RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

Doctorat de l'Université de Toulouse

délivré par l'ISAE-SUPAERO

Prise en compte des corrélations d'erreur d'observation en assimilation variationnelle de données océaniques: application aux données altimétriques

---

Thèse présentée et soutenue, le 20 février 2025 par

Olivier GOUX

École doctorale
SDU2E - Sciences de l'Univers, de l'Environnement et de l'Espace

Spécialité
Océan, Atmosphère, Climat

Unité de recherche
CECI - Climat, Environnement, Couplages et Incertitudes / CERFACS

Thèse dirigée par
Anthony WEAVER et Oliver GUILLET

Composition du jury
Mme Nadia FOURRIÉ, Présidente, CNRM
M. Andrew MOORE, Rapporteur, University of California, Santa Cruz
Mme Sarah DANCE, Rapporteure, University of Reading
M. Arthur VIDARD, Examinateur, INRIA
M. Massimo BONAVITA, Examinateur, European Centre for Medium-Range Weather Forecasts
M. Anthony WEAVER, Directeur de thèse, CERFACS / CECI UMR 5318
M. Oliver GUILLET, Co-directeur de thèse, CNRM

Membres invités
Mme Selime GÜROL, Co-Encadrante, CERFACS/CECI UMR 5318

Acknowledgements

First, I would like to thank the reviewers of this manuscript, whose feedback and interest in this work is very motivating. I also would like to thank the people who supervised me during this PhD: Anthony Weaver, Selime Gürol, Andrea Piacentini and Oliver Guillet. They taught me so much about mathematics or oceanography, while also proving equally inspiring in their human qualities. I feel incredibly grateful to have been able to work with researchers whose passion was so contagious during this PhD. I do not think any of this work would have been possible without their continuous support through both successes and setbacks. Beyond my supervisors, I found at CERFACS a great number of people with that were eager to share knowledge, enthusiasm, and coffee breaks, all of which contributed significantly to the results presented in this manuscript. The ALGO team was very welcoming from the very first day, and managed to show the "physics-inclined" intern I once was that mathematics were equally interesting. The fact that a chapter of this manuscript is based on a publication in a numerical linear algebra journal - one that I thoroughly enjoyed writing — is a testament to to their conviction and pedagogy. I am also grateful to many people in the GLOBC team for indulging my curiosity about to their work and being extremely friendly and welcoming while doing so. I want to thank in particular Anne, Camille, Juliette, Quentin and Théo, whose support in my fatigue-induced antics was crucial for maintaining morale while writing this manuscript. This work was also made possible by the support and responsiveness of the CSG team and the administration at CERFACS. I am also very grateful to the members of the gender equality group for the many eye-opening experiences they provided to me. I hope that their work will one day allow everybody to have the same opportunities I had, and to feel as welcome as I did at CERFACS.

I also received a lot of support outside of CERFACS: I would like to thank the friends who stood by me throughout these years. I feel extremely lucky to be surrounded by so many great people. I would like to mention in particular Camille, Gaspard, Rémi and Louise whose daily cute pictures of their cats gave me the courage to keep writing this manuscript. Although they did not send me cat pictures, I am also grateful to Amélie, Daria, Lisa, Louis, Mélodie, Paul, Thibault, and many others for their invaluable support (and more generally being genuinely nice people). The path to this PhD has also been made significantly easier by the continuous help and encouragement of my parents in pursuing higher studies. They have given me a chance many do not have, and I wish that one day everybody would have the opportunity I had to follow their passions.

Abstract

Data assimilation involves estimating the state of a system by combining observations from various sources with a background estimate of the state. For geophysical systems such as the ocean, the state would be defined as the values of physical variables (such as temperature or salinity) at every point on a three-dimensional ocean model grid. If we were trying to estimate the state of the ocean on a given day, the background state would typically be a forecast initiated the day before, which would then be corrected with the latest observations. Observations of the system can come from heterogeneous sources, such as satellites and in situ networks. The resulting corrected state of the system is called the analysis, and is typically used as the model initial condition for producing the forecast for the next day.

In order for the analysis to be an accurate representation of the state of the system, we need to have knowledge of the accuracy of the background state and observations. Both are affected by errors, which we assume can be described statistically by their covariance matrices: $\mathbf{B}$ for the background error, and $\mathbf{R}$ for the observation error. Variational data assimilation is a class of data assimilation methods where the analysis is obtained by minimising iteratively a cost function measuring the fit of a candidate state of the system to both the background and observations, with weights defined by $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$. However, most of the complexity of data assimilation pertains to adapting these basic methods to the constraints of the real system, in order to obtain methods that are practical for real applications. In Numerical Weather Prediction, the background state of the system can often represent billions of degrees of freedom, which need to be combined with millions of observations every day. In such a system, the explicit construction of the matrices $\mathbf{B}$ and $\mathbf{R}$ is too expensive to be done in practice. The construction of these matrices can be avoided using the iterative approaches of variational data assimilation, which only require operators that represent matrix-vectors products.

The assumption of uncorrelated observation errors (diagonal $\mathbf{R}$) is commonly made to simplify the inverse covariance operator $\mathbf{R}^{-1}$, which is required by many minimisation algorithms. However, this assumption is problematic when dealing with certain observation types, notably high-resolution satellite data. Neglecting observation error correlations during assimilation often results in suboptimal analyses, where observations tend to be overfit at large spatial scales and underfit at small spatial scales. To address this issue, we have developed an observation error correlation model based on a diffusion operator for the ocean data assimilation system NEMOVAR. Diffusion operators — initially designed for modelling correlations in background error — offer a cost-effective and flexible framework for modelling $\mathbf{R}^{-1}$ (and $\mathbf{R}$ itself) with unstructured data. While accounting for observation error correlations should improve the accuracy of the solution, it also affects the convergence rate of the minimisation algorithms used to approximate the solution. In operational applications, where the minimisation process is usually truncated before full convergence, even correctly accounted for observation error correlations might compromise the accuracy of the analysis. Through analytical and numerical results, we explore the influence of the observation error correlations on both the sensitivity and convergence rate of variational data assimilation algorithms. In particular, we provide insight into how the choice of an observation error correlation model must reflect a balance between computational efficiency and accuracy.

French Summary

L'assimilation de données consiste en un ensemble de méthodes qui peuvent être utilisées pour corriger une estimation initiale de l'état d'un système, appelée l'ébauche, à l'aide d'observations de ce système. En géophysique, le système en question peut être par exemple l'océan, dont l'état est défini comme les valeurs prises par plusieurs variables physiques (telles que la température ou la salinité) en chaque point d'une grille tridimensionnelle. Pour estimer l'état de l'océan un jour donné, l'ébauche pourrait être par exemple une prévision réalisée la veille, qui aurait besoin d'être corrigée avec des observations plus récentes. Les observations peuvent provenir de nombreuses sources hétérogènes, telles que des satellites ou des instruments *in situ*. L'état corrigé du système qui en résulte est appelé analyse, et peut être utilisé par exemple comme condition initiale par le modèle utilisé pour établir les prévisions du lendemain.

L'ébauche et les observations sont affectées par des erreurs, que nous supposons décrites statistiquement par leurs matrices de covariance : $\mathbf{B}$ pour l'erreur d'ébauche et $\mathbf{R}$ pour l'erreur d'observation. L'assimilation de données variationnelle est une catégorie de méthodes d'assimilation des données dans laquelle l'analyse est approchée progressivement en minimisant itérativement une fonction de coût mesurant simultanément l'adéquation d'un état du système à l'ébauche et aux observations, avec des poids définis par $\mathbf{B}^{-1}$ et $\mathbf{R}^{-1}$. Dans les prévisions météorologiques opérationnelles, l'ébauche contient souvent des milliards de degrés de libertés, et sera combinée avec des millions d'observations tous les jours. Cela rend notamment la construction explicite des matrices $\mathbf{B}$ et $\mathbf{R}$ trop coûteuse pour être réalisée. Les approches itératives utilisées en assimilation de données sont souvent utilisées pour contourner le problème car elles ne nécessitent des opérateurs modélisant des produits matrice-vecteur sans accès explicite aux matrices elles-mêmes.

Les erreurs d'observation sont souvent supposées non corrélées (et donc $\mathbf{R}$ diagonal) pour simplifier l'accès à l'opérateur de corrélation inverse, $\mathbf{R}^{-1}$, qui apparaît dans de nombreuses formulations. Cependant, cette hypothèse n'est pas réaliste pour certains types d'observations, en particulier les données satellitaires à haute résolution. Négliger les corrélations des erreurs d'observation pendant l'assimilation mène généralement à des analyses sous-optimales, trop proches des observations aux grandes échelles spatiales, et trop distantes des observations aux petites échelles spatiales. Pour résoudre ce problème, nous avons conçu un opérateur de corrélation associé aux erreurs d'observation basé sur un opérateur de diffusion pour le système d'assimilation de données océaniques NEMOVAR. Les opérateurs de diffusion permettent de modéliser de manière efficace et flexible l'inverse de l'opérateur de corrélation de l'erreur d'observation (et l'opérateur lui-même), et ce même avec des données non structurées.

La prise en compte des corrélations affectant les erreurs d'observation a pour but d'améliorer la qualité de l'analyse, mais elle risque également d'affecter le taux de convergence des algorithmes de minimisation utilisés pour approximer cette analyse. Dans un cadre opérationnel, le processus de minimisation est généralement tronqué avant d'atteindre la convergence totale. En conséquence, même si les corrélations d'erreurs d'observation sont correctement prises en compte, elles pourraient potentiellement compromettre la qualité de la solution obtenue. Sur la base de résultats analytiques et numériques, nous explorons l'influence des corrélations d'erreur

d'observation sur la sensibilité et le taux de convergence des algorithmes d'assimilation variationnelle de données. En particulier, nous cherchons à comprendre comment choisir un modèle de corrélation des erreurs d'observation pour refléter un équilibre entre l'efficacité du calcul et la précision de la solution.

Contents

1 Introduction **8**

2 Context **13**

2.1 Problem formulation: data assimilation . . . 13

2.1.1 Scalar data assimilation . . . 13

2.1.2 The Best Linear Unbiased Estimator . . . 14

2.1.3 Variational data assimilation . . . 15

2.1.4 Non-linear data assimilation . . . 16

2.2 Summary . . . 18

3 Motivation behind this study: the case of wide-swath altimetry **19**

3.1 The SWOT mission . . . 19

3.2 Theoretical error budget . . . 22

3.3 Estimation of error correlations . . . 28

3.4 Consequences for data assimilation . . . 33

3.5 Accounting for SWOT observation error correlations . . . 34

3.6 Summary . . . 35

4 Article 1: Impact of correlated observation errors on the conditioning of variational data assimilation problems **36**

5 Article 2: On the impact of observation-error correlations in data assimilation, with application to along-track altimeter data **72**

6 Ongoing developments and outlook **113**

6.1 Regularised diffusion operators . . . 113

6.1.1 Sensitivity at small spatial scales . . . 113

6.1.2 From control theory to regularised diffusion operators . . . 115

6.1.3 Summary . . . 118

6.2 Extension to other observation types . . . 120

6.2.1 Implementation for two-dimensional datasets . . . 120

6.2.2 Meshing in 2D . . . 127

6.2.3 Summary . . . 128

6.3 Convergence of the Conjugate Gradient beyond the condition number . . . 130

6.3.1 Limitations of the condition number . . . 130

6.3.2 Convergence analysis for variational data assimilation . . . 132

6.3.3 Summary 138

7 Conclusions **140**

Chapter 1

Introduction

Operational weather and climate forecasting centres such as the European Centre for Medium-Range Weather Forecasts (ECMWF), Météo-France or the UK Meteorological Office (Met Office) use vast amounts of in situ and satellite observations to estimate past and present states of Earth system models (atmosphere, ocean, land and sea-ice). On their own, the observations offer only a partial view of the state of the system components: in situ observations are usually sparse in space, while some satellites offer better spatial coverage but only re-observe the same location after multiple weeks, and for the ocean only provide information about ocean surface variables. In the end, most instruments do not observe all the physical variables of interest, but only a few of them or a quantity indirectly related to those of interest.

Data Assimilation (DA) aims at combining all of these partial observations of the state of the system into one, complete, estimate of the system, called the analysis. To do so, all of the observations are used to correct what is usually referred to as a background state: the best complete estimate of the state of the system available *prior* to this state being observed. A classical example would be the estimation of the weather on a given day, by correcting, with the latest observations, the weather forecast produced the day before. The corrected forecast would be the analysis in this scenario, and would typically be used as an initial condition for a forecast for the next day.

In order for the analysis to be as accurate as possible, it is important to account for the fact that neither the background state nor the observations are perfect. The errors affecting the background state and the observations are characterised statistically through their error covariance matrices denoted $\mathbf{B}$ and $\mathbf{R}$, respectively.

These covariance matrices are large, making it impractical to store or represent them explicitly as matrices. Instead, their effect is accessed through operators that perform matrix-vector products directly. Data assimilation algorithms are then chosen so that only operators representing matrix-vector products with $\mathbf{B}$ and $\mathbf{R}^{-1}$ are required to compute the analysis [*e.g.*, Derber and Rosati, 1989, Gürol et al., 2014]. A wide range of sophisticated and well-established methods are available to approximate or model products with $\mathbf{B}$ in current operational systems [*e.g.*, Bannister, 2008]. On the other hand, $\mathbf{R}$ is assumed to be almost or completely diagonal in most configurations for the sake of computational efficiency, meaning that observation error correlations are usually neglected. Multiple operational systems now account for inter-channel error correlations of satellite radiances [Weston et al., 2014, Bormann et al., 2016, Campbell et al., 2017, Geer, 2019], but spatial observation error correlations prove more challenging to account for. This difference is mostly associated with the computational challenge of processing much larger subsets of observations simultaneously. Inter-channel correlations can be represented explicitly with a block per atmosphere column in $\mathbf{R}$. When modelling spatial correlations, there is no such natural 'block' structure for most instruments, which makes it difficult to have an explicit representation

of $\mathbf{R}$ and $\mathbf{R}^{-1}$, especially if observations with correlated errors are scattered across different processors.

The assumption that observation errors are spatially uncorrelated is not realistic for many currently assimilated datasets, such as altimeter data [Storto et al., 2019], satellite-retrieved sea surface temperature [Reid et al., 2020], Doppler radial wind data [Waller et al., 2016b], and Spinning Enhanced Visible and InfraRed Imager (SEVIRI) radiance data [Waller et al., 2016a, Michel, 2018]. This issue is bound to become more pressing with the arrival of new high-resolution datasets, such as wide-swath altimeters [Esteban-Fernandez and Peral, 2018, King and Martin, 2021] or total surface current velocity measurements [Waters et al., 2024], for which accounting for spatial observation error correlations will prove critical.

Neglecting existing observation error correlations is bound to have a detrimental impact on the analysis. Multiple studies in simple systems show, in particular, that we can expect it to hinder the extraction of small-scale features from the observations [*e.g.*, Healy and White, 2005, Stewart et al., 2013, Rainwater et al., 2015, Pinnington et al., 2016]. In order to avoid this loss of valuable information, different practical methods aiming to account for spatial observation error correlations are being developed.

The Singular Value Decomposition (SVD) - Fast Multipole Method (FFM) method is a procedure that allows numerically efficient approximations of matrix-vector products in parallel environments [Simonin et al., 2019, Hu and Dance, 2021, 2024]. It can be used to apply efficiently a matrix $\mathbf{R}^{-1}$, pre-computed offline, to a vector at each iteration of the inner loop. The idea of the method is to partition the observations into 'boxes' of adjacent observations. The blocks of $\mathbf{R}^{-1}$ associated with observations located within the same box are defined and treated explicitly. The blocks of $\mathbf{R}^{-1}$ associated with observations within two separate boxes with a given distance from each other are represented by their truncated SVD, which reduces the cost of the matrix-vector product and compresses the information transferred between processors. Finally, the blocks of $\mathbf{R}^{-1}$ associated with distant boxes are neglected (treated as 0) as a form of localisation. This method offers flexibility as it imposes fewer constraints on the nature of the observation error covariances than other methods. However, it requires an explicit representation of $\mathbf{R}^{-1}$, which could be computationally challenging to build.

Brankart et al. [2009] (see also Ruggiero et al. [2016], Yaremchuk et al. [2018], Bédard and Buehner [2019]) developed a method based on the the assimilation of both observations and observation derivatives, which mimics the effect of using a non-diagonal $\mathbf{R}$. The effect of different $\mathbf{R}$ matrices can be mimicked depending on the weights attributed to observations and their derivatives. Finding which weights best approximate a given matrix $\mathbf{R}$ can be formulated as a minimisation problem to be solved prior to the assimilation [Ruggiero et al., 2016, Yaremchuk et al., 2018]. In the context of the assimilation of satellite images, Chabot et al. [2015] considered more generically how altering the observation space could indirectly allow observation error correlations to be accounted for. This alteration takes the form of an operator $\mathcal{A}$ applied as a pre-processing to the observations but also to their counterpart extracted from the model in the observation operator. For example, if the operator $\mathcal{A}$ returns the gradient of an image, then this procedure is equivalent to assimilating the gradient of the observations. However, the authors consider a broader range of operators $\mathcal{A}$ such as Fourier, wavelet, or curvelet transforms. This method thus provides a natural way of specifying different levels of observation errors associated with different spatial scales. The weights need to be determined offline by minimizing a cost function measuring how well the use of $\mathcal{A}$ can approximate the effect of a target observation error covariance matrix.

An advantage of these methods, which are based on an alternative observation space, is their simplicity as no actual observation error correlation operator needs to be implemented. However, in the case of the assimilation of observation derivatives along with the observations themselves, the size of the observation space is increased by approximately a factor five (or more if higher-order derivatives were to be used for more flexibility). The cost of the transform $\mathcal{A}$ should also be considered with care as it needs to be systematically applied alongside the observation operator. Another potential issue is the lack of adaptability to unstructured data: many observations

do not have a grid-like structure or could lose it through quality control if they did. It is interesting to note that these approaches can go beyond what can be achieved with a non-diagonal $\mathbf{R}$ if $\mathcal{A}$ is not invertible. For example, Metref et al. [2020] use a similar approach, not with the goal of accounting for observation error correlations, but to reject error-prone features in SWOT data. By using an operator $\mathcal{A}$ to remove, among other components, the linear trends perpendicular to the satellite swath known to be typical of large correlated errors, they limit the impact on the analysis of these errors. However, this error removal comes at the cost of rejecting geophysical signals in the observations whose features resemble those typical of the error.

Covariance operators based on the discretisation of a diffusion equation were originally developed to model background error correlations [Weaver and Courtier, 2001] in ocean data assimilation. Guillet et al. [2019] showed that 'diffusion operators' can be discretised on fully unstructured meshes using the finite element method, making them applicable for modelling observation error correlations with unstructured data. When the data lie on a Cartesian grid, using diffusion operators is mathematically equivalent to assimilating gradient or higher-order derivatives of the observations [Guillet et al., 2019, Section 4.1]. The diffusion method can thus be considered a generalisation of observation gradient assimilation methods. Diffusion operators avoid the explicit construction and inversion of $\mathbf{R}$ by building directly an operator associated with $\mathbf{R}^{-1}$. They use the input vector of the matrix-vector product as the initial condition for an implicit diffusion scheme, and output the solution after m iterations. Weaver and Mirouze [2013] show that this operator is, under the right conditions (far from the boundaries of the domain; spatially constant parameters), equivalent to a covariance operator whose correlation function belongs to the Matérn class [*e.g.*, Stein, 1978, Guttorp and Gneiting, 2006]. The number of diffusion iterations m affects the shape and regularity of the function while the diffusivity used in the numerical scheme can be related to function length scale. One of the advantages of this method is its low computational cost: the operator associated with $\mathbf{R}^{-1}$ is very sparse and local by construction, which allows for an efficient and scalable parallel implementation for large data sets. The main cost of diffusion operators for observation error correlation modelling is in the offline estimation of the normalisation factors [Weaver et al., 2020] required to replace their natural variance by those of the observation error.

Diffusion operators are more restrictive than the SVD-FFM in terms of the shape of the correlation function as they are constrained to Matérn-like correlation functions. However, they offer an inexpensive option for observation types for which this degree of flexibility is sufficient. For example, diffusion operators are flexible enough to approximate the error correlations of SWOT data as shown in Section 3.3. In this case, diffusion operators would prove of particular interest to avoid the construction and inversion of $\mathbf{R}$ for such a large volume of data.

In operational applications, the aim is to obtain an analysis as accurate as possible with a limited computational budget. In variational data assimilation, in which the analysis is obtained iteratively by minimizing a cost function, the limited computational budget means that the minimisation needs to be truncated before full convergence. The analysis whose quality should be evaluated is thus the solution of the *truncated* minimisation problem, denoted $\mathbf{x}_k$ where k denotes the iteration counter. In this PhD, the goal is to gain more insight on how accounting for observation error correlations through a non-diagonal $\mathbf{R}$ could allow $\mathbf{x}_k$ to be statistically closer to a true state $\mathbf{x}^{(\mathrm{t})}$ without significantly increasing the computational budget. The solution of the minimisation problem at full convergence, $\mathbf{x}^*$, differs from the true state $\mathbf{x}^{(\mathrm{t})}$. One of the sources of this difference can be related to an inconsistency between the specification of $\mathbf{R}$ and the actual observation error correlations. There is also a difference between the solution of the truncated minimisation problem $\mathbf{x}_k$ and the solution at full convergence $\mathbf{x}^*$. This difference also depends on $\mathbf{R}$ as its specification can influence the convergence rate of the minimisation, and thus how close $\mathbf{x}_k$ is from $\mathbf{x}^*$ at the truncation. Consequently, to understand the effect of a non-diagonal $\mathbf{R}$, we focus on two main directions:

- the *convergence rate* of the minimisation algorithm, which affects the difference between $\mathbf{x}^*$ and $\mathbf{x}_k$;
- the *accuracy of the solution* of the minimisation at full convergence with respect to the truth, which is the difference between $\mathbf{x}^*$ and $\mathbf{x}^{(\mathrm{t})}$.

