA produces less noise.*

538

539 The analysis presented above shows that the SDDA method is not the only one which can be applied
540 to model-to-model data assimilation, the 'standard' method also gives reasonable results. However,
541 when compared with the 'standard' DA method, the SDDA gives better accuracy- stronger reduction

Ocean Science

Discussions

Open Access

EGU

542   in the RMS errors and a complete removal of bias. The SDDA also is more computationally efficient.
543   We had no restrictions or limitations in computing the analysis for 1000×1000 km domain at 2.5 km
544   resolution when using the SDDA on our office PC. However, we were not able to carry out the
545   'standard' DA for a domain greater than 300×300km on the same PC due to computing resource
546   restrictions. In terms of speed of calculations, it took about 2 minutes to complete one full DA cycle
547   including calculation of the covariance matrix and weighting coefficients using the SDDA method in
548   the multiple eddy example. On the other hand, the standard method took 24 minutes, including the
549   creation of **B** and **R** matrices and calculation the analysis state.

550   Conclusion
551   This paper suggests a data assimilation approach where the data are assimilated into a high-resolution
552   model from a coarser good quality data assimilating model, not directly from observations. An
553   efficient and simple algorithm for model-to-model variational data assimilation method named Data
554   Assimilation with Stochastic-Deterministic Downscaling (SDDA) is developed. The theoretical
555   background behind the SDDA algorithm is discussed, and its application is illustrated in a number of
556   idealised synthetic situations which resemble real world practice in fine-resolution ocean modelling.
557   The results demonstrate that the model-to-model data assimilation is an efficient way of improving
558   the accuracy of fine resolution model. Such approach allows to avoid a repetition of a complex and
559   resource-hungry assimilation of actual observations which has already been done in the parent model.
560   It is likely that the same basic idea of model-to-model data assimilation would work also for other
561   methods currently used in observational data assimilation. In this paper, the SDDA was compared with
562   a commonly used Hollingsworth-Lönnberg method and shown to be more accurate and
563   computationally significantly less expensive.

564   Code and data availability
565   For code availability please contact the corresponding author.

566   Author contributions
567   GIS conceptualised and designed the study, developed the methodology,
568   contributed to software development, performed the analysis, and drafted the
569   paper. JMGO contributed to software development, analysis of the results, selection
570   of appropriate parameters and writing of the paper.

571   Competing interests
572   The authors declare that they have no conflict of interest.

573   Acknowledgements
574   The authors are thankful to the members of Plymouth Ocean Forecasting Centre for
575   their support.

576

577   References
578   Adhikary, B., Kulkarni,S., Dallura A., et al., 2008.; A regional scale chemical transport modeling of
579   Asian aerosols with data assimilation of AOD observations using optimal interpolation technique
580   Atmospheric Environment 42, 8600–8615.

Ocean Science
Discussions

EGU

581   Bell, M. J., R. M. Forbes, A. Hines, 2000 Assessment of the FOAM global data assimilation system
582   for real-time operational ocean forecasting. J. Mar. Sys., 25, 1-22.

583   Bouttier,F. and Courtier,P., 1999. Data assimilation concepts and methods, available at
584   https://www.ecmwf.int/en/elibrary/16928-data-assimilation-concepts-and-methods, last
585   access: 05 August 2021

586   Bruciaferri, D., Shapiro, G.I. & Wobus, F. 2018. A multi-envelope vertical coordinate system for
587   numerical ocean modelling. Ocean Dynamics 68, 1239–1258 https://doi.org/10.1007/s10236-
588   018-1189-x).

589   Carrassi, A., Bocquety ,M., Bertino, L. and Evensen,G. 2018. Data assimilation in the geosciences:
590   An overview of methods, issues, and perspectives. WIRES climate change, v9, No5,
591   https://doi.org/10.1002/wcc.535

592   Ciavatta, S, Brewin, R.J., Skákala, J., Polimene, L., de Mora, L., Artioli, Y., Allen, J.I. 2018.
593   Assimilation of Ocean-Color Plankton Functional Types to Improve Marine Ecosystem
594   Simulations. JOURNAL OF GEOPHYSICAL RESEARCH. OCEANS, ISSN: 2169-9291, doi:
595   10.1002/2017JC013490

596   CMEMS, 2021, available at https://marine.copernicus.eu/) last access: 20 May 2021

597   DARC, 2021 available at https://research.reading.ac.uk/met-darc/aboutus/what-is-data-
598   assimilation/ last access: 05 August 2021

599   Dobricic, S. and Pinardi, N. 2008. An oceanographic three-dimensional variational data
600   assimilation scheme, Ocean Modelling., 22, 89–105,

601   Dufour et al., 2015: Role of mesoscale eddies in cross-frontal transport of heat and
602   biogeochemical tracers in the Southern Ocean. Journal of Physical Oceanography, 45(12),
603   DOI:10.1175/JPO-D-14-0240.1

604   Fedorov, K.N, 1986. The Physical Nature and Structure of Oceanic Fronts, Springer, 333 pages,
605   ISBN 978-0-387-96445-4.

