Comparison of stochastic parameterizations in the framework of a coupled ocean-atmosphere model – Response to the 1st reviewer

Jonathan Demaeyer [1] and Stéphane Vannitsem [1]

[1]Institut Royal Météorologique de Belgique, Avenue Circulaire, 3, 1180 Brussels, Belgium

1 General comments

We thank the reviewer for his/her carefull reading of the manuscript. We will first adress the general comments and then address the proposed suggestions.

First, a comparison with a data-driven method would have been very interesting too. It could be easily done in the framework of the software code given as supplementary material. However, we wanted that the article remained focused on *first principle* methods and their technical details.

Secondly, the reviewer indicates that the consideration of the case $\varepsilon = \delta = 1$ only is another drawback of the study. We wanted to focus on the methods derivation within the ocean-atmosphere model, and their performance in its original form. The extension to other values has been done in Demaeyer and Vannitsem (2018) with a simpler system. It can of course be reproduced here but it is left for future works.

2 Modifications

We now address the proposed suggestions:

- p.2 l.22: over which $\rightarrow$ on which

  Answer: Ok.

- p.2 footnote: over which $\rightarrow$ on which

  Answer: Ok.

- p.3 l.5: mode $\rightarrow$ modes

  Answer: Ok.

– p.4 l.20: "It results": what does "it" refer to? Please clarify.

Answer: Indeed, the procedure was not correctly explained. We have modified the text according to:

> *By recasting these expansions into the partial differential equations of the model, one obtains a set of ODEs for its coefficients* $\boldsymbol{Z} = (\{\psi_{a,i}\}, \{\theta_{a,i}\}, \{\psi_{o,j}\}, \{\theta_{o,j}\})_{i\in\{1,\dots,n_a\},\, j\in\{1,\dots,n_o\}}$ *:*

– p.6 eq.17: there is an ε and ε^2 missing in the second and third term on the RHS

Answer: Ok.

– p.7 l.5: Vanden-Eijnden

Answer: Ok.

– p.7 l.10: In term $\rightarrow$ In terms

Answer: Ok.

– p.7 eq.23-24: note that in homogenization theory F_X and F_Y can also depend on both $\boldsymbol{X}$ and $\boldsymbol{Y}$

Answer: We have added this comment as footnote in the text:

> *Note that in homogenization theory* F_X *and* F_Y *can also depend on both* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$*, a possibility that was not considered here in order to effectively compare the MTV and WL parameterizations.*

– p.8 eq.32-33: The choice of decomposition is arbitrary without information of the size of different terms in the dynamical equation. If you agree, please insert a comment (here or elsewhere) that other decompositions may improve or decrease the performance. Could you also comment why you consider the assumptions of a time-scale separation or weak coupling to be valid in the current system?

Answer: We have change this section according to:

This decomposition can be chosen arbitrarily since the only requirement is that F_X depends solely on $\boldsymbol{X}$. However, in the following (and in the provided code), we will consider that the resolved-unresolved components form a coupled system, with "maximal" uncoupled dynamics. In this view for both parameterizations, the decomposition of the $\boldsymbol{X}$ component is the same:

$$F_X(\boldsymbol{X}) = \boldsymbol{H}^X + \mathbf{L}^{XX} \cdot \boldsymbol{X} + \mathbf{B}^{XXX} : \boldsymbol{X} \otimes \boldsymbol{X}$$

and

$$\Psi_X(\boldsymbol{X}, \boldsymbol{Y}) = \mathbf{L}^{XY} \cdot \boldsymbol{Y} + \mathbf{B}^{XXY} : \boldsymbol{X} \otimes \boldsymbol{Y} + \mathbf{B}^{XYY} : \boldsymbol{Y} \otimes \boldsymbol{Y} \; .$$

This choice was also consider in Franzke et al. (2005), dealing with the MTV parameterization. It is worth noting that other decompositions may improve or decrease the performance of the parameterization.