The thesis is structured as follows to present the stated contributions. The formalism of data assimilation is briefly introduced in Chapter 2. To motivate the necessity of this study and contextualise it, we present in Chapter 3 wide-swath altimetry data, whose potential in DA strongly depends on addressing its correlated observation error.

In Chapter 4 we analyse the effect of a non-diagonal $\mathbf{R}$ on the convergence rate of the minimisation problem at the core of variational data assimilation. Previous studies suggest that the use of a non-diagonal $\mathbf{R}$ is likely to affect the iterative methods used in variational DA, and in particular of the often used Conjugate Gradient (CG) algorithm. Most of these studies focus on the condition number of the Hessian matrix of the cost function minimised by the CG algorithm, with or without preconditioning it by $\mathbf{B}$. This choice is justified by the fact that the condition number can be used as a first-order indicator of the convergence rate of the CG algorithm [Saad, 2003, Section 6.11.3]. While keeping $\mathbf{R}$ diagonal, Haben et al. [2011] already showed that increasing the density of observations is likely to increase the condition number of the Hessian matrix of the cost function minimised by the $\mathbf{B}$-preconditioned CG. This should be accounted for here, as one potential improvement we hope a non-diagonal $\mathbf{R}$ would bring is the ability to assimilate observations of higher density. With a non-diagonal $\mathbf{R}$ and in an idealised framework, Tabeart et al. [2018] showed that, without preconditioning, the condition number is highly sensitive to the specification of observation error correlations: for a fixed background error correlation length scale, the condition number is minimised by reducing the observation error correlation length scale (or by reducing the background error correlation length scale if the observation error correlation length scale is fixed). This study is extended to the $\mathbf{B}$-preconditioned problem in Tabeart et al. [2020] in an idealised framework using Second Order Auto-Regressive (SOAR) functions for both the background and observation error correlations. They show experimentally that, for a fixed background error correlation length scale, the condition number is not minimised by the smallest possible observation error correlation length as in the unpreconditioned case but by intermediate length scales (depending on the background error correlation length scale). This result is of interest as it shows that moving from a diagonal to a non-diagonal $\mathbf{R}$ is not bound to degrade the convergence of the CG algorithm.

Building on these works, we use the same approach based on the condition number as a proxy for the convergence rate of the $\mathbf{B}$-preconditioned CG, but attempt to generalise these results by considering Auto-Regressive (AR) correlation functions of different orders. AR functions include, but are not limited to SOAR functions, and can be modelled using diffusion operators (as explained in Chapters 4 and 5). We also define the observation operator as a 'selection operator', where observations are distributed regularly in space but at a possibly lower resolution than that of the background. This allows us to study analytically the sensitivity of the condition number to the specification of observation error correlations with different length scales and functional shapes (*e.g.,* Gaussian *vs* SOAR) for the background and observations. We confirm the findings of Tabeart et al. [2020] and provide additional guidelines on how to avoid high condition numbers when specifying observation error correlation parameters. The relative shapes of the background and observation error correlation functions prove to be particularly critical in that effect. Numerical experiments are then used to illustrate variations in the convergence of CG due to the use of a non-diagonal $\mathbf{R}$. Additionally, these experiments show how changes in the convergence rate compare to changes in $\mathbf{x}^*$, as modifying $\mathbf{R}$ affects both simultaneously. These results have been published in an article [Goux et al., 2024], which can be found in Chapter 4.

One of the conclusions of Goux et al. [2024] is that, with a limited number of iterations, the observation error correlation model that best approximates the true observation error correlations is not necessarily the one

that leads to a solution closest to the true state $\mathbf{x}^{(\mathrm{t})}$. This highlights our need to gain a better understanding of what would constitute a 'good' observation error correlation model. In Chapter 5, we shift our focus from the impact of the choice of an observation error correlation model on the convergence rate of the minimisation (and thus on the difference between $\mathbf{x}_k$ and $\mathbf{x}^*$) to its impact on the solution of the minimisation at full convergence (and thus on the difference between $\mathbf{x}^*$ and $\mathbf{x}^{(\mathrm{t})}$). We study in an idealised framework how differences between the observation-error correlation model and true observation-error correlations affect differences between $\mathbf{x}^*$ and $\mathbf{x}^{(\mathrm{t})}$. To do so, we formulate the data assimilation problem in spectral space where the effect of error correlations can be visualised clearly. The relevancy of the insights gained in this simple system is evaluated experimentally using a global configuration of NEMOVAR, the variational ocean data assimilation system used operationally at ECMWF and the Met Office [Mogensen et al., 2009]. To do so, we have implemented in NEMOVAR an observation error correlation operator for nadir altimeter data based on diffusion operators. The implementation of this operator has been designed with the idea not only to be used for the experiments of Chapter 5 but also to serve as the basis for future operational applications. As such, its implementation has been made efficient in a parallel environment, scalable for large datasets, and created to facilitate the extension to other observation types. These results are presented in an article to be submitted to the *Quarterly Journal of the Royal Meteorological Society*.

Chapter 6 discusses ongoing and future work. In Section 6.1, we present a method which could be used to 'regularise' diffusion operators with an effect comparable to that of 'reconditioning' methods [*e.g.*, Tabeart et al., 2020] without the requirements to have access to an explicit representation of $\mathbf{R}$. We are developing this method to address the sensitivity issues appearing in some experimental configurations of Chapter 5. In Section 6.2, we discuss the extension of the observation error correlation operator implemented in NEMOVAR to other observation types. We present in more depth how our current implementation is built to facilitate a generalisation to other observation types, and the challenges that remain for this extension to take place. Finally, we consider in Section 6.3, the convergence properties of CG beyond what can be characterised by the condition number. We give an example highlighting the limitations of the condition number as a sole metric for the convergence rate, and demonstrate the importance of using complementary diagnostics. Conclusions are given in Chapter 7.

Chapter 2

Context

2.1 Problem formulation: data assimilation

2.1.1 Scalar data assimilation

Let us start with the example of scalar data assimilation where there is one observation and one grid point. While this configuration is very simple, insights developed with it prove useful to understand more complex cases. Let us denote y_o the observation and x_b the (scalar) background state. We assume that y_o and x_b differ from a true state x_t due to additive Gaussian errors ϵ_o and ϵ_b of mean zero and standard deviation σ_o and σ_b, respectively:

$$y_\mathrm{o} - x_\mathrm{t} = \epsilon_\mathrm{o} \sim \mathcal{N}(0, \sigma_\mathrm{o}^2),$$
$$x_\mathrm{b} - x_\mathrm{t} = \epsilon_\mathrm{b} \sim \mathcal{N}(0, \sigma_\mathrm{b}^2).$$

In order for the analysis error to also be unbiased, we constrain the analysis x_a to be a convex combination of y_o and x_b with a scalar k determining the balance between x_b and y_o:

$$x_\mathrm{a} = (1-k)x_\mathrm{b} + ky_\mathrm{o}. \tag{2.1}$$

Defining the analysis increment $\delta x_\mathrm{a} = x_\mathrm{a} - x_\mathrm{b}$ and the innovation $d = y_\mathrm{o} - x_\mathrm{b}$, we can rewrite Equation (2.1) as a 'filtering' problem:

$$\delta x_\mathrm{a} = kd,$$

in the sense that d constitutes a correction to the background which needs to be damped by k in order to obtain the optimal correction δx_a. Optimality is defined by the value of k that minimises the analysis error variance, which yields

$$k = \frac{\sigma_\mathrm{b}^2}{\sigma_\mathrm{b}^2 + \sigma_\mathrm{o}^2},$$

and the associated analysis error variance

$$\sigma_\mathrm{a}^2 = \left(\frac{1}{\sigma_\mathrm{o}^2} + \frac{1}{\sigma_\mathrm{b}^2}\right)^{-1}.$$

The response of data assimilation to the specification of the background and observation error standard deviations is easy to observe in this configuration:

- if we expect the observations to be much more accurate than the background state (*i.e.*, $\sigma_\mathrm{o} \ll \sigma_\mathrm{b}$), the innovation is not filtered (*i.e.*, $k \simeq 1$), the analysis is almost equal to the observation (*i.e.*, $x_\mathrm{a} = y_\mathrm{o}$) and its error variance is close to that of the observation error (*i.e.*, $\sigma_\mathrm{a}^2 = \sigma_\mathrm{o}^2$);
- if we expect the observations to be much less accurate than the background state (*i.e.*, $\sigma_\mathrm{o} \gg \sigma_\mathrm{b}$), the innovation is fully filtered (*i.e.*, $k \simeq 0$), the analysis is almost equal to the background (*i.e.*, $x_\mathrm{a} = x_\mathrm{b}$) and its error variance is close to that of the background error (*i.e.*, $\sigma_\mathrm{a}^2 = \sigma_\mathrm{b}^2$);
- if we expect the observations to be as accurate as the background state (*i.e.*, $\sigma_\mathrm{o} = \sigma_\mathrm{b}$), the innovation is halved (*i.e.*, $k = 1/2$), the analysis is equal to the mean of the background state and observation (*i.e.*, $x_\mathrm{a} = (x_\mathrm{b} + y_\mathrm{o})/2$) and its error variance is close to half of that of the background or observation error (*i.e.*, $\sigma_\mathrm{a}^2 = \sigma_\mathrm{b}^2/2 = \sigma_\mathrm{o}^2/2$).

2.1.2 The Best Linear Unbiased Estimator

We now consider the more general case where the state of the system is represented by n variables, and where p observations of the system are available. The background state and the analysis are now defined as vectors $\mathbf{x}_\mathrm{b}$ and $\mathbf{x}_\mathrm{a}$ of size n. The observations are stored in a vector $\mathbf{y}_\mathrm{o}$ of size p.

Part of the added complexity compared to the scalar case observed in Section 2.1.1 is that $\mathbf{x}_\mathrm{b}$ and $\mathbf{y}_\mathrm{o}$ are not directly comparable anymore, for example if observations are not available at every point of the system (*i.e.*, $p < n$). A transform is required to map the state space (associated with what we want to estimate) to the observation space (associated with what we can measure). We assume for now that this transform can be perfectly represented by a linear operator $\mathbf{H}$, referred to as the observation operator. If the state of the system was represented by a vector $\mathbf{x}$, then $\mathbf{Hx}$ would compute how the system appears through our observation network. For example, it often contains linear interpolations of the state of the system to observation locations.

We still assume that the background state and observations differ from a reference true state $\mathbf{x}^{(\mathrm{t})}$ due to additive Gaussian errors. Observation error and background error are now characterised by their covariance matrices $\mathbf{R}$ and $\mathbf{B}$, respectively:

$$\begin{aligned} \mathbf{y}_\mathrm{o} - \mathbf{H}x_\mathrm{t} &= \boldsymbol{\epsilon}_\mathrm{o} \sim \mathcal{N}(0, \mathbf{R}), \\ \mathbf{x}_\mathrm{b} - x_\mathrm{t} &= \boldsymbol{\epsilon}_\mathrm{b} \sim \mathcal{N}(0, \mathbf{B}). \end{aligned}$$

Following the same steps as in the scalar case, we can determine the analysis $\mathbf{x}_\mathrm{a}$ as a linear combination of $\mathbf{x}_\mathrm{b}$, $\mathbf{Hx}_\mathrm{b}$ and $\mathbf{y}_\mathrm{o}$. Under the constraint that $\mathbf{x}_\mathrm{a} = \mathbf{x}^{(\mathrm{t})}$ if $\mathbf{y}_\mathrm{o} = \mathbf{Hx}^{(\mathrm{t})}$ and $\mathbf{x}_\mathrm{b} = \mathbf{x}^{(\mathrm{t})}$ (or equivalently that the analysis error is unbiased), we can define this combination through a single $n \times p$ matrix $\mathbf{K}$:

$$\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{b} + \mathbf{K}(\mathbf{y}_\mathrm{o} - \mathbf{Hx}_\mathrm{b}), \tag{2.2}$$

or equivalently

$$\delta\mathbf{x}_\mathrm{a} = \mathbf{Kd},$$

where $\delta\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{a} - \mathbf{x}_\mathrm{b}$ is the analysis increment and $\mathbf{d} = \mathbf{y}_\mathrm{o} - \mathbf{Hx}_\mathrm{b}$ is the innovation vector.

We select the matrix $\mathbf{K}$ by minimising the trace of the analysis error covariance matrix, which leads to

$$\mathbf{K} = \mathbf{BH}^\mathrm{T}(\mathbf{HBH}^\mathrm{T} + \mathbf{R})^{-1} = (\mathbf{B}^{-1} + \mathbf{H}^\mathrm{T}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^\mathrm{T}\mathbf{R}^{-1}. \tag{2.3}$$

The matrix $\mathbf{K}$ defined in (2.3) is usually referred to as the Kalman gain, and the analysis $\mathbf{x}_\mathrm{a}$ derived from it using Equation (2.2) is called the Best Linear Unbiased Estimator (BLUE). As in the scalar case, the covariance matrix of the analysis error obtained with this definition of $\mathbf{K}$ can be expressed in terms of the background and observation error covariance matrices:

$$\mathbf{P}_\mathrm{a} = \left(\mathbf{B}^{-1} + \mathbf{H}^\mathrm{T}\mathbf{R}^{-1}\mathbf{H}\right)^{-1}.$$

The variances (diagonal elements) the covariance matrices play the same role as in the scalar case where they define the overall expected amplitude of the background and observation error. Variances can vary in space, to represent the fact that observations or the background might not be equally accurate everywhere. In addition to the variances, the covariance matrices contain information about background and observation error correlations. The correlations describe the statistical relations between the different elements of $\boldsymbol{\epsilon}_\mathrm{b}$ and $\boldsymbol{\epsilon}_\mathrm{o}$. As detailed in Chapter 5, under certain conditions, introducing correlations can also be interpreted as introducing *scale-dependent* variances.

For large applications such as those encountered in Earth system modelling, the dimension of the state of the system and the number of observations are too large (*e.g.*, $n \simeq 10^9$; $p \simeq 10^6$) for Equations (2.2) and (2.3) to be used directly to compute the analysis. Building and inverting explicitly the matrices involved in Equation (2.3) would lead to prohibitively large memory requirements. Moreover, our approximate knowledge of the actual background and observation error covariances does not warrant such a degree of fine-tuning where each element of $\mathbf{B}$ and $\mathbf{R}$ is specified individually. Multiple practical approaches have been developed to approximate $\mathbf{x}_\mathrm{a}$. The main approach considered in this manuscript is variational data assimilation, presented in Section 2.1.3. Other notable approaches include 'ensemble' methods, in which $\mathbf{B}$ is defined as a Monte-Carlo estimator based on a set of background error samples, which allows (2.2) to be re-worked in a subspace of lower dimension [e.g., Section 6 of Asch et al., 2016].

2.1.3 Variational data assimilation

The analysis $\mathbf{x}_\mathrm{a}$ from Equation (2.2) is also the solution of the minimisation problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} J(x) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}_\mathrm{b}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathbf{H}\mathbf{x} - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}}, \tag{2.4}$$

where, for any vector $\mathbf{x} \in \mathbb{R}^n$ and matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, the notation $\|\mathbf{x}\|^2_\mathbf{A}$ denotes the squared norm $\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}$. The cost function in Equation (2.4) is quadratic and can be minimised with an iterative method that can be implemented without ever requiring access to the elements of the matrices $\mathbf{B}$, $\mathbf{R}$, or their inverse. Instead, only operators modelling the product of a vector with these matrices is required. The iterative minimization of (2.4) is the basis of variational data assimilation.

For the quadratic cost function (2.4), the Conjugate Gradient (CG) algorithm [Hestenes and Stiefel, 1952] is commonly used, with $\mathbf{B}$ as a preconditioning matrix [*e.g.*, Lorenc, 1997]. With $\mathbf{B}$ preconditioning, CG can be formulated such that it does not require access to $\mathbf{B}^{-1}$ [*e.g.*, Gürol et al., 2014]. However, this formulation still requires the product of a vector with $\mathbf{R}^{-1}$ at each iteration, which is challenging when $\mathbf{R}$ is not diagonal and a reason why many operational centres do not account for observation error correlations.

As presented in Chapter 1, new methods are emerging that can model a matrix-vector product with $\mathbf{R}^{-1}$ at a reasonable computational cost. However, even under the assumption that a non-diagonal $\mathbf{R}$ does not affect the cost of an iteration of the minimisation, it could affect the overall efficiency of the minimisation by altering the convergence rate of the CG alogrithm. The performance of a CG algorithm is highly sensitive to the eigenvalues

of the Hessian matrix of the cost function. For the cost function in Equation (2.4), and with $\mathbf{B}$-preconditioning, the Hessian matrix takes the form

$$\mathbf{S} = \mathbf{I} + \mathbf{U}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{U} \tag{2.5}$$

where $\mathbf{U}$ is defined such that $\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathrm{T}}$. Altering $\mathbf{R}$ will alter $\mathbf{S}$ and in turn the convergence rate of the minimisation.

The condition number of the Hessian matrix $\mathbf{S}$ is a first order indicator of the convergence properties of CG, through which we can understand the dependency of the convergence rate to the specification of the observation error correlations. This is the subject of the article Goux et al. [2024], included in Chapter 4. In addition to the condition number, a deeper understanding of the convergence of CG can be gained by understanding the relationship between the eigendecomposition of $\mathbf{S}$ and the gradient of the cost function (2.4), as shown in Section 6.3 of Chapter 6.

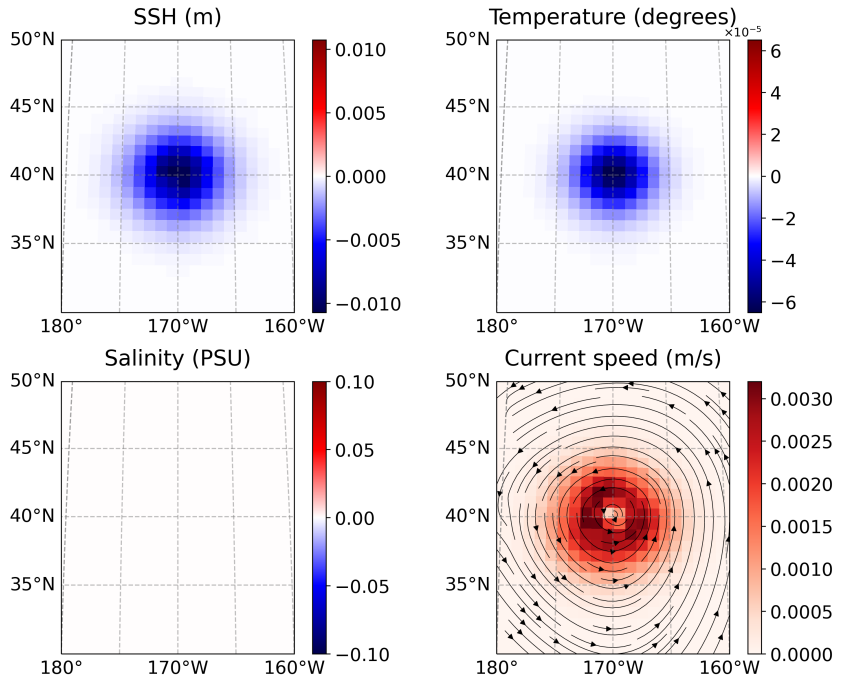
2.1.4 Non-linear data assimilation

We have assumed up to now that there exists a linear observation operator $\mathbf{H}$ that maps the state space to the observation space. This assumption is restrictive: the observation operator can contain simple operations such as linear interpolations to the observation locations, but also more complex, non-linear transformations if the variable observed is not directly the one we want to estimate. For example, if the variable to be estimated is humidity in the atmosphere, and the variable measured by a satellite is radiance at the top of the atmosphere, the observation operator would include a non-linear radiative transfer model.

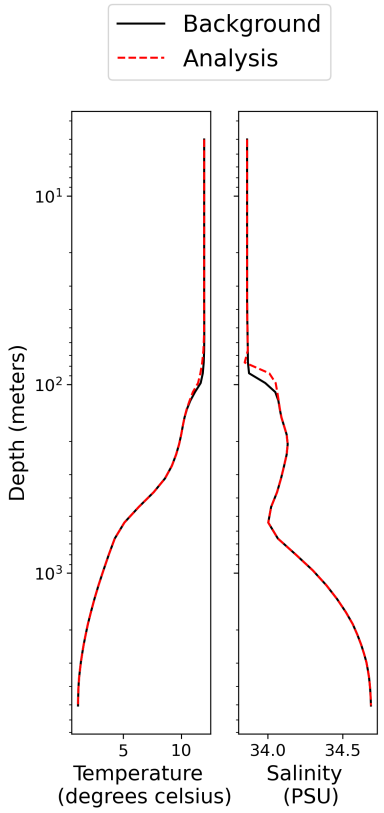
In practice, such non-linear transformations between variables are common even with 'simpler' observation types in multivariate DA systems, where different physical variables are estimated (which is the case in most operational applications). In order to facilitate the modelling of the operator associated with $\mathbf{B}$, the variables estimated in $\mathbf{x}$ – referred to as control variables – are often transformed variables chosen to minimise their inter-correlations. This allows $\mathbf{B}$ to be defined as a block-diagonal with each block corresponding to the auto-correlations of a variable. Balance relations [*e.g.*, Weaver et al., 2005], which are possibly non-linear, are used to transform the control variables to the physical variables. They can be interpreted as transformations within a generalised observation operator.