606   Gandin, L. S. 1959. The problem of optimal interpolation, Scientific papers, Main Geophysical
607   Observatory, 99, 67–75.

608   Gandin, L. S. 1965. Objective analysis of meteorological fields. Translated from the Russian,
609   Jerusalem, Israel Program for Scientific Translations, 242 pp..

610   GFDL, 2020: available at: https://www.gfdl.noaa.gov/high-resolution-modeling/, last access:
611   6 July 2020.

612   Ghil,M., Malanotte-Rizzoli, P., 1991. Data Assimilation in Meteorology and Oceanography,
613   Advances in Geophysics 33, 141-266, DOI: 10.1016/S0065-2687(08)60442-2;

614   Hollingsworth, A. and Lönnberg, P. 1986. The statistical structure of short-range forecast errors
615   as determined from radiosonde data, Part 1: The wind field, Tellus A, 38, 111–136.

616   Ide, K., Courtier,P., Ghil, M. and Lorenc ,A.C., 1997. Unified Notation for Data
617   Assimilation: Operational, Sequential and Variational J. Met. Soc. Japan, Vol. 75, No. 1B,
618   pp. 181-189.

Ocean Science
Discussions

619 Kalnay ,E. 2003. Atmospheric modeling, data assimilation and predictability , Cambridge
620 University Press, 369 pp.

621 Kubryakov1, A.I., Korotaev, G.K., Dorofeev, V.L. et al, 2012. Black Sea coastal forecasting system.
622 Ocean Science., 8, 183–196.

623 Lorenc, A. C. 1986. Analysis methods for numerical weather prediction, Q. J. Roy. Meteor. Soc.,
624 112, 1177–1194.

625 Mellor ,G.L., and A. F. Blumberg, 1985. Modeling vertical and horizontal diffusivities with the
626 sigma coordinate system. Mon. Wea. Rev., 113, 1379–1383.

627 Meunier,T., E. D. Barton, B. Barreiro, R. Torres Upwelling filaments off Cap Blanc: Interaction of
628 the NW African upwelling current and the Cape Verde frontal zone eddy field?Journal of
629 Geophysical Research Atmospheres, vol. 117, C08031, 2012,
630 https://doi.org/10.1029/2012JC007905.

631 Moore, A.M., Martin, M.J., Akella, S. et al ,2019. Synthesis of Ocean Observations Using Data
632 Assimilation for Operational, Real-Time and Reanalysis Systems: A More Complete Picture of the
633 State of the Ocean Frontiers of Marine Science, https://doi.org/10.3389/fmars.2019.00090

634 NEMO, 2021, available at https://www.nemo-ocean.eu/ last access: 16 June 2021

635 Robinson, A. R. (Ed.), 1983. Eddies in marine science. Springer-Verlag Berlin Heidelberg, 612 pp.,
636 ISBN 978-3-642-69003-7.

637 ROMS, 2021, available at https://www.myroms.org/; last access: 16 June 2021

638 Shapiro, G.I., Stanichny, S.V, Stanychna, R.R. 2010. Anatomy of shelf–deep sea exchanges by a
639 mesoscale eddy in the North West Black Sea as derived from remotely sensed data. Remote
640 Sensing of Environment , 114 , No 4 , 867-875

641 Shapiro, G. I., Gonzalez-Ondina, J. M., and Belokopytov, V. N., 2021. High-resolution stochastic
642 downscaling method for ocean forecasting models and its application to the Red Sea dynamics,
643 Ocean Sci., 17, 891–907, https://doi.org/10.5194/os-17-891-2021.

644 Stewart, L. M. , Dance,S.L., Nichols, N.K., Eyre, J.R., Cameron, J. , 2014. Estimating interchannel
645 observation-error correlations for IASI radiance data in the Met Office system. Quarterly Journal
646 of the Royal Meteorological Society, Volume140, Issue 681 Part B, Pages 1236-1244,
647 https://doi.org/10.1002/qj.2211.

648 Stull, R. B, 2003. An introduction to boundary layer meteorology, Kluwer, DOI: 10.1007/978-94-
649 009-3027-8, 670 pp.

650 Zingerlea, C. and Nurmib, P., 2008. Monitoring and verifying cloud forecasts originating from
651 operational numerical models, Meteorol. Appl., 15, 325–330.

652