In addition, we do not assume that the resolved-unresolved components decomposition possesses a weak coupling or a timescale separation since these can be set arbitrarily by the user of the code. If the decomposition is for instance a resolved ocean and an unresolved atmosphere, then a timescale separation exists (weak-coupling is less obvious due to the temperature scheme). On the other hand, for all the cases considered in the result section, the decomposition is made in the atmosphere, where no weak-coupling or timescale separation (spectral gap) is found.

So the assumptions are valid depending on the "experiment" which is made. We have developed the parameterizations *based* on these assumptions, but then one can "push" these methods over their limits, in order to see if they would satisfy more difficult applications, like the parameterization of some atmospheric large scales. This is what we have tried to do.

– p.9 l.4-10: Either method should work as long as the fast dynamics has an ergodic invariant measure (with some extra assumptions). Please comment on the reason for choosing two different decompositions for the two methods.

Answer: We have modified this section according to:

The definition of F_Y and Ψ_Y is also arbitrary, but it is of particular importance since it is the measure of the system whose tendencies are given by $F_Y(\boldsymbol{Y})$ over which the averages are performed (Demaeyer and Vannitsem, 2018). A requirement is thus that the dynamics $\dot{\boldsymbol{Y}} = F_Y(\boldsymbol{Y})$ has an ergodic invariant measure. In the framework of the WL method, it is natural to consider the measure of the intrinsic $\boldsymbol{Y}$-dynamics as:

$$F_Y(\boldsymbol{Y}) = \boldsymbol{H}^Y + \mathbf{L}^{YY} \cdot \boldsymbol{Y} + \mathbf{B}^{YYY} : \boldsymbol{Y} \otimes \boldsymbol{Y}$$

to perform the averaging.

In the framework of the MTV method, the measure of the $O(1/\delta^2)$ singular system $\dot{\boldsymbol{Y}} = F_Y(\boldsymbol{Y})$ is used for the averaging and it is usually assumed that the quadratic $\boldsymbol{Y}$-terms of the unresolved component tendencies represent the fastest timescale of the system (see Majda et al. (2001); Franzke et al. (2005)):

$$F_Y(\boldsymbol{Y}) = \mathbf{B}^{YYY} : \boldsymbol{Y} \otimes \boldsymbol{Y}.$$

and those are the ones over which the averaging has to be done.

In both cases, the others terms belong then to Ψ_Y.

It is interesting to note that there is no a priori justification for one or the other assumption. For both parameterization methods, the decomposition of the unresolved dynamics could be based on Eq. (34) or on Eq. (35). The code provided as supplementary materials allows to select the F_Y-dynamics as either Eq. (34) or Eq. (35).

The MTV and WL parameterizations described above are presented in more details in the appendices A and B, respectively.

We have chosen the dynamics of Eq. (34) to test the case with maximal uncoupled resolved-unresolved components. The test with Eq. (35) can easily be done with the code provided. In addition, it has already been used in Franzke et al. (2005).

Following the reviewer comment, we have also noticed that the line 2 p. 10 :

The $F_Y(\boldsymbol{Y}) = (F_{Y,a}(\boldsymbol{Y}), F_{Y,o}(\boldsymbol{Y}))$ function can be specified by either Eq. (35) or (34) (only (34) for the WL parameterization).

should be modified. The text inside the parenthesis should be removed since the dynamics (35) could also be used to perform the WL parameterization. The code should allow this dynamics to be used by this parameterization, it will be changed in a future version.

– p.9 l.13-14: “we consider these two different assumptions”: this is contradictory to p.10 l.27: “we will here consider only the dynamics of (34)”.

Answer: Indeed, there is a contradiction. It has been solved by removing the sentence “we consider these two different assumptions…” (p.9 l.13-14).