The balance operator allows observational information from one variable to be transferred to other variables. For example, in NEMOVAR, the main physical variables are temperature, salinity, the horizontal velocity components and Sea Surface Height (SSH). Figure 2.1 shows the impact of assimilating a single SSH observation in NEMOVAR. The observation generates a negative SSH analysis increment not only at the analysis location but also at surrounding grid points due to the horizontal spatial background error correlations for SSH. Two of the balance relations applied in NEMOVAR are the steric height relation and the approximate mass conserving temperature-salinity relation. In response to a decrease in the SSH, the water column is cooled. As seen in Figure 2.1b, this cooling is mostly distributed at depths where the vertical gradient of the background temperature is largest, which effectively targets the thermocline. To preserve the stratification of the original profile, it is assumed that the local changes in temperature are mostly due to a vertical displacement of the water column, which are converted to a similar vertical displacement of the salinity profile [Ricci et al., 2005]. No balance between temperature and salinity is applied in areas of weak stratification such as the mixed layer, which explains the lack of salinity increment at the surface (Figure 2.1a). Finally, the meridional and zonal velocity components are updated by maintaining geostrophic balance with the updated density gradients.

With a non-linear observation operator, $\mathcal{H}$, the counterpart of the cost function in Equation (2.4) is no longer



(a) Analysis increment at the surface in terms of SSH, temperature, salinity and velocity.



(b) Temperature and salinity profiles of the background and analysis under the location of the observation.

Figure 2.1: Impact of a single SSH observation in NEMOVAR. The observation is taken to be in the Pacific at latitude 40°N and longitude 170°W.

quadratic:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}_\mathrm{b}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{H}(\mathbf{x}) - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}}. \tag{2.6}$$

This cost function is usually referred to as the 3D-Var cost function. Another source of non-linearity appears when one attempts to account for the fact that not all observations are measured simultaneously. In this case, the state of the system $\mathbf{x}$ in the cost function is considered to be the state of the system at the beginning of the assimilation window. A non-linear model can then be used to obtain a physically meaningful trajectory starting from $\mathbf{x}$, so that each observation can be compared to the model state closest to the observation time. Altering the cost function (2.6) to include time-dependency leads to the 4D-Var cost function:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}_\mathrm{b}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{G}(\mathbf{x}) - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}}, \tag{2.7}$$

where $\mathcal{G}$ is a 'generalised' observation operator and $\mathbf{y}_\mathrm{o}$ is a concatenation of observations from different times. To compare a system state $\mathbf{x}$ to an observation, the operator $\mathcal{G}$ can involve:

- applying balance relations to convert $\mathbf{x}$ into physical variables which can be propagated by the model;
- propagating $\mathbf{x}$ up to the observation time using the non-linear model;
- applying a transformation to the model variable to convert it to the observed quantity;

- interpolating the result to the observation location.

The minimisation of the non-quadratic cost function (2.7) is usually achieved by minimising successive quadratic approximations of it using a Gauss-Newton algorithm [Gratton et al., 2007]. The quadratic cost functions are obtained by linearizing $\mathcal{G}$ locally around successive approximations of $\mathbf{x}$, denoted $\mathbf{x}_k$:

$$\mathcal{G}(\mathbf{x}) \simeq \mathcal{G}(\mathbf{x}_k) + \mathbf{G}_{\mathbf{x}_k}\delta\mathbf{x},$$

where $\mathbf{x} = \mathbf{x}_k + \delta\mathbf{x}$ and $\mathbf{G}_{\mathbf{x}_k}$ is the Jacobian matrix of $\mathcal{G}$ (often referred to as the tangent linear model of $\mathcal{G}$) around $\mathbf{x}_k$. This approximation can be used to derive the quadratic cost function $J_{\mathbf{x}_k}$ osculating $\mathcal{J}$ around $\mathbf{x}_k$:

$$\begin{aligned}\mathcal{J}(\mathbf{x}_k + \delta\mathbf{x}) &= \frac{1}{2}\|\delta\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{G}(\mathbf{x}_k + \delta\mathbf{x}) - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}},\\ J_{\mathbf{x}_k}(\delta\mathbf{x}) &= \frac{1}{2}\|\delta\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathbf{G}_{\mathbf{x}_k}\delta\mathbf{x} - \mathbf{d}\|^2_{\mathbf{R}^{-1}},\end{aligned}$$

where $\mathbf{d} = \mathbf{y}_\mathrm{o} - \mathcal{G}(\mathbf{x}_k)$.

For the first Gauss-Newton iteration, the iterate $\mathbf{x}_k$ is initialised using the background ($\mathbf{x}_0 = \mathbf{x}_\mathrm{b}$). The quadratic cost function is minimised with a CG algorithm as explained in Section 2.1.3. Then, the solution of that minimisation problem is used to update $\mathbf{x}_k$ which serves as a new linearisation point for $\mathcal{G}$. In the data assimilation community, the minimisation of the quadratic cost function is usually referred to as the 'inner loop', while the successive approximations of the non-quadratic cost function is referred to as the 'outer loop'. This approach is often referred to as the 'incremental formulation' [Courtier et al., 1994].

2.2 Summary

- Data assimilation consists of finding a statistically optimal estimate of the state of a system, called the analysis, by combining an a priori estimate of the system state, called the background state, and observations of the system. In variational data assimilation, the analysis can be defined as the solution of a weighted least-squares problem. The weights attributed to the background state and observations are defined in terms of their error covariance matrices $\mathbf{B}$ and $\mathbf{R}$, respectively.

- In order to compare estimates of the state of the system to observations, we need an observation operator $\mathcal{G}$ which maps the state space to the observation space. The operator $\mathcal{G}$ can contain linear relations such as interpolations to observation locations and non-linear relations such as a dynamical model propagating the state of the system to the observation time.

- The non-linearity introduced by $\mathcal{G}$ means that the cost function to be minimised to find the analysis is not quadratic. The solution of the non-linear problem is obtained iteratively using a Gauss-Newton algorithm, where successive quadratic approximations of the cost function are obtained by linearizing locally the operator $\mathcal{G}$.

- The quadratic cost functions can be minimised using a CG algorithm, which removes the need to build the prohibitively large matrices $\mathbf{B}$ and $\mathbf{R}$. Instead, only operators modelling the product of a vector with $\mathbf{B}$ and $\mathbf{R}^{-1}$ are required.

- We will focus on the solution of the linear problem with CG in this manuscript. The results are however also relevant to the non-linear problem as they apply to each inner loop of the Gauss-Newton algorithm.

Chapter 3

Motivation behind this study: the case of wide-swath altimetry

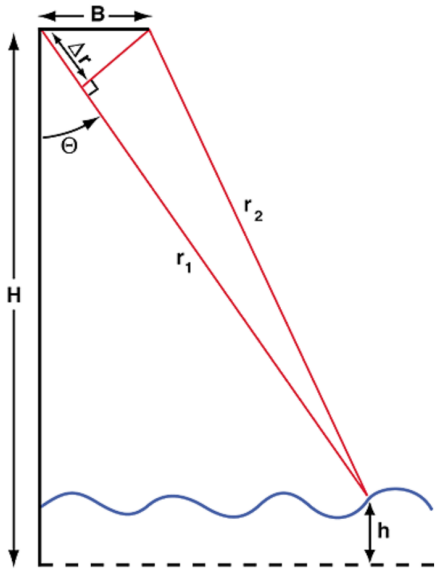


Figure 3.1: Reproduction of Figure 1 from Esteban-Fernandez and Peral [2018]: principle of the interferometric triangulation. Given the distance between the two antennas B, the platform elevation H and the range measurements from each antenna r_1 and r_2, one can deduce both the look angle θ and the elevation h of the sea surface relative to the ellipsoid (knowing the platform altitude H).

Designing a computationally efficient non-diagonal $\mathbf{R}$ is challenging for an operational DA system. Observation errors are often assumed be uncorrelated for this reason, so that $\mathbf{R}$ can be approximated by a diagonal matrix. To highlight why the assumption of uncorrelated observation error needs to be challenged, and to better justify the effort put into developing methods to model observation error correlations, we will provide in this chapter an example of a dataset for which observation error correlations could be critical: wide-swath altimetry.

3.1 The SWOT mission

The Surface Water Ocean Topography (SWOT) mission is a new satellite altimetry mission launched in December 2022. Its main instrument is a Ka band Radar Interferometer (KaRIn), which provides Sea Surface Height (SSH) measurements at an unprecedented resolution.

In altimetry, the measure is based on the emission of an electromagnetic pulse and the measure of its echo reflected by the sea surface. The received echo is not instantaneous but a function of time. The received signal is separated in time bins, which correspond to different ranges, *i.e.*, different distances from the satellite. Each bin corresponds to an iso-range surface, a sphere centred on the satellite. The intensity of the signal in a bin depends on the reflectiveness of the intersection of the iso-range sphere and the sea-surface. In traditional nadir altimetry, the distance to the sea surface is deduced from the power received at each iso-range through a process called waveform retracking.

In interferometric altimetry, we have access to more information: the sea surface is observed from two different angles in order to triangulate the position of the points at the surface, as shown in Figure 3.1. The SWOT satellite uses two antennas positioned on each side of a mast perpendicular to the satellite track.

The antennas are separated by a distance B, referred to as the baseline. The difference between the two range measurements is extremely small compared to the actual range, but can be measured with high accuracy through the phase difference between the two signals. As shown in Figure 3.1, the phase can be used to determine the look angle; *i.e.*, the angle between the nadir and the position vector of the source of the echo. For each pulse sent, it is thus possible to obtain a one-dimensional (across-track) map of the SSH. If a pulse is sent regularly along the satellite track, a two-dimensional map of the SSH can be generated.

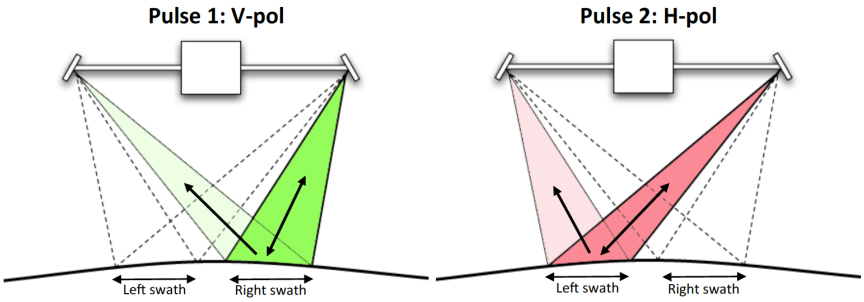


Figure 3.2: Reproduction of Figure 4 from Esteban-Fernandez and Peral [2018]. Each antenna sends one pulse but measures both echoes; *e.g.*, the right antenna produces a pulse with horizontal polarization towards the left half-swath and both antennas measure its reflection to build the left SSH half-swath.

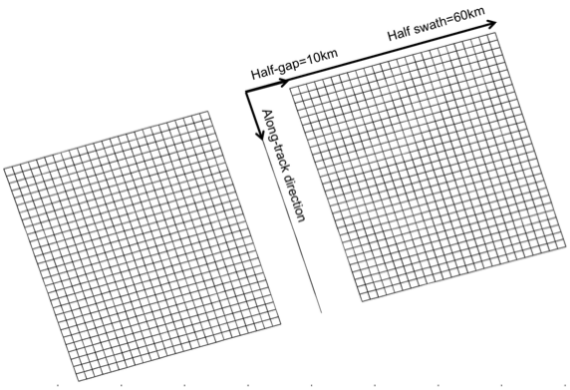


Figure 3.3: Visualisation of a segment of the SWOT grid at a 2 km resolution from the SWOT simulator documentation (`https://swot-simulator.readthedocs.io/en/latest/sampling/swath.html`).

In practice, the resolution and accuracy achieved with this method are sub-optimal (and insufficient for most science requirements). A better resolution and accuracy can be achieved by using SAR (Synthetic Aperture Radar) compression methods. The idea behind SAR imagery is that each pulse can actually be used to obtain a two-dimensional map of the SSH (and not just one-dimensional in the across-track direction). Mapping the SSH of a two-dimensional surface can actually be seen as geolocating in three dimensions sea surface 'pixels'. As seen before, the iso-range surfaces are spheres centred on the satellite. The iso-phase surfaces are cones centred about a vector parallel to the baseline. We need a third measurement to determine the three-dimensional position of the 'pixels' reflecting the pulse. In SAR imagery, this third measurement is the Doppler frequency shift of the echo pulse. This shift is a function of the angle between the satellite velocity and the position vector linking the satellite and the sea surface 'pixel'. The iso-Doppler surfaces are cones centred about the velocity vector of the satellite. Only two points share the same range, phase difference, and Doppler shift; *i.e.*, are at the intersection of an iso-range sphere, an iso-phase 'across-track' cone, and an iso-Doppler 'along-track' cone. One is on the left of the satellite, another is on the right. The ambiguity between the two points can be eliminated by emitting the pulse at an angle away from the nadir, so that the image is only on one side of the satellite track.

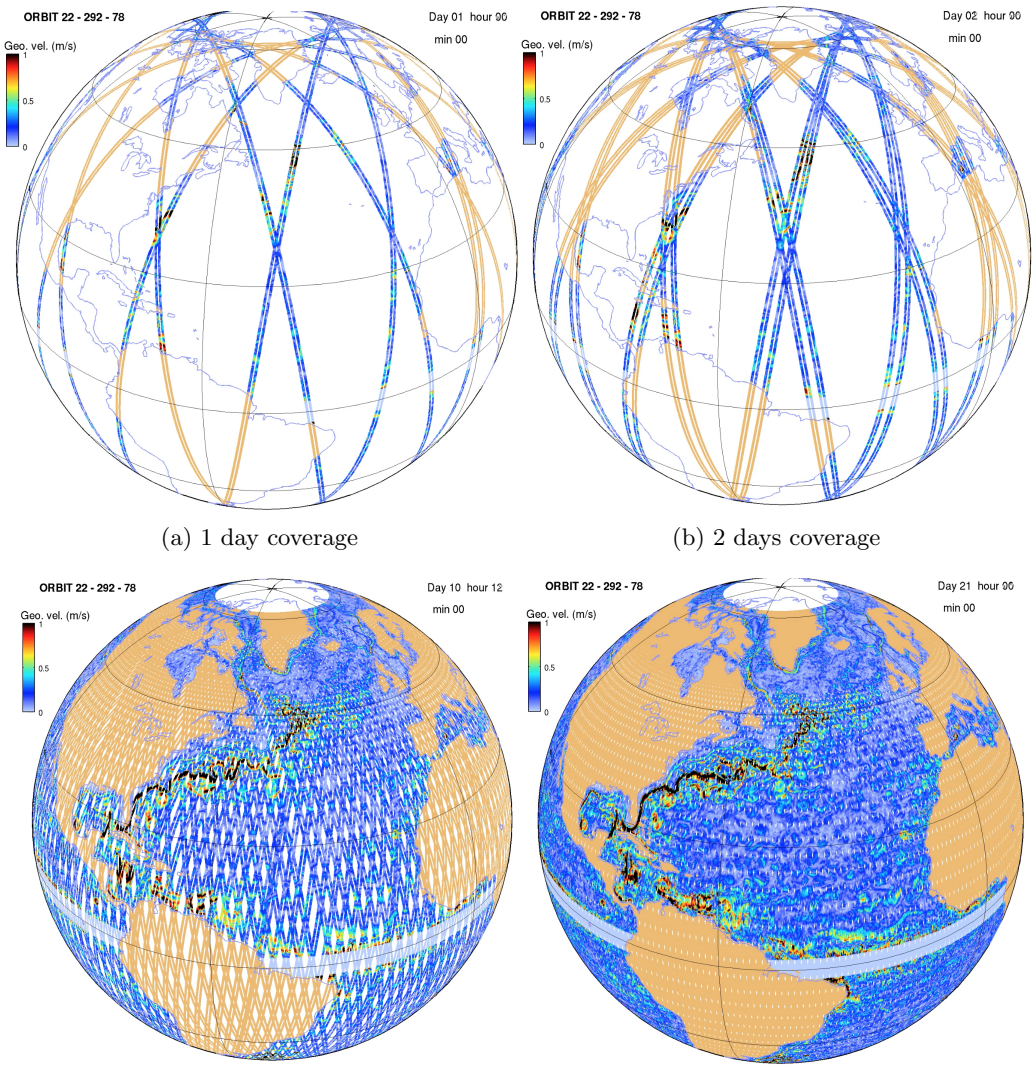
Then, the sea surface pixel can be geolocated in three dimensions without ambiguity to create a two-dimensional map of the SSH. Pulses are then sent regularly along the satellite track and generate overlapping two-dimensional maps of the SSH which can be combined in a single, high resolution, accurate two-dimensional

SSH map[1]. In this method, one antenna emits the pulse while both receive it to generate an image on one side of the track. In the SWOT mission, the antennas send a pulse each on different sides of the track and measure both echoes to generate an image on each side of the track (the two pulses use different polarizations so that their echoes can be separated), as illustrated in Figure 3.2.

Consequently, while previous altimetry missions only provided one-dimensional measurements at the nadir along the satellite track, the KaRIn measurements are two-dimensional over two 50 km wide half-swaths separated by a 20 km gap around the nadir, as illustrated in Figure 3.3. Over the ocean, SSH measurements are provided up to a resolution of 2 km along and across track. In practice, based on the theoretical error budget [Esteban-Fernandez and Peral, 2018], an effective resolution of 15 km should be expected for the ocean [Morrow et al., 2018]. Beyond its main instrument, SWOT also carries a Jason class nadir altimeter, and a 2-beam radiometer used to correct the path delay due to the water vapour content of the atmosphere.

After an initial calibration-validation phase of six months, the satellite was placed on a 'science' orbit. The science orbit is a non sun-synchronous phased orbit, with a repeat cycle of 21 days. The 21-day cycle contains two sub-cycles of 1 and 10.5 days, respectively. During the first day, the consecutive orbits are separated by approximately 26°of longitude at the Equator (Figure 3.4a). The next day, the pattern of the first day is repeated around 2°westward, leaving a gap of the width of a SWOT swath between the two closest tracks (Figure 3.4b). After 10.5 days, half the globe has been observed, with gaps of at most the width of a swath (Figure 3.4c) between each track. The satellite then fills the gaps between the swaths in another 10.5 sub-cycle and reach a near global coverage between +/- 77.6 °after 21 days (Figure 3.4d)

[1]The SAR principle presented here is a simplification of the complete procedure, which is beyond the scope of this report. A more detailed overview can be found in Chapter 2 of Stammer and Cazenave [2018]



(a) 1 day coverage

(b) 2 days coverage

(c) 10 days coverage

(d) 21 days coverage

Figure 3.4: Visualisation of the SWOT orbit coverage, after 1, 2, 10 and 21 days. The full animation can be found at `https://www.aviso.altimetry.fr/gallery/entry_182_swot_scientific_orbit_animated_ground_tracks_on_a_globe.html` .

3.2 Theoretical error budget

A complete overview of the error budget and the science requirements of the SWOT mission has been conducted by Esteban-Fernandez and Peral [2018]. Based on a previous version of this report, Gaultier et al. [2016] developed a Python package to simulate SWOT SSH measurements with a realistic geometry and error. This package is referred to as the SWOT simulator in this manuscript. The simulator interpolates global SSH data on a regular grid provided by the user onto the SWOT grid at a resolution up to 2 km, and adds random errors with power spectrum and geometry in accordance with the results of Esteban-Fernandez and Peral [2018]. Six main error sources are implemented in the SWOT simulator, as described below.

KaRIn error: The KaRIn error refers to white noise from multiple sources such as the speckle noise (random additive/destructive interference between the multiple reflectors contained in a single pixel), or the 'surfboard' effect. The latter represents the variability in the range measurement due to the variable angle of incidence between waves and the iso-range lines, introducing an error which decreases with the across-track distance. On the other hand, the speckle increases with the across-track distance. The 'surfboard effect' error increases quickly with the Significant Wave Height (SWH) until it becomes predominant over the speckle noise, as shown in Figure 3.5.

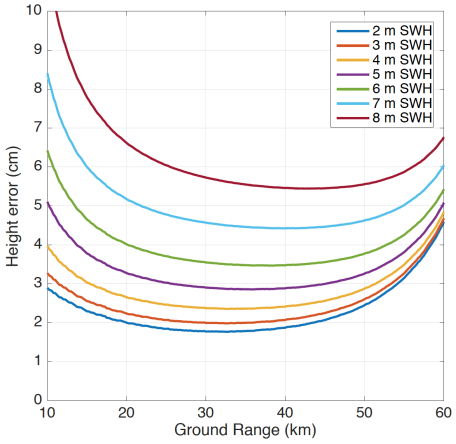


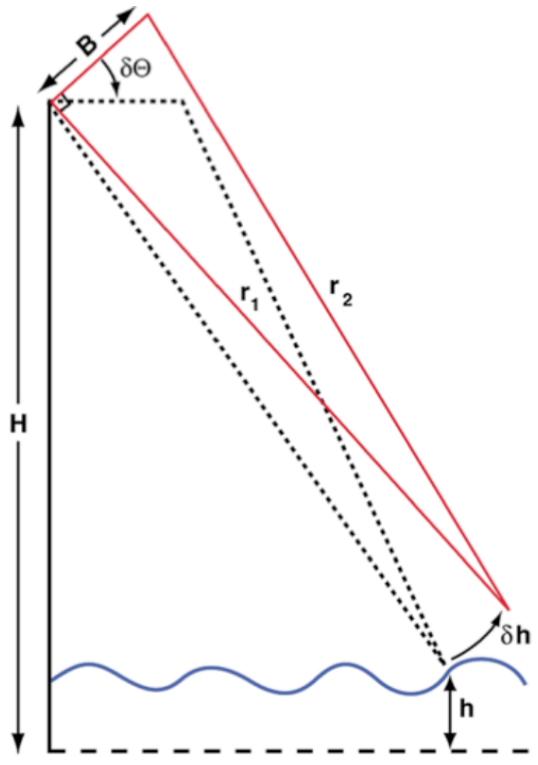
Figure 3.5: Reproduction of figure 34 from Esteban-Fernandez and Peral [2018]. Standard deviation of the KaRIn error at 1 km resolution as a function of the distance from the satellite nadir.

In the SWOT simulator, this error is generated directly as a sample from a standard normal distribution to which the standard deviations from Figure 3.5 are applied.

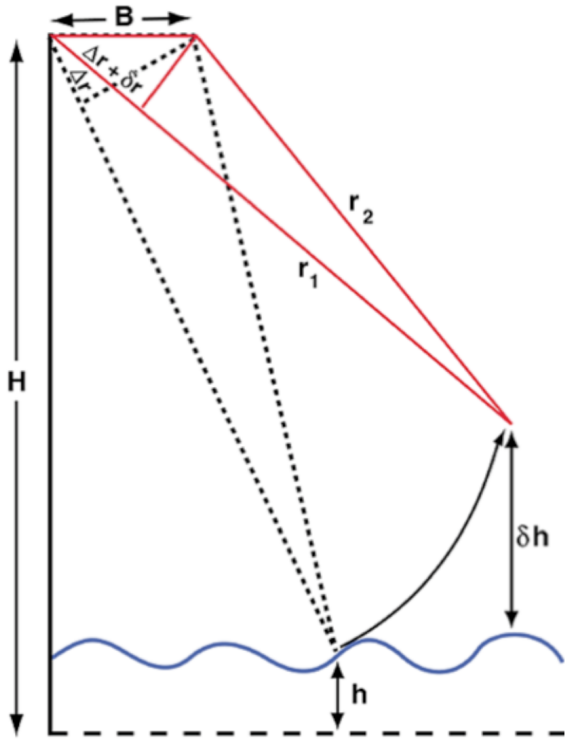
Roll error The measure of the SSH is very sensitive to the roll of the satellite, *i.e.*, rotations of the satellite in the plane normal to its velocity vector. The main source of roll error is a knowledge error on the attitude of the whole satellite. The main measurement of the satellite attitude is based on gyroscopes, which measure attitude variations rather than absolute angles. As such, roll angle measurements tend to drift slowly. The mission specifications thus require the roll angle knowledge error to drift 'slowly enough' to be calibrated rather than requesting an absolute error bound. This results in a roll angle error relatively large in amplitude but constrained to low frequencies; and for which cross-calibration should be critical. The roll error is thus highly correlated along track. Deformations of the KaRIn mechanical system due to vibrations or thermal changes also contribute to the knowledge error on the roll angle, but to a lesser extent, and are also mostly composed of slow variations. As shown in Figure 3.6a, an error on the roll angle translates as an error both on the position and elevation of SSH measurements. While the position error can be neglected compared to the grid resolution, the elevation error is a significant contributor to the SSH error budget. Due to the geometry of the measurements, the SSH error is zero at the nadir and increases linearly with the across-track distance. The error is thus fully correlated across-track within a half-swath while the two half-swaths are fully anti-correlated.

In the SWOT simulator, a random one-dimensional (along track) roll angle is generated according to the power spectrum described in Esteban-Fernandez and Peral [2018]. The linear SSH error can then be deduced from the roll angle through simple geometry as in Figure 3.6a.

Phase error The phase error refers to the sum of all the contributions to the error on the measure of the phase difference between the pulse echoes received on each antenna, and thus the measure of the difference between the range measurements of each antenna (δ_r in Figure 3.6b). As for the roll error, the phase error translates into a (neglected) position error and an elevation error that increases linearly with the across-track distance. However, while the roll angle error affects the whole swath (*i.e.*, both half-swaths have the same error slope), the phase error is not the same on the left and right half-swaths, which have different across-track error slopes. The phase error is thus fully correlated across track within each half-swath, but uncorrelated between half-swaths. Like the roll error, the phase error is specified as a drift in the mission requirements. One of its main causes is temperature driven drifts between the electronic subsystems of each antenna, or distortions of the KaRIn



(a) Reproduction of Figure 38 of Esteban-Fernandez and Peral [2018]. Propagation of a roll angle error δ_θ to an SSH error.



(b) Reproduction of Figure 42 of Esteban-Fernandez and Peral [2018]. Propagation of a relative range error δ_r linked to a phase error to an SSH error.

Figure 3.6: Effect of the roll and phase drift errors on the SSH measurement.

mechanical system. As the thermal environment of the satellite evolves slowly, this error is highly correlated in the along-track direction.

In the SWOT simulator, a random one-dimensional (along track) phase error is generated for each half-swath according to the power spectrum described in Esteban-Fernandez and Peral [2018]. The linear SSH error of each half-swath can then be deduced from the phase error through simple geometry as in Figure 3.6b.

Timing error The timing error refers to errors on the measurement of the back and forth propagation time of the pulse. This error translates into an error on the range measurement fully correlated across track within each half-swath, while the two swaths are not intercorrelated. Drifts in the timing system evolve slowly, and the error is correlated along track.

In the SWOT simulator, a random one-dimensional (along track) timing error is generated for each half-swath independently according to the power spectrum described in Esteban-Fernandez and Peral [2018]. The linear SSH error of each half-swath can then be deduced from the timing by converting it into a range using the estimated propagation speed of the pulse.

Baseline error Baseline errors refer to an error in the estimation of the distance separating the two antennas (B in Figure 3.1). Baseline errors are mostly due the dilation of mechanical system due to changes in the thermal environment. As these changes occur slowly, the error is correlated along track. The error is zero at the nadir and increases quadratically with the across-track distance. The error is fully correlated across track throughout

the whole swath.

In the SWOT simulator, a random one-dimensional (along track) baseline error is generated for the whole swath according to the power spectrum described in Esteban-Fernandez and Peral [2018]. The quadratic SSH error can then be deduced from the baseline error by converting it into a range using simple geometry.

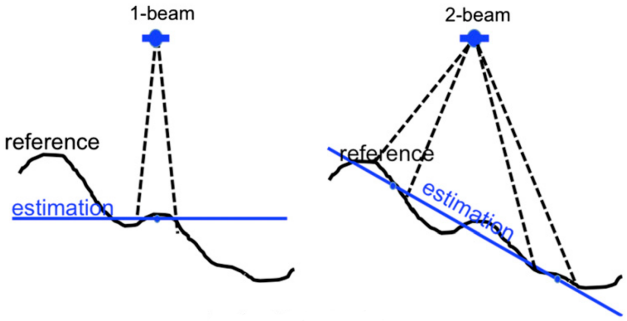


Figure 3.7: Difference between the estimated and actual water vapour content with an ideal (*i.e.*, error free) one-beam or two-beam radiometer.

Wet troposphere error The wet troposphere error represents a range error due to an inaccurate estimation of the propagation speed of the pulse and its echo throughout the atmosphere. The variability of the propagation speed of the pulse is (mostly) linked to the variability of the water vapour content in the troposphere. In traditional nadir altimetry, a radiometer estimates the integrated water vapour content under the nadir, and the propagation speed can be corrected accordingly. In this set up, the wet troposphere error is only due to an inaccurate estimation of the water vapour content by the radiometer. While radiometer errors also contribute to the wet troposphere error in wide-swath altimetry, it is no longer the dominant term. In wide-swath altimetry, we need to account for the across-track variability of the water vapour content in the troposphere. The SWOT mission carries a two-beam radiometer that collects measurements 40 km away from the nadir on both sides of the track. As a consequence, even with an ideal radiometer, the water vapour content can only be approximated by a slope across track. The difference between the actual water vapour content and this estimated slope creates a range error correlated along and across track.

In the SWOT simulator, a random two-dimensional water vapour signal is generated following the global average of the power spectrum derived by Ubelmann et al. [2014]. The two radiometer measurements are estimated from the water vapour signal over Gaussian footprints with additional errors as specified in Esteban-Fernandez and Peral [2018]. A variation of the propagation speed and thus of the range are then deduced from the difference between the actual water vapour signal and the estimated slope.

Figure 3.8 shows an example of simulated SSH errors generated with the SWOT simulator with a SWH of 2 m. The observations are generated at a 12.5 km resolution as the full 2 km resolution is excessive for the 1/4° resolution of the target ORCA configuration of NEMO (ORCA025_Z75). The along-track length of the sample corresponds to a 20 minute time window (*i.e.*, one model time step in the ORCA025_Z75 configuration). The largest error terms are the roll and phase errors, which are both characterised by long-range along-track correlation length scales (of the order of magnitude of the time window or more). As mentioned earlier, this result is expected in the mission requirements, where these errors are not required to be below an absolute upper bound but rather to drift slowly enough to be cross-calibrated.

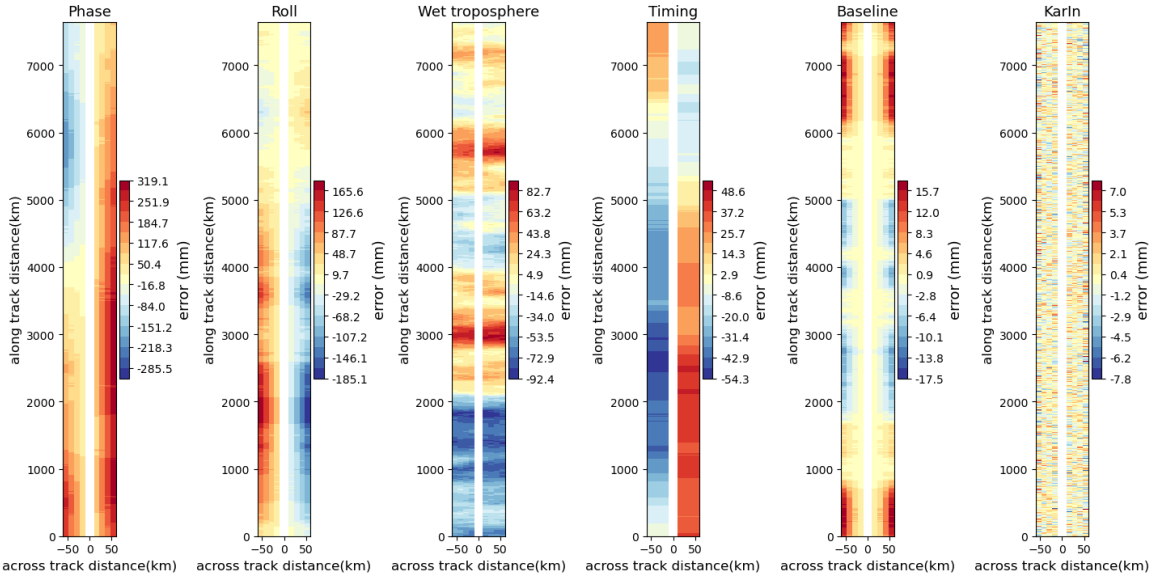
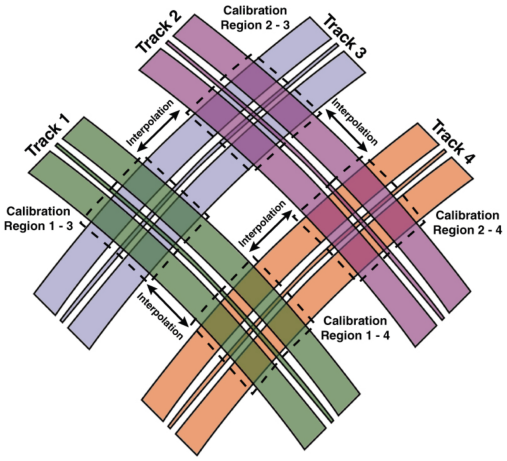


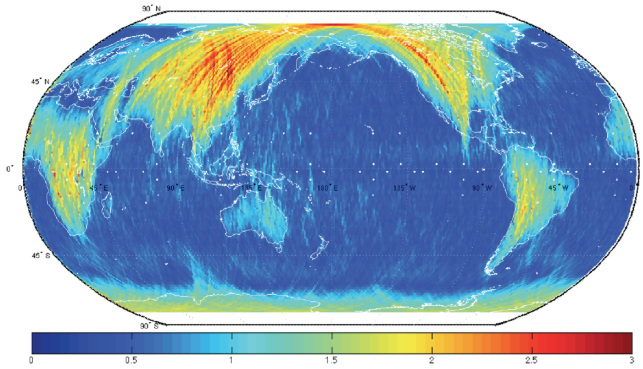
Figure 3.8: Sample of simulated SSH errors at a 2 km resolution with a SWH of 2 metres. The length of the sample corresponds to 20 minutes of flight, or about 8000 km along track.

Cross-calibrated phase and roll errors pre-generated using inversion methods detailed in Section 6.2 of Esteban-Fernandez and Peral [2018] are supplied with the SWOT simulator and interpolated on the SWOT grid. The across-track slopes of the roll and phase errors are estimated at cross-over points and interpolated between them, as illustrated in Figure 3.9a. The method cannot distinguish between roll and phase errors due to their similar geometry across track, and the correction is thus applied directly to their sum in the simulator. As the method can only use cross-over points on the ocean, the residual roll and phase error is larger on land and near the coast. This is illustrated with the residual RMS of the error (expressed as a slope angle in μrad) in Figure 3.9b. Figure 3.10 shows an example of simulated SSH errors with cross-calibration on the roll and phase errors, and the same parameters as in Figure 3.8 (resolution of 12.5 km, SWH of 2 m).

The cross-calibration significantly reduces the amplitude and correlation length scale of the roll and phase errors; however, it might not be appropriate for all applications. Indeed, the cross-calibration algorithm correction of a given measurement depends on cross-over points located further in time in the cycle. As a consequence, the cross-calibrated data are not available in real time, but with a delay of approximately 10 days. While this lag is not an issue for reanalyses, it makes the data less relevant for near-real time ocean forecast applications and Numerical Weather Prediction (NWP) applications involving coupled ocean-atmosphere models. Ideally, this external cross-calibration would be replaced by an online bias estimation procedure (*e.g.*, variational bias corrections). One of the difficulties that would come with this method is to estimate which portion of the error would be caught by the variational bias correction, and which portion should be accounted for in $\mathbf{R}$. Alternatively, other cross-calibrated products more appropriate for near-real time DA might become available in the future. These products could use fewer cross-over points, or only cross-over points located before the actual measurements.



(a) Reproduction of Figure 42 of Esteban-Fernandez and Peral [2018]. Visualization of the cross-over points of the SWOT swaths. The differences between swaths at cross-over points can be used to estimate the errors taking the form of across-track slopes. The error slopes are interpolated between the cross-over points.



(b) Reproduction of Figure 60 of Esteban-Fernandez and Peral [2018]. RMS of the residual roll and phase error (expressed as a slope angle in μrad) after the cross-calibration.

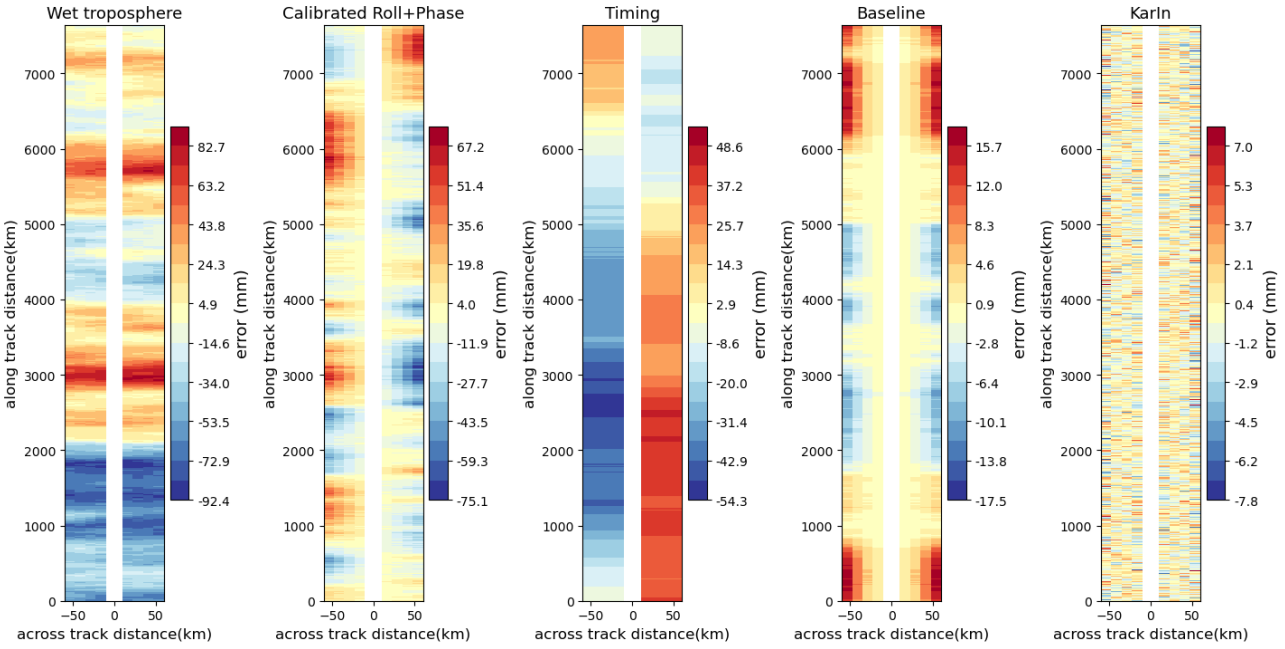


Figure 3.10: Sample of simulated SSH errors at a 2 km resolution with a SWH of 2 m. The length of the sample corresponds to 20 minutes of flight, or about 8000 km along track. The roll+phase error refers to the residual of the roll and phase errors after cross-calibration.

Nadir altimeter errors Two error sources are modelled for the nadir altimeter. The altimeter error is a random noise correlated along track whose spectrum is based on the error spectrum of previous Jason-class nadir altimeters. The second error source is a residual path delay error after the wet troposphere correction. We can expect this error to be correlated with the KaRIn wet troposphere error, as the water vapour content and radiometer measurements are common to both measurements, but the SWOT simulator does not account for this correlation. Instead, the nadir wet troposphere error is a slightly correlated along-track error independent of the KaRIn measurements. Figure 3.11 shows a sample of nadir SSH errors generated by the SWOT simulator.

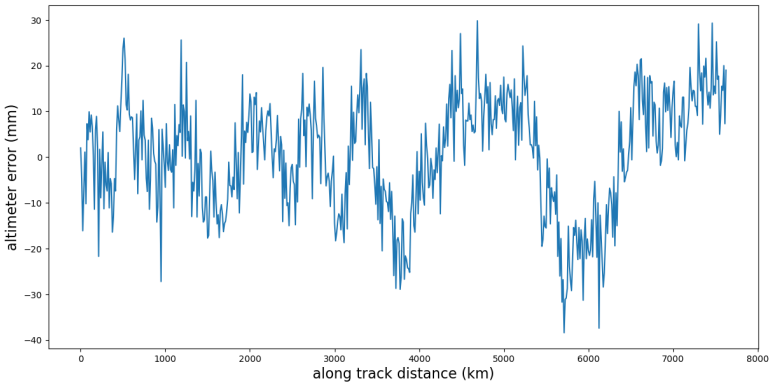
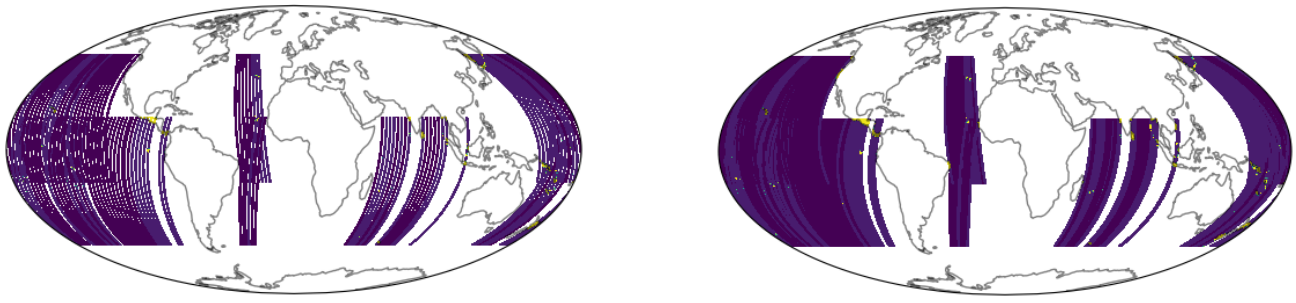


Figure 3.11: Sample of simulated nadir SSH errors at a 2 kilometres resolution. The length of the sample corresponds to 20 minutes of flight, or about 8000 km along track.