– p.10 l.1: please add a few words explaining the problem of “dead” scales

Answer: We have decided to remove the expression *“dead” scales* which we have in fact not found in the literature. In place we have written :

> […] *or to increase the small-scale variability to address the problem of scales that are passively slaved and whose variability comes uniquely from their interactions with others.*

– p.13 l.6: setting $\rightarrow$ settings

Answer: Ok.

– p.30 l.12: “We will also assume…”: this is not an extra assumption if you already consider a Gaussian Ornstein-Uhlenbeck process.

Answer: Indeed, we have thus modified and moved this sentence just after equation (A1):

> *This dynamics thus generates Gaussian distributions, such that the odd-order moments vanish and that the even-order moments are related to the second-order one.*

– p.30 eq. A5 and A7: the $1/\delta^2$ factors here should be 1

Answer: These factors in front of the time integrals are correct. This is due to the fact that we use the $1/\delta^2$ scaled dynamics Eq. (A2) to compute the average. Therefore, as stated p.31 line 1-2, the integrals over the lagtime s is of order δ^2 and the r.h.s. of Eq. (A5) and (A7) are of order 1. The integrals are of order δ^2 since the exponential decorrelation of the scaled dynamics (A2) are scaled accordingly. In addition, the integrated correlation matrices $\mathbf{\Sigma}$ and $\mathbf{\Sigma}_2$ are also $O(\delta^2)$ matrices, as stated p. 34 line 4.

In principle, the unscaled dynamics $\dot{\boldsymbol{Y}} = F_Y(\boldsymbol{Y})$ corresponding to the $\mathcal{L}_1$ operator should be used to compute the average. But here, the need for a code that allows to compute both the MTV and WL parameterization, with the same initialization files, led to this particular formulation. The text of the appendix try to stay close to how the code performs the computation of the parameterization.

To clarify this for the reader, we added the following footnote p.31 on line 2 :

> *It is due to the fact that we use directly the measure $\tilde{\rho}_Y$ of the $O(1/\delta^2)$ dynamics (A2) to performs the averaging, and not the measure of the dynamics $\dot{\boldsymbol{Y}} = F_Y(\boldsymbol{Y})$.*

– p.31 l.17: the backward KE describes the evolution of expectation values, not densities

Answer: To our knowledge, the backward Kolmogorov equation:

$$-\frac{\partial}{\partial t}p(x,t) = b(x,t)\,\frac{\partial}{\partial x}p(x,t) + \frac{1}{2}\,\Sigma(x,t)\,\frac{\partial}{\partial x^2}p(x,t)$$

describes the evolution of the probability density $p(x,t)$ with a final condition $p(x,s) = u_s(x)$ with $t < s$. But it is true that the Feynman-Kac formula allows for its interpretation in terms of the expectation of the final condition $u_s(x)$ over all paths starting from x at time t, see for instance Pavliotis (2014). In that sense, the probability density of a SDE is always an expectation over paths, the transition probability giving the "weight" (measure) of each path. Therefore both formulations are correct and we stick with the word "density" but we add "probability" in front for the sake of clarity. See also Gardiner (2009) and Risken (1996).

– p.3p.32 l.5-6: incomplete sentence

Answer: Has been completed by adding the word "vanishes" at the end.

– p.32 footnote: Additional to the diagonality, one needs assumption A4 of Majda et al. (2001)

Answer: A diagonal matrix $\mathbf{A}$ in Eq. (A1) implies that assumptions A4 is fulfilled, due to the following property of the quadratic terms in quasi-geostrophic systems:

$$\frac{\partial}{\partial Y_i} B^{YYY}_{i\;j\;k}\, Y_j\, Y_k = 0 \qquad \forall j,k \quad . \tag{1}$$

Indeed, if $\mathbf{A}$ is diagonal then the multidimensional Ornstein-Uhlenbeck process is made of uncorrelated one-dimensional Ornstein-Uhlenbeck processes, and thus:

$$\left\langle B^{YYY}_{i\;j\;k}\, Y_j\, Y_k \right\rangle = 0 \qquad \forall i,j,k \tag{2}$$

which is precisely the assumption A4. In other words, in this case the covariance matrix of the multidimensional O-U process is also diagonal and with the special property (1), it implies that the assumption A4 is fulfilled.