3.3 Estimation of error correlations

To estimate the correlations in the SWOT errors, we need to compute statistics from a sample of independent realizations of the error. The along-track length of the error samples is chosen to correspond to a 20-minute time window, or approximately 8000 km. All the errors in the SWOT simulator are generated independently of the absolute position of the satellite. However, as mentioned in the previous section, the effectiveness of the cross-calibration algorithm depends on the absolute position (see Figure 3.9b), which implies that the residual roll+phase error after the cross-calibration depends on the absolute position of the satellite. We ignore this variability in this first example, but further studies could consider correlation length scales varying with their proximity to land to account for this spatial disparity. In order to avoid unwanted inter-correlations between the samples, we choose to extract a sample starting at 50° S in each ascending satellite pass and at 50° N in each descending satellite pass to ensure a separation of at least 10000 km along-tracks between samples. From these samples, only those containing 5% of land area or less are used. The locations of these samples for a half or full orbit repeat cycle are shown in Figure 3.12b.



(a) Coverage after 10.5 days.

(b) Coverage after 21 days.

Figure 3.12: Area covered by all the SWOT samples after a 10.5 day sub-cycle and a full 21-day cycle, respectively. The (very small) areas highlighted in yellow are the land areas covered by the samples within the 5% tolerance on each sample.

Using the ensemble of SWOT data, we estimate the standard deviations and correlations associated with each component of the SWOT error. As the standard deviations proved to be invariant along the satellite track, only across-track profiles are shown in Figure 3.13. The across-track position dependence of the roll, phase, timing and baseline error appears clearly in their standard deviations. Indeed, the roll and phase errors increase linearly with the distance from the nadir, the timing error is constant across-track, and the baseline error increases quadratically with the across-track distance. The shape of the across-track standard deviation profile of the KarIn error is consistent with the theoretical profile for a SWH of 2 m from Figure 3.5 (the amplitude differs because of the difference in grid size). At this grid resolution, the small-scale KaRIn errors have a much lower amplitude than large-scale errors such as the residual roll and phase error or the wet troposphere error. The phase, roll and wet troposphere errors dominate the error budget.

An example of correlation functions of the error components are shown in Figure 3.14. The roll and phase without cross-calibration as well as the residual error which would be obtained after their cross-calibration are both shown. The total error of non cross-calibrated data would thus be the sum of all components shown in Figure 3.14 but the roll+phase residual. The total error in the cross-calibrated data would be obtained by replacing the roll and phase error by the residual roll+phase error.

As expected, apart from a residual sampling noise, the KaRIn error is uncorrelated. Very large along-track correlations appear in the phase and roll errors. Within each half-swath, these errors are completely correlated across-track. For the phase error, the two half-swaths are completely decorrelated from each other. For the roll error, the two half-swaths are completely anti-correlated from each other. After cross-calibration, the residual roll+phase error has a much smaller standard deviation than both the uncalibrated roll and phase errors and its along-track correlation length scale is significantly smaller than that of the roll and phase errors. Its across-track correlation structures are close to those of the phase error: fully correlated within each half-swath, but no correlations between the two half-swaths. The timing error also shows a large along-track correlation length scale. Like the phase error, the timing error is fully correlated across track within each half-swath, but the two half-swaths are uncorrelated. The residual wet troposphere error is roughly ellipsoidal across both half-swaths, with a larger length scale along-track. The baseline error show similar correlations along-track but is fully correlated across half-swaths.

The standard deviation of the nadir alimeter error appears to be almost ten times smaller than that of the swath measurements, even with cross-calibration: the swath data contains large correlated errors that do not affect the nadir measurements. As shown in Figure 3.14, the KaRIn noise is the only error term significant at small scales as it is the only one not to contain large correlations. Its standard deviation is under 2.5 mm, lower than the $3 \sim 4$ mm standard deviation displayed by the nadir altimeter, which explains why KaRIn swath measurements are expected to be more accurate than Jason class altimeters at small scales [Morrow et al., 2018].

However, the standard deviation of the nadir altimeter error appears to be more than ten times smaller than that of the total swath error, which is dominated by a large correlated errors. This difference of magnitude between errors prevalent at large spatial scales or small spatial scales highlights the need for a data assimilation algorithm that can account for correlations in SWOT measurement errors. Indeed, if the swath data are assimilated with a diagonal error covariance matrix $\mathbf{R}$ based on the standard deviation profile of the total swath error from Figure 3.13, the analysis will extract less information at small spatial scales from the swath measurement than from the nadir measurement, even though the latter has a larger error budget at this spatial scale.

While separating the different error sources allows a better understanding of the nature of the covariances, what is observed in practice is the sum of all these errors. In Figure 3.16, the correlations are estimated from the sum of all the error terms. The estimation is performed with and without cross-calibration (Figures 3.15 and 3.16, respectively). Without cross-calibration, the correlation length scale is generally very large, and seems larger on

the outer edge of the half-swath than on the inner edge. The across-track correlations are significant within a half-swath but much weaker from one half-swath to another, due to the contribution of the anti-correlated roll error.

With cross-calibration, the along-track length scale is much smaller, and the across-track correlations between the two half-swaths are larger due to a reduced contribution from the roll error. The correlations seem roughly ellipsoidal in that setting, with a much larger length scale along track than across track. The length scale of the across-track correlations is of the same order of magnitude as the width of the swath: the points located on opposite outer edges of the swaths share a correlation of about 50%.

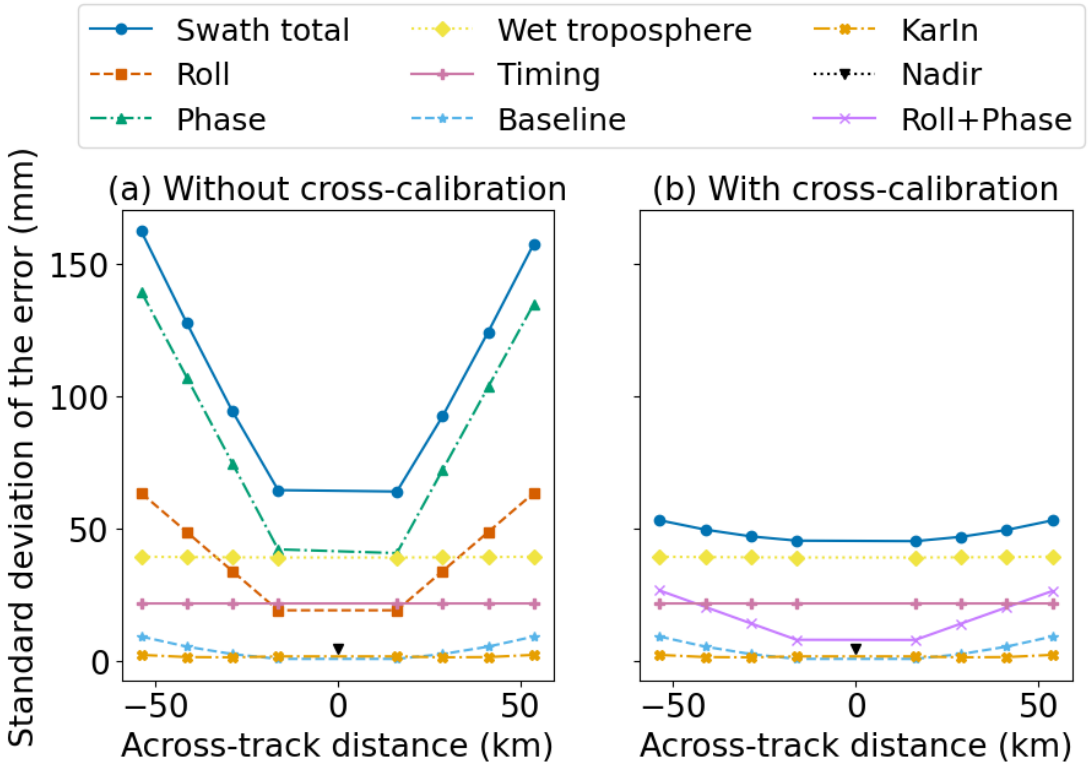


Figure 3.13: Across-track standard deviation profile of the total error on the SWOT swath computed from 924 samples, with a grid size of 12.5 km, and a SWH of 2 m. The standard deviation of the onboard nadir altimeter is also added as a reference.

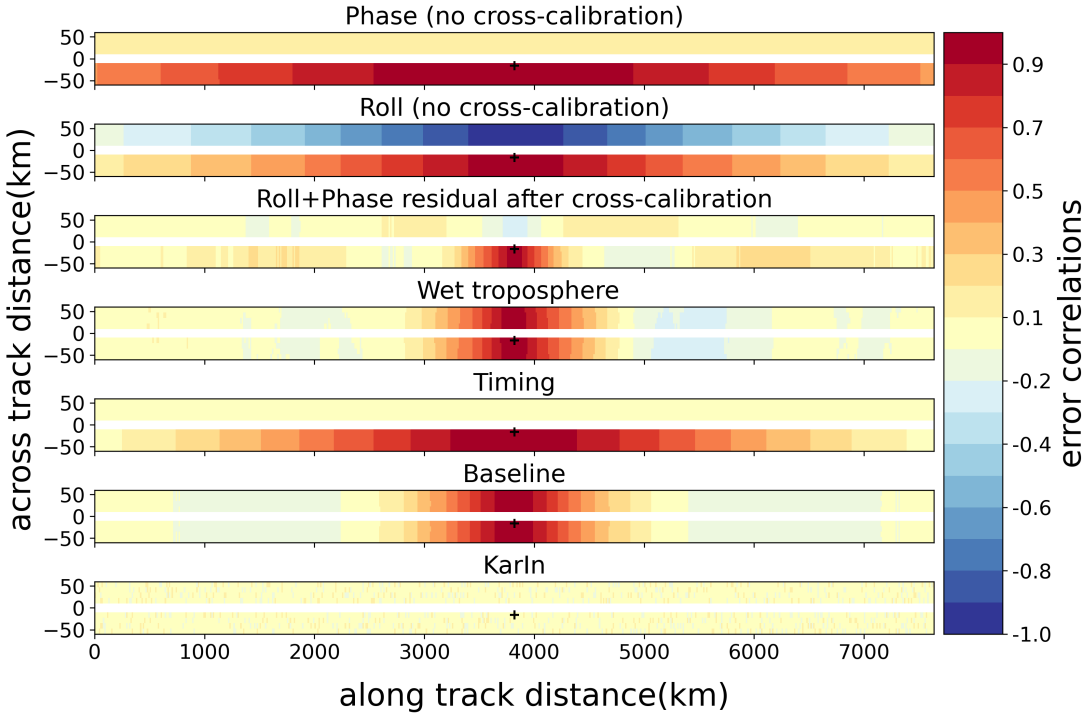


Figure 3.14: Correlations of the different components of the SWOT error with respect to a reference pixel, estimated from 924 samples with a grid size of 12.5 km.

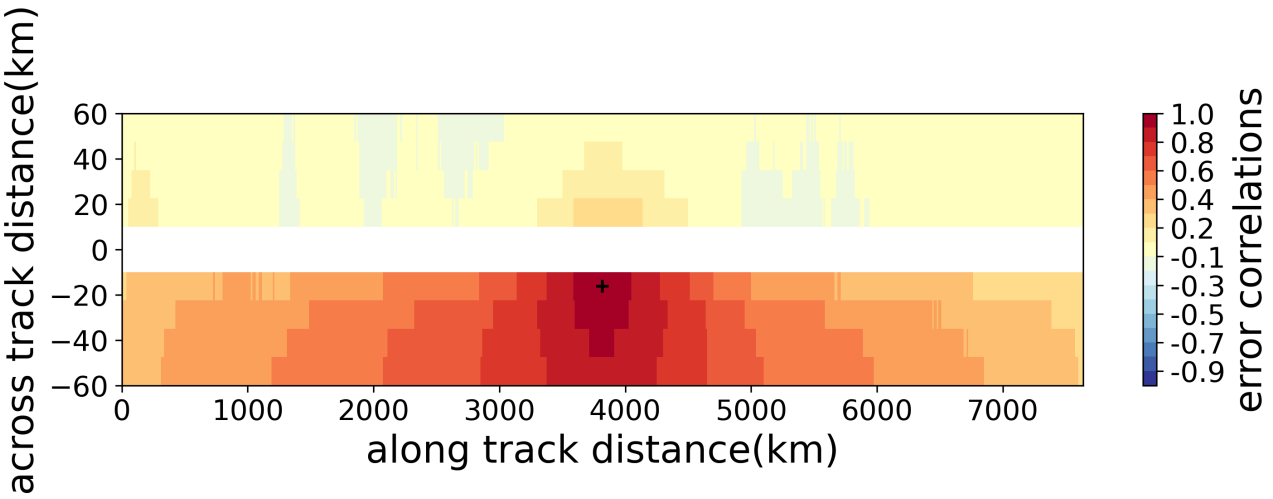


Figure 3.15: Correlations of the total SWOT error with respect to a reference pixel on the inner edge of the lower half-swath, around 1900 km across-track. This estimate is computed from 924 samples, with a grid size of 12.5 km, and a SWH of 2 m. **No cross-calibration** is applied.

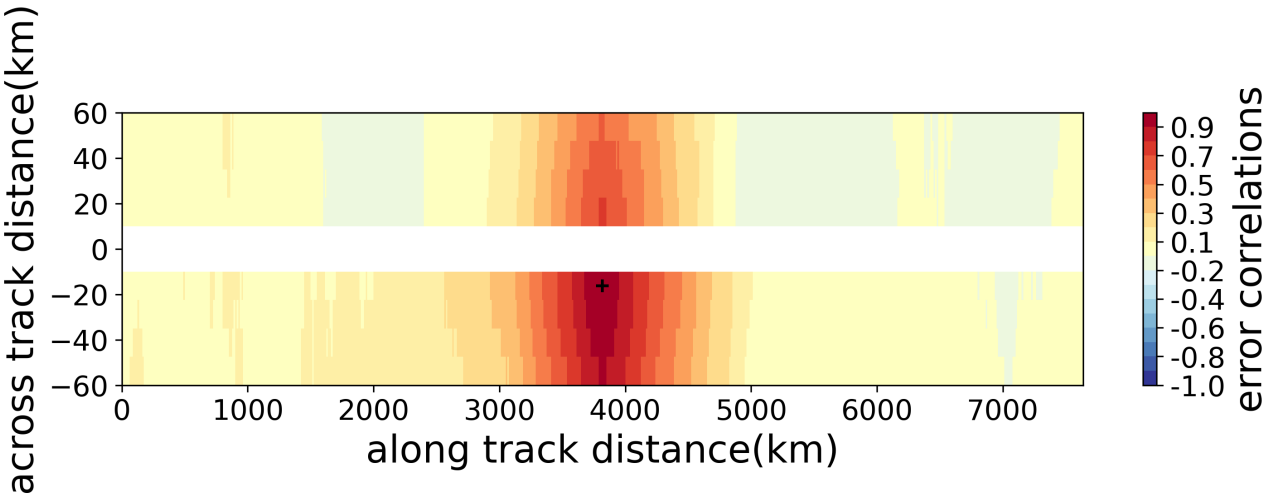


Figure 3.16: Correlations of the cross-calibrated SWOT errors with respect to the reference pixel, estimated from 924 samples with a grid size of 12.5 km. **Cross-calibration** is applied.

3.4 Consequences for data assimilation

As shown in Section 3.3, we can expect SWOT observation errors to present particularly large correlations, and in particular at a potentially longer range than background error correlations. This would create a considerable challenge for current operational data assimilation systems based on a diagonal $\mathbf{R}$. On the other hand, beyond its improved sampling and coverage, we can expect SWOT data to be more accurate at small spatial scales than traditional nadir altimeter data. There is thus a strong potential for improving mesoscale and sub-mesoscale features in ocean analyses.

Due to the challenging nature of this data, most of the attempts at assimilating simulated SWOT data[2] in Observing System Simulation Experiments (OSSEs) found in the literature do not use a full error budget but only white noise when creating synthetic data. For example, [Lopez-Radcenco et al., 2019], [Benkiran et al., 2021], and [Tchonang et al., 2021], only use the (uncorrelated) KaRIn error from the SWOT simulator; [Bonaduce et al., 2018] uses white Gaussian noise with a standard deviation of 1 centimetres; [D'Addezio et al., 2019] uses white Gaussian noise with a standard deviation of 3 centimetres; and [Souopgui et al., 2020] does not add any error. All of these experiments show that assimilating simulated SWOT data free of large correlated errors leads to significant improvements in the reconstruction of mesoscale and sub-mesoscale features in SSH and surface velocity.

However, obtaining similar performance with realistic correlated observation error does not seem possible with current DA systems. In their OSSE with simulated correlated observation error, [Carrier et al., 2016] are able to mitigate the effect of observation error correlations by thinning SWOT data with a 40 km radius, which does not allow them to retrieve the small-scale features that could be resolved by SWOT. [King and Martin, 2021] assimilated data simulated with only the KaRIn error as well as data simulated with the full error budget in an experiment using the Met Office 1.5 km North-West European Shelf analysis and forecast system. While the data with KaRIn error only lead to significant improvements compared to a control run (*e.g.*, -21% RMSE for SSH off-shelf), the data with a full error budget lead to a substantial degradation of the quality of the analysis (*e.g.*, +21% RMSE for SSH off-shelf). [King and Martin, 2021] found that to mitigate this negative impact of correlated observation error on the system, the data had to undergo extensive thinning (and the removal of the outer half of the swath), which also negated most of the positive impact of SWOT data. This result was confirmed by [Le Hénaff et al., 2008], who assimilated SWOT data with either white noise or with white noise and an additional correlated roll error (without using the SWOT simulator). They also show that assimilating SWOT data with only KaRIn error strongly reduces the error variances of the analysis for multiple variables. Additionally, they show that the analysis obtained by assimilating of SWOT data with correlated observation error *accounted for* with a non-diagonal $\mathbf{R}$ – which can be built and inverted explicitly due to the small size of their system – retains most of the benefits observed with data without correlated observation error. On the other hand, with a diagonal $\mathbf{R}$ (with variances accounting for the roll error), the assimilation of SWOT data with correlated observation error leads at best to a performance similar to that of a traditional nadir altimeter.

These results confirm that simulating only KaRIn error is not sufficient to estimate the impact of SWOT data on a current data assimilation system, as it removes the main hurdle that data assimilation will need to overcome to see the benefits expected of SWOT. However, we can anticipate similar performance if the data assimilation systems used in these OSSEs were able to account for observation error correlations.

[2]Real SWOT data are still for the most part in an experimental stage at the time of writing.

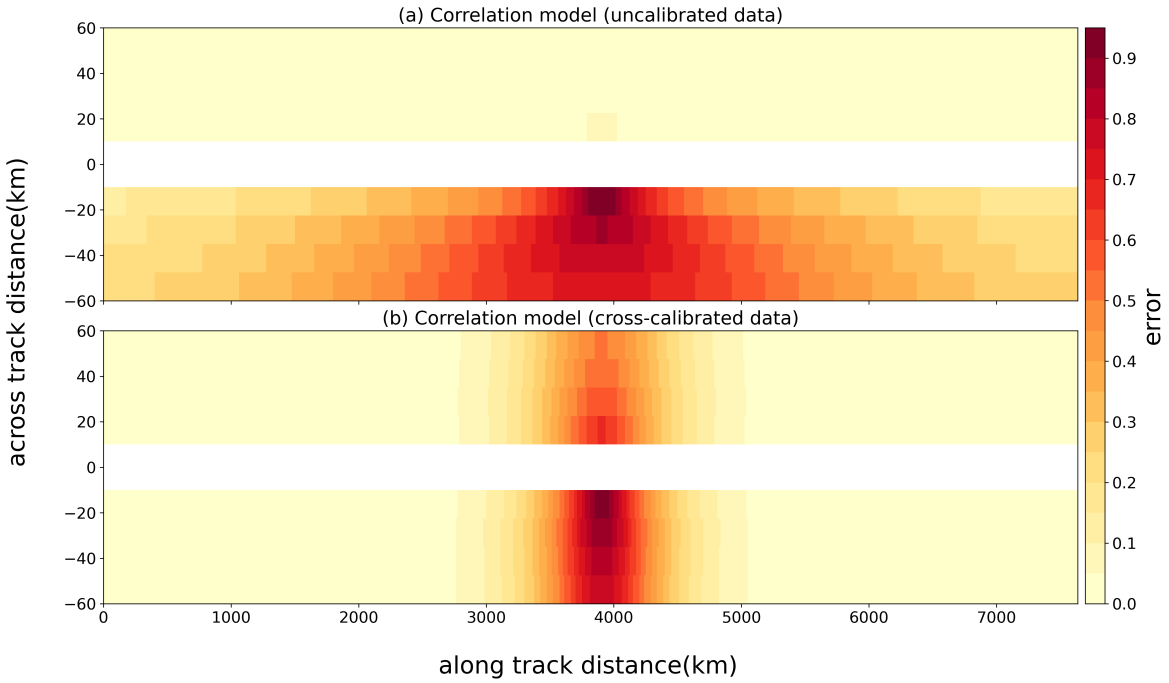


Figure 3.17: Diffusion-based correlation model of SWOT error correlations. (a) Without cross-calibration, the along-track length scale ranges from 3000 km on the outside of the swath to 500 km inside the swath; the across-track length scale is approximately equal to 75 km except between the two half-swaths where it is equal to 0. (b) With cross-calibration, the along-track length scale ranges from 310 km on the outside of the swath to 260 km inside the swath; the across-track length scale varies between 10 km and 60 km. The number of diffusion iterations is set to 2 in both cases. Mass lumping (presented in Section 6.2.1) is used.

3.5 Accounting for SWOT observation error correlations

We demonstrate here how a finite-element method (FEM)-based diffusion operator [Guillet et al., 2019] can be used to account for correlated errors in SWOT data. In Figure 3.17, we show an example of spatial correlations associated with a diffusion operator used to model the error correlations of uncalibrated data (shown in the top panel of Figure 3.15) and the error correlations of the cross-calibrated data (shown in the bottom panel of Figure 3.15). The correlations have been computed using an FEM-based diffusion operator coded in Python, which has been used as a prototype for the developments in NEMOVAR. Details about the implementation of such operators for 2D data can be found in Section 6.2.1. The length scales used to parameterise the diffusion operator have been computed by minimising the difference in the 2-norm between correlations estimated from simulated SWOT error samples and their modelled counterparts. The minimum has been computed iteratively using the Scipy implementation of the Nelder-Mead algorithm [3]. To reduce the numbers of parameters of the minimisation, we assumed the correlation length scales to be invariant along-track, and symmetric across the nadir of the satellite. Comparing Figure 3.17 with Figures 3.16 and 3.15 shows that the diffusion operator is able to approximate SWOT-like correlations with a degree of accuracy likely to surpass our knowledge of the actual error correlations.

Instead of accounting for observation error correlations with a non-diagonal $\mathbf{R}$, another possibility is to avoid altogether assimilating the most error-prone features of SWOT data. Metref et al. [2019] and Metref et al. [2020]

[3] https://docs.scipy.org/doc/scipy/reference/optimize.minimize-neldermead.html

suggest assimilating SWOT data with a 'correlated error-reduction' scheme. This scheme fits a parametric model based on typical SWOT error patterns (*e.g.*, constant, linear and quadratic components across-tracks) to the SWOT observations. The error estimated by this "error-model" is then removed from the observations and from vectors to which the observation operator is applied. They show that this approach effectively mitigates the contamination of the analysis with unphysical features associated with these errors. However, this method does not discriminate between patterns across-track due to error or geophysical signals, and might thus prevent the extraction of the latter from the observations. For that reason, and because not all correlated errors affecting SWOT follow such predictable across-track patterns (*e.g.*, wet troposphere residual error), this method does not offer a complete treatment of correlated observation error. However, this method could be used in addition to a non-diagonal $\mathbf{R}$ in the event that the extremely large scale of the roll or phase errors make them difficult to account for without affecting the convergence and robustness of the minimisation algorithm (as discussed in Chapters 4 and 5).

3.6 Summary

The new technology of the wide-swath altimeter used by SWOT should provide unprecedented insight on the sub-mesoscale dynamics of the ocean which were up to now poorly observed. The interferometry technique used with the KaRIn instrument allows two-dimensional measurements with a 2 km sampling within a 100 km wide satellite swath rather than one-dimensional measurements at the satellite nadir from traditional altimeters. SWOT data could thus prove to be a great asset for data assimilation systems for high-resolution models by providing an otherwise lacking observational constraint on the small spatial scales resolved by these models. Many studies demonstrate this potential by showing substantial decreases in analysis RMSE in OSSEs when assimilating simulated SWOT data. However, when simulating SWOT data, these studies mostly assume that the only error source is white noise with a standard deviation of a few centimetres. According to the SWOT error budget [Esteban-Fernandez and Peral, 2018], this assumption is very optimistic. We have described here multiple sources of correlated error of large amplitude which are likely to affect the real data. The few studies that attempt to assimilate simulated SWOT data with this realistic error budget in systems using a diagonal $\mathbf{R}$ show that instead of a substantial improvement of the analysis, a substantial degradation is obtained due to neglecting observation error correlations. Observation thinning and observation error variance inflation can mitigate the degradation but also severely limits the positive impact of SWOT data, generally reducing its value to that of a traditional nadir altimeter.

Based on synthetic data from the JPL/CNES SWOT simulator, we have estimated the correlations of the SWOT error and shown that significant correlations exist between observation locations thousands of kilometres apart from each other. This explains the limited effect of thinning to mitigate observation error correlations. We also observed that, if the correlated errors are included in the error budget, the standard deviation of the SWOT error largely exceeds that of the error affecting traditional nadir altimeters. It is only at small spatial scales, which are only affected by small-amplitude uncorrelated error, that SWOT data outperforms nadir altimeter data. This strong discrepancy in error amplitude at different spatial scales highlights the need for a *scale-dependent* description of observation error in order to extract features of interest at small spatial scales in SWOT data while rejecting the large-scale error-prone features. We have shown that it would be possible to model the correlations of SWOT error using a diffusion operator with large correlation length scales.

Chapter 4

Article 1: Impact of correlated observation errors on the conditioning of variational data assimilation problems

Using a non-diagonal $\mathbf{R}$ in DA significantly alters the ($\mathbf{B}$-preconditioned) Hessian matrix $\mathbf{S}$ (defined in Equation (2.5)) of the quadratic cost function (2.6). As the convergence of the CG algorithm is known to be strongly related to the eigenspectrum of the Hessian matrix, there could be a risk of degrading the convergence rate of the minimisation. As the minimisation is usually truncated early due to a limited computational budget, a degraded convergence rate could increase the difference between $\mathbf{x}_k$, the solution of the truncated minimisation, and $\mathbf{x}^*$, the solution of the minimisation at full convergence. The choice of observation-error correlation model should thus account for its impact on the convergence rate of the minimisation. As a first step towards planning an operational non-diagonal $\mathbf{R}$, we thus evaluate this impact in a simple system in an article published in *Numerical Linear Algebra with Applications* in 2024 and included hereafter.

Received: 25 November 2022 | Revised: 25 May 2023 | Accepted: 25 July 2023

DOI: 10.1002/nla.2529

RESEARCH ARTICLE

WILEY

Impact of correlated observation errors on the conditioning of variational data assimilation problems

Olivier Goux[1,2] | **Selime Gürol**[1] | **Anthony T. Weaver**[1] | **Youssef Diouane**[3] | **Oliver Guillet**[4]

[1]ALGO-COOP Team, CERFACS / CECI CNRS UMR 5318, Toulouse, France

[2]ISAE-SUPAERO, University of Toulouse, Toulouse, France

[3]Department of Mathematics and Industrial Engineering, Polytechnique Montréal, Montréal, Quebec, Canada

[4]CNRM UMR 3589, Météo-France and CNRS, Toulouse, France

Correspondence
Olivier Goux, CERFACS / CECI CNRS UMR 5318, 42 avenue Gaspard Coriolis, 31057 Toulouse Cedex 01, France.
Email: goux@cerfacs.fr

Funding information
Copernicus Climate Change Serive, Grant/Award Number: C3S_601_INRIA; French National Programme LEFE/INSU

Summary

An important class of nonlinear weighted least-squares problems arises from the assimilation of observations in atmospheric and ocean models. In variational data assimilation, inverse error covariance matrices define the weighting matrices of the least-squares problem. For observation errors, a diagonal matrix (i.e., uncorrelated errors) is often assumed for simplicity even when observation errors are suspected to be correlated. While accounting for observation-error correlations should improve the quality of the solution, it also affects the convergence rate of the minimization algorithms used to iterate to the solution. If the minimization process is stopped before reaching full convergence, which is usually the case in operational applications, the solution may be degraded even if the observation-error correlations are correctly accounted for. In this article, we explore the influence of the observation-error correlation matrix (**R**) on the convergence rate of a preconditioned conjugate gradient (PCG) algorithm applied to a one-dimensional variational data assimilation (1D-Var) problem. We design the idealized 1D-Var system to include two key features used in more complex systems: we use the background error covariance matrix (**B**) as a preconditioner (B-PCG); and we use a diffusion operator to model spatial correlations in **B** and **R**. Analytical and numerical results with the 1D-Var system show a strong sensitivity of the convergence rate of B-PCG to the parameters of the diffusion-based correlation models. Depending on the parameter choices, correlated observation errors can either speed up or slow down the convergence. In practice, a compromise may be required in the parameter specifications of **B** and **R** between staying close to the best available estimates on the one hand and ensuring an adequate convergence rate of the minimization algorithm on the other.

KEYWORDS

condition number, conjugate gradient, convergence rate, covariance operators, diffusion operators, nonlinear weighted least-squares

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2023 The Authors. *Numerical Linear Algebra with Applications* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

An important class of nonlinear weighted least-squares problems arises from the assimilation of observations in atmospheric and ocean models, a procedure known as *data assimilation*. In data assimilation, observations of the state of a system are combined with an *a priori* estimate of the state, called the *background*, to produce an optimal estimate of the state of the system, called the *analysis*. In *variational data assimilation*, the optimal estimate is obtained iteratively by minimizing a nonlinear weighted least-squares cost function that is the sum of two terms: one measuring the model fit to the background state (the background term $\mathcal{J}_b$); the other measuring the model fit to the observations (the observation term $\mathcal{J}_o$), subject to constraints (generally nonlinear) that relate the model state to the observations. The weighting matrices for $\mathcal{J}_b$ and $\mathcal{J}_o$ are defined by an estimate of the inverse of the background and observation-error covariance matrices ($\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$), respectively. Variational data assimilation is widely used for operational state estimation in meteorology and oceanography as it is a practical method for solving nonlinear least-squares problems when the dimensions of the state and observation vectors are huge (typically 10^6 to 10^9).

In variational data assimilation, the cost function is minimized approximately using a truncated Gauss–Newton (GN) algorithm[1] or incremental variational data assimilation as it is known in the data assimilation community.[2] This reduces the nonlinear problem to a sequence of linear sub-problems (quadratic cost functions), each of which is solved iteratively using a preconditioned conjugate gradient (PCG) method.[3] Standard implementations of PCG for data assimilation employ **B** as a first-level preconditioner,[3,4] which we refer to as **B**-PCG hereafter. Besides significantly improving the conditioning of the Hessian matrix,[5] **B**-preconditioning allows the **B**-PCG algorithm to be formulated in a way that avoids the need to specify $\mathbf{B}^{-1}$ explicitly. This is important as most **B** formulations used in practice are not associated with convenient representations of $\mathbf{B}^{-1}$.

There is still a requirement to specify $\mathbf{R}^{-1}$, however. To simplify its specification, practical implementations of **R** tend to have relatively simple structural forms. In the extreme yet common case, **R** is taken to be a diagonal matrix, which amounts to assuming that the observation errors are uncorrelated. This is a poor assumption for certain observations, especially from satellites.[6,7] If observation-error correlations are neglected when they are known to be important then the solution of the weighted least-squares problem will result in a degraded (sub-optimal) analysis and poor exploitation of the assimilated data. To mitigate the former while still using a diagonal **R**, observation data sets are either 'thinned' into a subset of observations or aggregated into 'super-observations' that have reduced error correlations.[8] Furthermore, the observation-error variances are often multiplied by an 'inflation' factor in order to prevent the analysis from overfitting observations that may still have a substantial component of correlated error. However, when observation error is correlated over distances similar to or greater than those of the background error, inflation can actually degrade the analysis.[9] While these methods can alleviate to some extent the inaccuracies associated with a diagonal **R**, they still lead to sub-optimal solutions since potentially valuable observations are excluded and any remaining error correlations from the pre-processed observations are ignored.[10]

Several studies have examined the impact from using non-diagonal representations of **R** to account for spatially correlated errors.[11-14] A general conclusion that arises in most of these works is that accounting for spatial correlations in **R** leads to a more accurate solution, especially for the smaller spatial scales, even with a rather crude correlation model. However, even crude correlation models can lead to impractical representations of $\mathbf{R}^{-1}$. Various correlation models with accessible inverse representations have been proposed in the literature.[15-19] One of the challenges with specifying **R** and $\mathbf{R}^{-1}$ is that observation locations tend to be arranged in an arbitrary and unpredictable way due to the measurement method or quality control procedures that result in observations being removed. This means that correlation models developed for structured grids, like those typically associated with **B**, are not necessarily applicable for **R**.

In this study, we use diffusion operators to model spatial correlations in both **B** and **R**. Diffusion operators can be used to model correlation functions from the Matérn class[20] and provide convenient and inexpensive representations of the associated inverse correlation operators.[21,22] They are popular for representing spatially correlated background error in complex boundary domains such as those encountered in ocean data assimilation.[4,23-26] Furthermore, Reference 18 describes how to adapt these operators to unstructured meshes and hence to make them suitable for **R** and $\mathbf{R}^{-1}$.

The rate of convergence of the conjugate gradient (CG) method is mainly determined by the characteristics of the eigenvalue spectrum of the Hessian matrix.[3,27] As the eigenvalues of the Hessian matrix are strongly dependent on **B** and **R**, we can expect a non-diagonal **R** to have a significant impact on the rate of convergence of **B**-PCG. In operational data assimilation, analyses must be delivered subject to strict computational constraints, which means that the stopping criterion for **B**-PCG is usually set by a maximum allowed number of iterations rather than a measure of the convergence of the solution. Therefore, it is essential to ensure that the rate of convergence of **B**-PCG from the use

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of a non-diagonal **R** is not deteriorated to an extent that it outweighs the benefits brought from specifying a more accurate **R**.

In previous work, Reference 28 analysed the convergence rate for the special case of a diagonal **R** and discussed the influence on the condition number of observation and background accuracy, observation density, and the background-error correlation length-scale. Reference 29 studied the effects of a non-diagonal **R** on the convergence rate of the unpreconditioned CG method and derived general theoretical bounds for the condition number. These results were extended by Reference 30 to the **B**-preconditioned case (**B**-PCG). In both studies, theoretical and experimental results were obtained for the special case where the background- and observation-error correlation matrices are defined as circulant matrices. Their numerical experiments were performed using a particular circulant matrix built from a second order auto-regressive (SOAR) correlation function defined on the one-dimensional (1D) circular domain $\mathbb{S}$.

In this article, we are interested in understanding the sensitivity of the condition number with respect to the basic parameters of the diffusion operators that are used to model background- and observation-error correlations. Correlation functions derived from diffusion operators are controlled by a smoothness parameter M (the number of diffusion iterations) as well as a length-scale parameter L (the square-root of the diffusion coefficient), and thus are more flexible than the SOAR function, which is controlled by a single length-scale parameter. In fact, the SOAR function corresponds to a particular member ($M = 2$) of the family of correlation kernels represented by the 1D diffusion operator. We illustrate how the relative choice of M for **B** and **R** can have a profound effect on the conditioning of the minimization problem.

The organization of the article is as follows. In Section 2, we introduce the weighted least-squares problem underlying variational data assimilation and we outline the solution algorithm based on truncated GN combined with CG. We provide the background theory on CG (and **B**-PCG) that is needed in this article for establishing the theoretical results and for interpreting the results from the numerical experiments with a 1D variational data assimilation (1D-Var) system. We conclude this section with a description of **B** and **R**, exposing the fundamental covariance parameters that control the shape characteristics of the correlation functions as well as the conditioning of the CG minimization. The formulation of **B** and **R** in terms of diffusion operators depends on theoretical results that are summarized in Appendix A. In Section 3, we study the eigenvalue spectrum of the Hessian matrix under the assumption that the observations are sampled from a regular grid and evenly distributed. While restrictive, this assumption allows us to derive analytical bounds on the condition number and to gain crucial insights on the convergence of **B**-PCG in more general frameworks. First, we present the general bounds that were derived by References 29 and 30. Then, we derive specific bounds that take into account the structural properties of the diffusion operators used to model **B** and **R**. We relegate the technical details of the proofs of a key theorem and associated corollaries to Appendices B,C, and D. In Section 4, we present the results from numerical experiments with the 1D-Var system to examine the sensitivity of the convergence of **B**-PCG to the parameters of **B** and **R**. These results show a strong sensitivity of the condition number, and hence convergence rate, to the correlation parameters. We argue that the parameter values should be chosen as a compromise between specifying the most accurate correlation model on the one hand and achieving a satisfactory convergence rate for the CG minimization on the other. We provide a summary and conclusions in Section 5.

2 | THE WEIGHTED LEAST-SQUARES PROBLEM

2.1 | Problem formulation: Variational data assimilation

Variational data assimilation provides an estimate of the physical state of a system by combining *a priori* information (the background state) and observations, together with information about their uncertainties. Here, we will use mathematical notation that is standard in meteorological and ocean data assimilation.[31] Assuming unbiased Gaussian error statistics for the background state and observations, the estimation problem can be formulated as a nonlinear weighted least-squares problem defined by the cost function

$$\min_{\mathbf{x}} \mathcal{J}(\mathbf{x}) = \frac{1}{2} \| \mathbf{x} - \mathbf{x}_\mathrm{b} \|^2_{\mathbf{B}^{-1}} + \frac{1}{2} \| \mathcal{H}(\mathbf{x}) - \mathbf{y}_\mathrm{o} \|^2_{\mathbf{R}^{-1}}, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector to be optimized and $\mathbf{x}_\mathrm{b} \in \mathbb{R}^n$ is the background estimate of the state vector. The vector of observations is $\mathbf{y}_\mathrm{o} \in \mathbb{R}^m$, and $\mathcal{H}(\cdot)$ is the observation operator, which maps an estimate of the state of the system to its equivalent in observation space. In general, $\mathcal{H}(\cdot)$ is nonlinear and non-bijective. In four-dimensional variational

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

assimilation (4D-Var), $\mathcal{H}(\cdot)$ would contain the forecast model operator, $\mathbf{x}$ would be the initial state vector, and $\mathbf{y}_o$ would be a vector that concatenates observations distributed over a given time window. The unbiased Gaussian distributions of the background and observation errors are characterized statistically by the covariance matrices $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{m\times m}$, respectively. By definition, $\mathbf{B}$ and $\mathbf{R}$ are symmetric positive-definite (SPD) matrices. The inverse covariance matrices $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$ define inner products in the background and observation spaces, and are used as weighting matrices in the cost function (1) where $\|\mathbf{x}\|^2_{\mathbf{P}} = \mathbf{x}^{\mathsf{T}}\mathbf{P}\mathbf{x}$ denotes the squared $\mathbf{P}$-norm of a vector. The analysis is the global minimizing solution: $\mathbf{x}_a = \arg\min \mathcal{J}(\mathbf{x})$.

Truncated Gauss–Newton (GN),[1] which is known as incremental variational assimilation in the meteorological and ocean data assimilation communities,[2] is a common method for finding an approximate minimum of the nonlinear cost function (1). Truncated GN approaches the solution iteratively by solving, on each GN iteration k, the linearized sub-problem

$$\min_{\delta\mathbf{x}} J(\delta\mathbf{x}) = \frac{1}{2}\|\mathbf{x}_k - \mathbf{x}_b + \delta\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathbf{H}_k\delta\mathbf{x} - \mathbf{d}_k\|^2_{\mathbf{R}^{-1}}, \tag{2}$$

which is a quadratic approximation of the non-quadratic cost function (1) in a neighborhood of the current iterate $\mathbf{x}_k$. In Equation (2), $\mathbf{H}_k \in \mathbb{R}^{m\times n}$ is the observation operator linearized about $\mathbf{x}_k$, and $\mathbf{d}_k = \mathbf{y}_o - \mathcal{H}(\mathbf{x}_k) \in \mathbb{R}^m$ is the misfit between the observation vector and the current iterate mapped to observation space. If $\delta\mathbf{x}_{k+1}$ denotes the solution of (2) then the estimate of the state is updated according to

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \delta\mathbf{x}_{k+1},$$

where $k = 0, \ldots, K-1$ and $\mathbf{x}_0 = \mathbf{x}_b$ (in general). In data assimilation applications with atmospheric or ocean models, the maximum number of GN iterations is typically very small ($K < 10$) for computational reasons.

The quadratic sub-problem (2) can be rewritten in standard quadratic form

$$\min_{\delta\mathbf{x}} J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^{\mathsf{T}}\mathbf{A}_k\delta\mathbf{x} - \mathbf{b}_k^{\mathsf{T}}\delta\mathbf{x} + c_k, \tag{3}$$

where

$$\mathbf{A}_k = \mathbf{B}^{-1} + \mathbf{H}_k^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{H}_k,$$

is the SPD approximation of the Hessian matrix of the nonlinear cost function,

$$\mathbf{b}_k = \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}_k) + \mathbf{H}_k^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{d}_k,$$

is the negative gradient of the nonlinear cost function with respect to the current iterate $\mathbf{x}_k$, and $c_k = J(\mathbf{0})$ is a scalar. Satisfying the optimality condition of the quadratic sub-problem (3) requires solving the linear system

$$\mathbf{A}_k\delta\mathbf{x} = \mathbf{b}_k.$$

For our target applications, the dimension (n) of the state vector is large and the matrices are generally only available as operators (i.e., via matrix-vector products, not explicit matrices). For this reason, it is very common to solve the quadratic sub-problem iteratively using CG methods.