To explain this better in the article, we removed the footnote and modified the text according to:

The first solvability conditions is obviously satisfied but the the second one is not necessarily satisfied since

$$\mathbb{P}\mathcal{L}_2\rho_0 = \left(\mathbf{B}^{XYY} : \left\langle \boldsymbol{Y} \otimes \boldsymbol{Y} \right\rangle_{\tilde{\rho}_Y}\right) \cdot \partial_{\boldsymbol{X}}\rho_0 = \left(\mathbf{B}^{XYY} : \boldsymbol{\sigma}_Y\right) \cdot \partial_{\boldsymbol{X}}\rho_0$$

where $\mathbb{P}$ is the expectation with respect to the measure of the process (A1) and $\boldsymbol{\sigma}_Y$ is its covariance matrix. If the matrix $\mathbf{A}$ in Eq. (A1) is considered to be diagonal, as in Majda et al. (2001) and in Franzke et al. (2005), then it is satisfied. Indeed, in this case $\boldsymbol{\sigma}_Y$ is diagonal and $\mathbb{P}\mathcal{L}_2\rho_0 = 0$ due to the following property of the quadratic terms in the model:

$$\frac{\partial}{\partial Y_i} B^{YYY}_{i\,j\,k}\, Y_j Y_k = 0 \quad .$$

However, here we do consider the general case: $\mathbf{A}$ is not diagonal and thus we can have $\mathbb{P}\mathcal{L}_2\rho_0 \neq 0$. It then indicates that $1/\delta$ effects have to be taken into account in the parameterization.

– p.34 eqs A34: check the factors $1/\delta^2$

Answer: As stated above, the "integrated" correlation matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^2$ are of order δ^2, therefore these factors are correct.

– p.36 l. 11-12: Have any tests been performed to conclude that the process can be approximated with an Ornstein-Uhlenbeck?

Answer: No other tests than a rapid evaluation of the PDFs of the processes and a check of the fast exponential decorrelation properties have been done. In general, the PDFs looked Gaussian. Since then, we have computed the standardized third and fourth moments which show that indeed the statistics of the unresolved processes

$$\dot{\boldsymbol{Y}} = F_Y(\boldsymbol{Y})$$

is nearly Gaussian for every case that we considered. We comment now on that at the end of the introduction of section 4:

In each case, we have also checked the statistics of the dynamics (34) and have concluded that it is nearly Gaussian.

However, as stated in the conclusion, the parameterization methods could be extended to deal with the non-Gaussian case.

We thank again the reviewer for his/her suggestions.

References

Demaeyer, J. and Vannitsem, S.: Stochastic Parameterization of Subgrid-Scale Processes: A Review of Recent Physically Based Approaches, in: Advances in Nonlinear Geosciences, pp. 55–85, Springer, 2018.

Franzke, C., Majda, A. J., and Vanden-Eijnden, E.: Low-order stochastic mode reduction for a realistic barotropic model climate, Journal of the atmospheric sciences, 62, 1722–1745, 2005.

Gardiner, C. W.: Handbook of stochastic methods, Springer Berlin, fourth edn., 2009.

Majda, A. J., Timofeyev, I., and Vanden Eijnden, E.: A mathematical framework for stochastic climate models, Communications on Pure and Applied Mathematics, 54, 891–974, 2001.

Pavliotis, G. A.: Stochastic processes and applications: Diffusion processes, the Fokker-Planck and Langevin equations, vol. 60, Springer, 2014.

Risken, H.: Fokker-planck equation, in: The Fokker-Planck Equation, pp. 63–95, Springer, 1996.