2.2 | Solving the quadratic sub-problem with the conjugate gradient method

Conjugate gradient (CG) is a Krylov subspace method (see Golub and Van Loan [Reference 32, Sec. 6.7] and Saad [Reference 33, Sec. 10.2]) for solving linear systems where the system matrix is SPD. CG seeks an approximate solution

$$\delta\mathbf{x}_\ell \in \delta\mathbf{x}_0 + \mathcal{K}^\ell(\mathbf{A}_k, \mathbf{b}_k),$$

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where $\delta\mathbf{x}_0$ is the initial approximation and

$$\mathcal{K}^{\ell}(\mathbf{A}_k, \mathbf{b}_k) = \text{span}\{\mathbf{b}_k, \mathbf{A}_k\mathbf{b}_k, \ \dots \ , \mathbf{A}_k^{\ell-1}\mathbf{b}_k\},$$

is the Krylov subspace generated by $\mathbf{A}_k$ and $\mathbf{b}_k$. When $\mathbf{x}_0 = \mathbf{x}_b$, the initial iterate $\delta\mathbf{x}_0 = \mathbf{0}$. Hereafter, we will drop the truncated GN iteration index k for clarity of notation. In order to find a unique solution, CG imposes the orthogonality condition

$$\mathbf{r}_{\ell} \perp \mathcal{K}^{\ell}(\mathbf{A}, \mathbf{b}),$$

where $\mathbf{r}_{\ell} = \mathbf{b} - \mathbf{A}\delta\mathbf{x}_{\ell}$ is the residual at the ℓ-th iteration of CG. As a result, CG minimizes the quadratic cost function given by (3) over the subspace $\delta\mathbf{x}_0 + \mathcal{K}^{\ell}(\mathbf{A}, \mathbf{b})$ [Reference 34, theorem 5.2], so that the ℓ-th iterate $\delta\mathbf{x}_{\ell}$ minimizes the error $\mathbf{e}_{\ell} = \delta\mathbf{x}^* - \delta\mathbf{x}_{\ell}$ in the $\mathbf{A}$-norm over the same Krylov subspace, $\delta\mathbf{x}^*$ being the exact solution [Reference 35, lemma 2.1.1]. The convergence properties of CG can then be analysed in terms of the error in the $\mathbf{A}$-norm [Reference 33, pp. 204–205]:

$$\frac{\|\mathbf{e}_{\ell}\|_{\mathbf{A}}}{\|\mathbf{e}_0\|_{\mathbf{A}}} \leq 2\left(\frac{\sqrt{\kappa(\mathbf{A})} - 1}{\sqrt{\kappa(\mathbf{A})} + 1}\right)^{\ell}, \quad \ell \in \mathbb{N}, \tag{4}$$

where $\kappa(\mathbf{A})$ is the condition number of $\mathbf{A}$, which is defined in the 2-norm as

$$\kappa(\mathbf{A}) = \frac{\lambda_{\max}(\mathbf{A})}{\lambda_{\min}(\mathbf{A})},$$

with $\lambda_{\max}(\mathbf{A})$ and $\lambda_{\min}(\mathbf{A})$ being the largest and smallest eigenvalues of $\mathbf{A}$, respectively. Equation (4) shows that convergence will tend to be fast when the condition number is close to 1. The condition number can thus be used as an indicator of the convergence rate of CG. Note that the initial error may also have an influence on the convergence behavior. For simplicity, however, we focus only on the effect of the condition number on the convergence rate.

Since Equation (4) depends on the condition number, it does not account for the distribution of the eigenvalues between the smallest and largest values. As a consequence, it can lead to a pessimistic error bound, especially when κ is very large. More advanced error bounds exist, which account for a more complex representation of the spectrum (e.g., see Chap. 13 in the book of Reference 36). However, for the problem considered in this article, those error bounds provide little improvement over the error bound given by Equation (4). They have been shown to be more accurate only in the limit of a very large number of iterations, while here we are interested mainly in the solution accuracy in the early iterations of CG. As they are more complex and less general than Equation (4), we did not apply them in this article.

In order to accelerate the convergence rate of CG, it is common to use a preconditioner. For data assimilation problems that solve quadratic problem (3), it is customary to use $\mathbf{B}$ as a preconditioner, as it usually yields a significantly smaller condition number compared to that of the unpreconditioned problem, and a more clustered spectrum of eigenvalues.[3,5,37] Therefore, we will focus on solving the $\mathbf{B}$-preconditioned linear system. Since $\mathbf{B}$ is SPD, it can be factored as

$$\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathsf{T}},$$

where $\mathbf{U} \in \mathbb{R}^{n\times n}$. We can then introduce $\mathbf{B}$-preconditioning symmetrically using a split-preconditioner,

$$\mathbf{U}^{\mathsf{T}}\mathbf{A}\ \mathbf{U}\ \delta\mathbf{v} = \mathbf{U}^{\mathsf{T}}\mathbf{b}, \tag{5}$$

where $\delta\mathbf{x} = \mathbf{U}\delta\mathbf{v}$. An unpreconditioned CG can be applied to Equation (5) by taking $\mathbf{U}^{\mathsf{T}}\mathbf{A}\mathbf{U}$ as the (SPD) system matrix and $\mathbf{U}^{\mathsf{T}}\mathbf{b}$ as the right-hand side.

In this article, we will evaluate the condition number of the preconditioned Hessian matrix,

$$\mathbf{S} = \mathbf{U}^{\mathsf{T}}\mathbf{A}\ \mathbf{U} = \mathbf{I}_n + \mathbf{U}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\mathbf{R}^{-1}\ \mathbf{H}\ \mathbf{U}, \tag{6}$$

and determine its sensitivity to parameters in the covariance matrices $\mathbf{B}$ and $\mathbf{R}$ when their spatial correlations are modeled by diffusion operators.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.3 | Weighting matrices formulated as the inverse of diffusion operators

The operators **B** and **R** describe the covariance structures of the background and observation errors. In the idealized 1D-Var system used in this study, the covariance matrices are small enough to be constructed explicitly using a functional expression to determine the matrix elements (e.g., as done in Reference 30). However, when one considers a realistic system, the size of **B** and **R** become too large to perform direct matrix-vector products. Hence, we prefer to adopt an approach that scales with the size of the problem and avoids the explicit construction of covariance matrices.

Covariance matrices can be factored as $\mathbf{\Sigma C \Sigma}$, where $\mathbf{\Sigma}$ is a diagonal matrix of standard deviations and **C** is an SPD correlation matrix. The computational difficulties are inherent in the specification and application of **C**. Reference 23 showed that multiplying an arbitrary vector by a Gaussian correlation matrix can approximately be achieved by numerically 'time'-stepping a diffusion equation with that arbitrary vector taken as the 'initial' condition[1]. This procedure defines an operator that models the product of a correlation matrix with the 'initial' condition without defining each element of the matrix. Since each 'time'-step involves the manipulation of sparse matrices, this strategy is naturally appropriate for large problems. The product of the diffusion coefficient μ and the total action 'time' $T = M\Delta t$ of the diffusion process, where M is the total number of diffusion steps and Δt is the 'time' step, controls the length-scale D_g of the Gaussian function that would be used to construct the correlation matrix, where $D_g^2 = 2\mu T$. Reference 24 describe the technique in detail and generalize it to account for anisotropic correlations. References 21 and 22 describe an extension of the technique that involves solving the diffusion equation using an implicit 'time'-stepping scheme instead of the explicit scheme of the original approach.

With the implicit scheme, the total number of diffusion steps M becomes a free parameter together with the diffusion coefficient multiplied by the 'time' step ($\mu\Delta t$ [With the explicit scheme, their product is the single free parameter controlling the length-scale D_g of the Gaussian function]). This extra degree of freedom allows the diffusion operator to represent matrix-vector products with correlation matrices from the Matérn family,[20] where M is linked to the standard smoothness parameter ν of the underlying Matérn correlation functions in $\mathbb{R}^d$ via the relation $\nu = M - d/2$. In $\mathbb{R}$, these functions are characterized by a polynomial times the exponential function and are also known as Mth-order auto-regressive (AR) functions. The parameter $\mu\Delta t$ is precisely the square of the standard length-scale parameter L of the Matérn functions. Without loss of generality, Δt can be set to 1, so that $\mu = L^2$. The connection between the Matérn correlation functions and the differential operator describing the inverse of an implicit diffusion process has its roots in the seminal work of Reference 38. We outline the connection in Appendix A and provide the key theoretical relations for the 1D-Var problem under consideration in this study.

While L is the length-scale parameter appearing explicitly in the definition of the diffusion operator, it is common in data assimilation to use an alternative length-scale parameter, the Daley length-scale D, to control the spatial smoothing properties of the diffusion kernel. The Daley length-scale can be understood as the half-width of the parabola osculating the correlation function at its origin.[39,40] It is defined for at least twice differentiable correlations functions, which in our case corresponds to AR functions with $M > 1$. As discussed in Appendix A.1, on $\mathbb{R}$, D, and L are related through the equation

$$D = L\sqrt{2M - 3}, \tag{7}$$

where the square-root term generalizes to $\sqrt{2M - d - 2}$ in $\mathbb{R}^d$.[22] An advantage of the parameter pair (D, M) over (L, M) is that it allows better control of the spatial range and spectral properties of the AR functions (see Figure A1 in Appendix A.1). An important property is that AR functions converge to a Gaussian function as M tends to infinity with L simultaneously reduced to zero to keep D constant. A closely-related length-scale parameter,

$$\rho = wL\sqrt{2M - 1}, \tag{8}$$

where w is a constant, is used in geostatistical and machine-learning applications.[41-43]2 AR functions defined in terms of ρ also have the property that they converge to a Gaussian function as M tends to infinity with ρ fixed. In $\mathbb{R}^d$, the square-root term in Equation (8) generalizes to $\sqrt{2M - d}$. An advantage of ρ over D is that it is valid for $M = 1$ as well as $M > 1$, while an advantage of D over ρ (and L) is that it is easier to estimate in practical applications when D is spatially dependent [Reference 26, Sec. 2.4]. In Section 3, the analytical results are first derived in terms of L and then interpreted in terms of both D and ρ, where we will refer to ρ with $w = 1$ as the *Stein* length-scale for convenience. In Section 4, the numerical experiments are discussed mainly in terms of D.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

On a circular domain of radius a (see Appendix A.2), we can define the discrete, symmetric diffusion-modeled correlation operator $\mathbf{C}$ as a sequence of linear operators represented by their respective matrices [Reference 25, Sec. 3.1]:

$$\mathbf{C} = \mathbf{\Gamma}\,\mathbf{L}\,\mathbf{W}^{-1}\,\mathbf{\Gamma}, \tag{9}$$

where $\mathbf{L} = \mathbf{T}^{-M}$ is a self-adjoint diffusion operator, $\mathbf{T}$ being a discrete representation of the shifted Laplacian operator $\mathcal{T}$. On the circular domain with constant L, we have from Equation (A5) that $\mathcal{T} \equiv I - L^2\partial^2/a^2\partial\phi^2$ where $-\pi \le \phi \le \pi$. The matrix $\mathbf{W}$ contains geometry- and grid-dependent weights. It defines the weighting matrix of the discrete form of the $L^2(\mathbb{S})$-inner product with respect to which $\mathbf{L}$ is self-adjoint; that is, $\mathbf{L} = \mathbf{W}^{-1}\mathbf{L}^{\mathsf{T}}\mathbf{W}$. Sections 3.2 and 3.3 of Reference 18 provide a comprehensive discussion of this point within the context of a finite element method discretization of the diffusion equation. The diagonal matrix $\mathbf{\Gamma}$ contains normalization factors so that the diagonal elements of $\mathbf{C}$ are approximately equal to one. On the circular domain with constant $L \ll a$, we can set $\mathbf{\Gamma} = \gamma\mathbf{I}$ where γ is well approximated by the constant product $\sqrt{\nu L}$ where ν is a monotonically increasing function of M given by Equation (A3). For example, the error in γ^2 is smaller than 0.001% when $L/a = 0.3$. Reference 26 provide an overview of methods for estimating $\mathbf{\Gamma}$ on other domains and when the correlation parameters are not constant.

Taking M to be an even number allows us to split $\mathbf{T}^{-M} = \mathbf{T}^{-M/2}\,\mathbf{T}^{-M/2}$ and hence to derive a simple 'square-root' factorization of Equation (9). The 'square-root' operator is convenient for generating random correlated samples and has been used for this purpose for the numerical experiments in Section 4. Another convenient property that comes specifically from the implicit formulation is that it provides immediate access to an inexpensive formulation of the inverse of the operator. This can be noticed from the inverse of Equation (9),

$$\mathbf{C}^{-1} = \mathbf{\Gamma}^{-1}\,\mathbf{W}\,\mathbf{L}^{-1}\,\mathbf{\Gamma}^{-1},$$

where $\mathbf{L}^{-1} = \mathbf{T}^{M}$ simply involves M applications of the shifted Laplacian operator.

Here, we consider a finite-difference discretization of the diffusion equation under the assumption that the grid resolution is uniform so that $\mathbf{W} = h\mathbf{I}$, where h is the grid size. Furthermore, we assume that $\mathbf{\Sigma} = \sigma\mathbf{I}$ where σ is a constant standard deviation. Given these assumptions together with the assumption that L is constant, we can simplify the expressions for the diffusion-modeled covariance operators for $\mathbf{B}$ and $\mathbf{R}$ as

$$\mathbf{B} = \frac{\sigma_{\mathrm{b}}^2\nu_{\mathrm{b}}L_{\mathrm{b}}}{h_{\mathrm{b}}}\mathbf{T}_{\mathrm{b}}^{-M_{\mathrm{b}}} = \frac{\sigma_{\mathrm{b}}^2\nu_{\mathrm{b}}L_{\mathrm{b}}}{h_{\mathrm{b}}}\left(\mathbf{I}_n - L_{\mathrm{b}}^2\mathbf{\Delta}_{h_{\mathrm{b}}}\right)^{-M_{\mathrm{b}}}, \tag{10}$$

$$\mathbf{R} = \frac{\sigma_{\mathrm{o}}^2\nu_{\mathrm{o}}L_{\mathrm{o}}}{h_{\mathrm{o}}}\mathbf{T}_{\mathrm{o}}^{-M_{\mathrm{o}}} = \frac{\sigma_{\mathrm{o}}^2\nu_{\mathrm{o}}L_{\mathrm{o}}}{h_{\mathrm{o}}}\left(\mathbf{I}_m - L_{\mathrm{o}}^2\mathbf{\Delta}_{h_{\mathrm{o}}}\right)^{-M_{\mathrm{o}}}, \tag{11}$$

where the subscripts 'b' and 'o' refer to quantities relative to the background and observations, respectively. The symbol $\mathbf{\Delta}_h$ denotes the finite-difference representation of the Laplacian operator, which depends on the grid resolution for the background and observations. The matrices $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ and $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ are identity matrices. We are also interested in the expression for $\mathbf{R}^{-1}$, which follows immediately from Equation (11):

$$\mathbf{R}^{-1} = \frac{h_{\mathrm{o}}}{\sigma_{\mathrm{o}}^2\nu_{\mathrm{o}}L_{\mathrm{o}}}\mathbf{T}_{\mathrm{o}}^{M_{\mathrm{o}}} = \frac{h_{\mathrm{o}}}{\sigma_{\mathrm{o}}^2\nu_{\mathrm{o}}L_{\mathrm{o}}}\left(\mathbf{I}_m - L_{\mathrm{o}}^2\mathbf{\Delta}_{h_{\mathrm{o}}}\right)^{M_{\mathrm{o}}}.$$

By taking $\mathbf{W}_{\mathrm{o}} = h_{\mathrm{o}}\mathbf{I}_m$, we are assuming that the observations are regularly distributed with a separation distance of h_{o}. This is done for mathematical convenience. Relative to the domain size $2\pi a$, h_{o} is an explicit parameter that reflects observation density and can be compared to $h_{\mathrm{b}}/2\pi a$, the density of background points. With the simplifying assumptions above, we are able to establish explicit theoretical bounds on the condition number of the preconditioned Hessian matrix as detailed in Section 3. Note that these assumptions, while necessary for establishing theoretical results, are not met in realistic atmospheric or ocean applications. In particular, most observation types have unstructured spatial distributions. Nevertheless, we can expect the results derived in Section 3, which focus on the sensitivity of the condition number of the Hessian matrix to correlation model parameters, to be relevant even with more complicated observation networks.

We remark now on the actual values of the parameter pairs $(M_{\mathrm{b}}, M_{\mathrm{o}})$, $(L_{\mathrm{b}}, L_{\mathrm{o}})$, and $(\sigma_{\mathrm{b}}^2, \sigma_{\mathrm{o}}^2)$ that will be considered in this study. First, values of $(M_{\mathrm{b}}, M_{\mathrm{o}}) \ge 10$ lead to AR functions that are practically Gaussian, so we will not consider values

RIGHTSLINK
10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

beyond 10. Values of (L_b, L_o) should be large enough compared to the grid size (h_b, h_o) (at least $L_b/h_b \geq 1$ and $L_o/h_o \geq 1$) in order to avoid large discretization errors in the finite-difference representation of the diffusion operator. Ideally, the parameters should be chosen to provide the optimal fit to our available estimate of the error covariances (with (L_b, L_o) and (σ_b^2, σ_o^2) made spatially dependent in general). Background-error correlations are often specified as quasi-Gaussian functions (large values of (M_b, M_o)). The reason for this choice can be mainly computational; that is, efficient models, like diffusion, exist for applying quasi-Gaussian functions.[24,44,45] Another reason is that quasi-Gaussian functions are sufficiently regular that they can be differentiated, which is important for defining cross-variable (multivariate) covariances in atmospheric and ocean data assimilation.[39,46,47] In comparison, estimates of the spatial correlations of observation error often display a sharp decrease at short range and slow decay at longer range,[7,16,48] which with an AR function is best modeled with a small value of (M_b, M_o). In view of these remarks, the case where $M_o < M_b$ seems to be of particular interest. Nevertheless, both this case and the case where $M_o \geq M_b$ will be considered as different data-sets may give rise to different error characteristics.

3 | CONDITIONING OF THE PRECONDITIONED LINEAR SYSTEM

In this section, we are interested in analysing the convergence of CG applied to the linear system (5) where the system matrix $\mathbf{S}$ depends on the diffusion-modeled covariance matrices described in the previous section. In particular, we are interested in analysing the sensitivity of the convergence in terms of the parameters of these covariance matrices. For this purpose, we will focus on the condition number of $\mathbf{S}$, denoted $\kappa(\mathbf{S})$.

We start by recalling some results from References 28 and 30 on the upper bound of $\kappa(\mathbf{S})$ for general covariance matrices.

Theorem 1 (Theorem 3 of Reference 30). *Let $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{m\times m}$ be symmetric, positive-definite matrices. Let $\mathbf{U} = \mathbf{U}^T \in \mathbb{R}^{n\times n}$ be the (unique) symmetric square root of $\mathbf{B} = \mathbf{U}\mathbf{U}^T = \mathbf{U}^2$ and let $\mathbf{V} = \mathbf{V}^T \in \mathbb{R}^{m\times m}$ be the (unique) symmetric square root of $\mathbf{R} = \mathbf{V}\mathbf{V}^T = \mathbf{V}^2$. If $\mathbf{H} \in \mathbb{R}^{m\times n}$, with $m < n$, and $\mathbf{S} = \mathbf{I}_n + \mathbf{U}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{U}$, then*

$$\kappa(\mathbf{S}) \leq 1 + \|\mathbf{V}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{V}^{-1}\|_\infty. \tag{12}$$

The upper bound given in Equation (12) is quite general and it is not straightforward to understand the impact of each component of the matrix $\mathbf{S}$ on the condition number. Moreover, caution is needed when applying Equation (12) in physical applications as the eigen-decomposition of a covariance matrix will not be independent of the physical units of the variables [Sec. 4.3.4; Reference 49]. For this reason, it is generally more meaningful to consider an eigen-decomposition on the associated (non-dimensional) correlation matrix. In our idealized study, there is a single "physical" variable, with arbitrary units, and direct observations of that variable. Furthermore, both the background- and observation-error variances are taken to be constant, so can be factored out of $\mathbf{B}$ and $\mathbf{R}$. In this case, the eigenvalues of $\mathbf{B}$ and $\mathbf{R}$ will be identical to those of their respective correlation matrices up to a multiplicative factor given by their respective variances. We can thus continue to consider the eigen-decompositions of $\mathbf{B}$ and $\mathbf{R}$ without ambiguity.

Other bounds have been proposed by References 28 and 30 to understand the impact of each covariance matrix.

Theorem 2 (Corollary 1 of Reference 30). *Let $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{m\times m}$ be symmetric, positive-definite matrices. Let $\mathbf{U} = \mathbf{U}^T \in \mathbb{R}^{n\times n}$ be the (unique) symmetric square root of $\mathbf{B} = \mathbf{U}\mathbf{U}^T = \mathbf{U}^2$. If $\mathbf{H} \in \mathbb{R}^{m\times n}$, with $m < n$, and $\mathbf{S} = \mathbf{I}_n + \mathbf{U}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{U}$, then*

$$\kappa(\mathbf{S}) \leq 1 + \frac{\lambda_{\max}(\mathbf{B})}{\lambda_{\min}(\mathbf{R})}\lambda_{\max}(\mathbf{H}\mathbf{H}^T). \tag{13}$$

The upper bound in Equation (13) attempts to separate the influence of $\mathbf{B}$ and $\mathbf{R}$ by considering their eigenvalues separately. While this separation makes the bound easier to use for sensitivity analyses, it can also degrade the accuracy of the bound. Indeed, compared to the bound given by Equation (12), the bound given by Equation (13) does not account for any interaction between $\mathbf{B}$ and $\mathbf{R}^{-1}$, which results in a more pessimistic bound.

As an example, we consider the case where both $\mathbf{B}$ and $\mathbf{R}$ are modeled using diffusion operators as described in Section 2.3, and $\mathbf{H}$ is a selection matrix. We consider a domain of length 2000 km, composed of $n = 500$ points that are equally spaced every $h_b = 4$ km. We assume that direct observations are available at every other grid point ($m = 250$ and

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

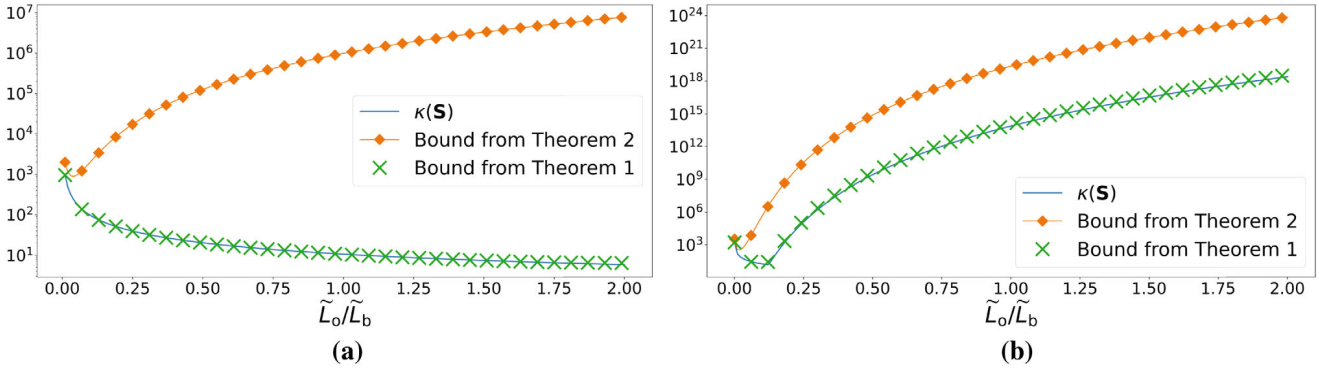


FIGURE 1 Upper bounds on the condition number of **S** compared to the exact condition number as a function of $\widetilde{L}_o/\widetilde{L}_b = L_o h_b/L_b h_o$. (a) The correlation function of the background error is Gaussian-like ($M_b = 8$), while the correlation function of the observation error is a SOAR function ($M_o = 2$). (b) The correlation function of the background error is a SOAR function ($M_b = 2$), while the correlation function of the observation error is Gaussian-like ($M_o = 8$).

$h_o = 8$ km). We define $\widetilde{L}_b = L_b/h_b$ and $\widetilde{L}_o = L_o/h_o$ where $L_b = 60$ km is fixed and L_o is allowed to vary. Figure 1 compares the two upper bounds from Theorems 1 and 2 with the exact condition number[3] for different parameter specifications in **B** and **R**. Figure 1a shows the results for $M_b = 8$ and $M_o = 2$, while Figure 1b shows the results for $M_b = 2$ and $M_o = 8$, for values of $\widetilde{L}_o/\widetilde{L}_b$ ranging from 0.01 to 2. While the bound of Theorem 1 matches closely the condition number for both settings, the bound of Theorem 2 is far less accurate. This discrepancy shows that even though the diffusion models use different parameters, the structural similarities between **B** and **R** lead to a crucial interaction between them. Therefore, separating the effect of these matrices in the bound results in a pessimistic upper bound.

In this article, we are interested in obtaining explicit and accurate theoretical bounds on the condition number in terms of the key parameters of the diffusion-modeled correlation operators presented in Section 3.1. We restrict **H** to a class of uniform selection operators, which are associated with uniformly distributed observations. In Section 3.2, we study how these selection operators interact with diffusion operators. Based on the results of these two sections, we examine the conditioning of **S** in Section 3.3. Additional results are derived in Section 3.3.1 using a simplified matrix $\mathbf{S}_o$, which is equal to **S** when **H** is the identity matrix but is an approximation otherwise.

3.1 | Spectral properties of diffusion operators

The diffusion-modeled covariance operators for **B** and **R** in Equations (10) and (11) are defined in terms of Laplacian matrices $\mathbf{\Delta}_{h_b} \in \mathbb{R}^{n\times n}$ and $\mathbf{\Delta}_{h_o} \in \mathbb{R}^{m\times m}$, formed from a centred finite-difference discretization of the Laplacian operator on a uniform grid of resolution h_b and h_o, respectively. On a periodic domain, $\mathbf{\Delta}_{h_b}$ and $\mathbf{\Delta}_{h_o}$ are circulant matrices[4] that are *tridiagonal* except for the first and last lines where additional non-zero elements appear in the corners due to the periodic boundary conditions. Likewise, the shifted Laplacian matrices $\mathbf{T}_b \in \mathbb{R}^{n\times n}$ and $\mathbf{T}_o \in \mathbb{R}^{m\times m}$ are circulant, near-tridiagonal matrices. Specifically, for **R**, we have

$$\mathbf{\Delta}_{h_o} = \frac{1}{h_o^2}\begin{pmatrix} -2 & 1 & 0 & & 0 & 1 \\ 1 & -2 & 1 & 0 & & 0 \\ 0 & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & 0 \\ 0 & & 0 & 1 & -2 & 1 \\ 1 & 0 & & 0 & 1 & -2 \end{pmatrix} \tag{14}$$

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

RIGHTSLINK

and thus

$$\mathbf{T}_{\mathrm{o}} = \begin{pmatrix} 1+2\widetilde{L}_{\mathrm{o}}^2 & -\widetilde{L}_{\mathrm{o}}^2 & 0 & & 0 & -\widetilde{L}_{\mathrm{o}}^2 \\ -\widetilde{L}_{\mathrm{o}}^2 & 1+2\widetilde{L}_{\mathrm{o}}^2 & -\widetilde{L}_{\mathrm{o}}^2 & 0 & & 0 \\ 0 & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & 0 \\ 0 & & 0 & -\widetilde{L}_{\mathrm{o}}^2 & 1+2\widetilde{L}_{\mathrm{o}}^2 & -\widetilde{L}_{\mathrm{o}}^2 \\ -\widetilde{L}_{\mathrm{o}}^2 & 0 & & 0 & -\widetilde{L}_{\mathrm{o}}^2 & 1+2\widetilde{L}_{\mathrm{o}}^2 \end{pmatrix} \tag{15}$$

where $\widetilde{L}_{\mathrm{o}} = L_{\mathrm{o}}/h_{\mathrm{o}}$ is a non-dimensional parameter that roughly corresponds to the number of grid points over which observation-error correlations are significant. The expressions for $\mathbf{\Delta}_{h_{\mathrm{b}}}$ and $\mathbf{T}_{\mathrm{b}}$ are the same as Equations (14) and (15) with $(h_{\mathrm{b}}, \widetilde{L}_{\mathrm{b}})$ instead of $(h_{\mathrm{o}}, \widetilde{L}_{\mathrm{o}})$.

An important property of circulant matrices is that, for a given size, they all share the same eigenvectors, which form a Fourier basis.[50] Consequently, their eigenvalues can be found by taking the discrete Fourier transform of one of the rows. Let $[\mathbf{f}^{(i)}]_p$ be the p-th component of the i-th eigenvector of $\mathbf{T}_{\mathrm{o}}$ and let λ_i be the corresponding eigenvalue:

$$\left[\mathbf{f}^{(i)}\right]_p = \frac{1}{\sqrt{m}} e^{-2\pi j \frac{ip}{m}}, \qquad \lambda_i = 1 + 4\widetilde{L}^2 \sin^2\left(\pi \frac{i}{m}\right), \; i, p \in [\![0, m-1]\!],$$

where j denotes the imaginary unit ($j^2 = -1$). Since $\mathbf{T}_{\mathrm{o}}$ is symmetric, its eigenvalues are real and each of them is repeated twice; that is, $\lambda_i(\mathbf{T}_{\mathrm{o}}) = \lambda_{m-i}(\mathbf{T}_{\mathrm{o}})$, except $\lambda_0(\mathbf{T}_{\mathrm{o}})$ as the index i stops at $m-1$. If m is even, $\lambda_{m/2}(\mathbf{T}_{\mathrm{o}})$ is also unique.

The covariance matrices $\mathbf{B}$ and $\mathbf{R}$ in Equations (10) and (11) are proportional to a power of the inverse of $\mathbf{T}_{\mathrm{b}}$ and $\mathbf{T}_{\mathrm{o}}$, respectively. They are also circulant matrices and diagonal in a Fourier basis described by the vectors $\mathbf{f}^{(i)}$. Two circulant covariance matrices with different parameter specifications then share the same eigenvectors as long as they are applied on the same domain. In addition, since $\mathbf{T}_{\mathrm{o}}$ and $\mathbf{T}_{\mathrm{b}}$ are symmetric, the diffusion matrices are symmetric and their eigenvalues are real and proportional to powers of the inverse of the eigenvalues of $\mathbf{T}_{\mathrm{o}}$ and $\mathbf{T}_{\mathrm{b}}$. The eigenvalues of $\mathbf{B}$ and $\mathbf{R}$ are then

$$\lambda_i(\mathbf{B}) = \frac{\sigma_{\mathrm{b}}^2 \nu_{\mathrm{b}} L_{\mathrm{b}}}{h_{\mathrm{b}}} \lambda_i(\mathbf{T}_{\mathrm{b}})^{-M_{\mathrm{b}}} = \sigma_{\mathrm{b}}^2 \nu_{\mathrm{b}} \widetilde{L}_{\mathrm{b}} \left[1 + 4\widetilde{L}_{\mathrm{b}}^2 \sin^2\left(\pi \frac{i}{n}\right)\right]^{-M_{\mathrm{b}}}, \quad i \in [\![0, n-1]\!], \tag{16}$$

$$\lambda_i(\mathbf{R}) = \frac{\sigma_{\mathrm{o}}^2 \nu_{\mathrm{o}} L_{\mathrm{o}}}{h_{\mathrm{b}}} \lambda_i(\mathbf{T}_{\mathrm{o}})^{-M_{\mathrm{o}}} = \sigma_{\mathrm{o}}^2 \nu_{\mathrm{o}} \widetilde{L}_{\mathrm{o}} \left[1 + 4\widetilde{L}_{\mathrm{o}}^2 \sin^2\left(\pi \frac{i}{m}\right)\right]^{-M_{\mathrm{o}}}, \quad i \in [\![0, m-1]\!]. \tag{17}$$

3.2 | Influence of H as a uniform selection operator

In the previous section, we derived the eigenvalue spectra of the covariance matrices when they are defined as circulant matrices. Our simplifying assumption that the covariance and grid parameters are constant is crucial to ensure the circulant property of the covariance matrices. In this section, we exploit these results to analyse the spectrum of the matrix $\mathbf{HBH}^{\mathsf{T}}$ that appears in $\mathbf{S}$. We have already assumed that observations are evenly distributed with a grid spacing h_{o}. We add here another assumption such that $\zeta = h_{\mathrm{o}}/h_{\mathrm{b}}$ is an integer, meaning that there are observations every ζ grid points. In this case, we refer to the resulting observation operator $\mathbf{H}$ as a *uniform selection* operator. The total number of observations is then given by $m = n/\zeta$ assuming that ζ is a divisor of n. While this distribution of observations is highly simplified, it allows us to analyse explicitly the sensitivity of the condition number to the correlation model parameters since it preserves to some extent the structure of $\mathbf{B}$, as described in the following lemma.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lemma 1. *Let $\mathbf{B} \in \mathbb{R}^{n\times n}$ be a symmetric circulant matrix and let $\mathbf{H} \in \mathbb{R}^{n\times m}$ be a uniform selection operator where $\zeta m = n$ with ζ a positive integer. The matrix $\mathbf{HBH}^\mathsf{T} \in \mathbb{R}^{m\times m}$ is then a symmetric circulant matrix with eigenvalues*

$$\lambda_i\left(\mathbf{HBH}^\mathsf{T}\right) = \frac{1}{\zeta}\sum_{r=0}^{\zeta-1}\lambda_{i+rm}(\mathbf{B}), \tag{18}$$

for $i \in [\![0, m-1]\!]$.

Proof. Let the notation $[\mathbf{A}]_{p,q}$ denote the element on the p-th row and q-th column of any matrix $\mathbf{A}$. As $\mathbf{B}$ is a circulant matrix, it can be diagonalized in a Fourier basis:

$$\mathbf{B} = \mathbf{F}_n\mathbf{\Lambda}_\mathrm{b}\mathbf{F}_n^\mathrm{H}, \tag{19}$$

where $\mathbf{\Lambda}_\mathrm{b}$ is a diagonal matrix, and the elements of $\mathbf{F}_n$ are the (normalized) n-th roots of unity:

$$\forall p, q \in [\![0, n-1]\!], \quad [\mathbf{F}_n]_{p,q} = \frac{1}{\sqrt{n}}\omega_n^{pq} \quad \text{with} \quad \omega_n = e^{\frac{2\pi j}{n}}.$$

The superscript "H" stands for conjugate (Hermitian) transpose and $\mathbf{F}_n^\mathrm{H} = \mathbf{F}_n^{-1}$. Starting from Equation (19), we have

$$\mathbf{HBH}^\mathsf{T} = \mathbf{HF}_n\mathbf{\Lambda}_\mathrm{b}(\mathbf{HF}_n)^H. \tag{20}$$

The matrix $\mathbf{HF}_n$ is of dimension $m \times n$ and is composed of the rows of $\mathbf{F}_n$:

$$\forall p \in [\![0, m-1]\!], q \in [\![0, n-1]\!] \quad [\mathbf{HF}_n]_{p,q} = [\mathbf{F}_n]_{\zeta p,q} = \frac{1}{\sqrt{n}}\omega_n^{\zeta pq}.$$

The matrix $\mathbf{HF}_n$ can be linked to $\mathbf{F}_m$, the matrix that diagonalizes circulant matrices of dimension $m \times m$. The elements of the latter are the m-th root of unity:

$$\forall p, q \in [\![0, m-1]\!], \quad [\mathbf{F}_m]_{p,q} = \frac{1}{\sqrt{m}}\omega_m^{pq} \quad \text{with} \quad \omega_m = e^{\frac{2j\pi}{m}},$$

which, as $n = \zeta m$, can be linked to the n-th root of unity as $\omega_m = \omega_n^\zeta$. Consequently, for the first m columns of $\mathbf{HF}_n$, we have

$$\forall p \in [\![0, m-1]\!], q \in [\![0, m-1]\!] \quad [\mathbf{HF}_n]_{p,q} = \frac{1}{\sqrt{n}}\omega_m^{pq} = \frac{1}{\sqrt{\zeta}}[\mathbf{F}_m]_{p,q}.$$

The other $n - m$ columns of $\mathbf{HF}_n$ can be characterized by using the periodicity of the m-th root of unity: $\omega_m^{pq} = \omega_m^{p(q+rm)}$ for any positive integer r. Therefore, $\mathbf{HF}_n$ is a matrix concatenated with ζ copies of $\mathbf{F}_m$:

$$\mathbf{HF}_n = \frac{1}{\sqrt{\zeta}}[\mathbf{F}_m \cdots \mathbf{F}_m].$$

Equation (20) can thus be rewritten as

$$\mathbf{HBH}^\mathsf{T} = \mathbf{F}_m\frac{1}{\sqrt{\zeta}}\left(\mathbf{I}_m \cdots \mathbf{I}_m\right)\mathbf{\Lambda}_\mathrm{b}\begin{pmatrix}\mathbf{I}_m\\ \vdots \\ \mathbf{I}_m\end{pmatrix}\frac{1}{\sqrt{\zeta}}\mathbf{F}_m^\mathrm{H} = \mathbf{F}_m\mathbf{\Lambda}'_\mathrm{b}\mathbf{F}_m^\mathrm{H},$$

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

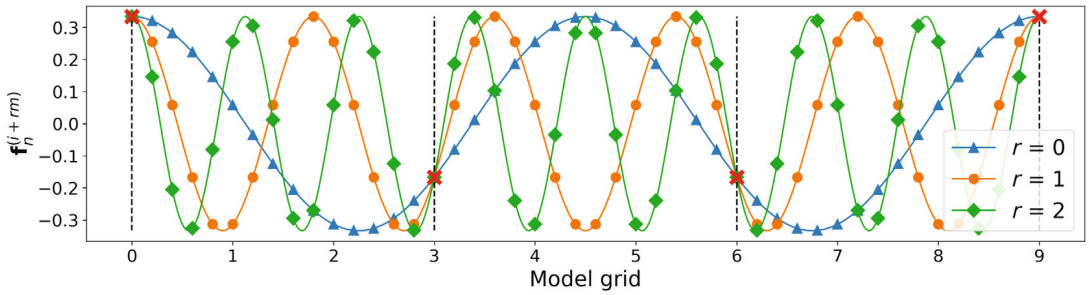


FIGURE 2 Eigenvectors $\mathbf{f}_n^{(i+rm)}$ of a circulant matrix of size $n \times n$ with $n = 9$, for $i = 2$, $r \in [\![0, \zeta - 1]\!]$ with $\zeta = 3$ and $m = n/\zeta = 3$. On the model grid with n points, these eigenvectors can be distinguished. On the observation grid with m points (marked by dashed lines), they all take the same values, which are the elements of the corresponding eigenvector $\mathbf{f}_m^{(i)}$ of a circulant matrix of size $m \times m$. Note that the first and last points in the figure represent the same grid point due to the periodic boundary conditions. The vectors have been plotted with a higher resolution than the one of the model grid for the sake of clarity.

where $\mathbf{\Lambda}'_{\mathrm{b}}$ is a diagonal matrix of dimension $m \times m$. As $\mathbf{HBH}^{\mathsf{T}}$ is diagonal for the basis defined by the columns of $\mathbf{F}_m$, it is a circulant matrix. Its eigenvalues are the non-zero elements of the diagonal matrix $\mathbf{\Lambda}'_{\mathrm{b}}$, which are given by Equation (18). ∎

A matrix-vector product with $\mathbf{HBH}^{\mathsf{T}}$ is therefore in the range of the column vectors $\mathbf{f}_m^{(i)}$, weighted by the *average of ζ evenly-distributed eigenvalues* of $\mathbf{B}$. This result can be linked to the notion of *aliasing*. Different vectors from the Fourier basis of dimension n (e.g., column $\mathbf{f}_n^{(i)}$) become indistinguishable and equal to the same frequency mode in the Fourier basis of dimension m (e.g., column $\mathbf{f}_m^{(i)}$) once they are sub-sampled, as the highest frequencies cannot be resolved by the observation grid. This point is illustrated in Figure 2, which shows multiple distinct columns of $\mathbf{F}_n$ that all take the same values on the observation grid, which are the values of a column of $\mathbf{F}_m$. As shown in Equation (18), the weight associated with a frequency mode $\mathbf{f}_m^{(i)}$ in $\mathbf{HBH}^{\mathsf{T}}$ is the average of the weights associated with the frequency modes $\mathbf{f}_n^{(i)}$, $\mathbf{f}_n^{(i+m)}$, … , $\mathbf{f}_n^{(i+(\zeta-1)m)}$ in $\mathbf{B}$, which become equal to $\mathbf{f}_m^{(i)}$ once sub-sampled by $\mathbf{H}$.

3.3 | Spectral properties of the B-preconditioned Hessian matrix

Lemma 1 implies that $\mathbf{HBH}^{\mathsf{T}}$ shares the same eigenvectors as any circulant matrix of size $m \times m$, including the matrices associated with the constant-parameter diffusion operator applied on the uniform observation grid. This allows us to derive a new expression for the spectrum of $\mathbf{S}$.

Lemma 2. *Let $\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathsf{T}} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{m\times m}$, with $m < n$, be circulant matrices and let $\mathbf{H} \in \mathbb{R}^{n\times m}$ be a uniform selection operator. If $\mathbf{S} = \mathbf{I}_n + \mathbf{U}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{U}$ then the i-th eigenvalue of $\mathbf{S}$ is*

$$\lambda_i(\mathbf{S}) = \begin{cases} 1 + \frac{\lambda_i(\mathbf{HBH}^{\mathsf{T}})}{\lambda_i(\mathbf{R})} & \text{if } i \in [\![0, m-1]\!], \\ 1 & \text{otherwise.} \end{cases}$$

Proof. Let $\Lambda^*(\cdot)$ denote the spectrum of a matrix without its zero eigenvalues. For any matrix $\mathbf{P}$ and $\mathbf{Q}$ of respective sizes $n \times m$ and $m \times n$, with $n > m$, we know that [Theorem 21.10.1, Reference 51]

$$\Lambda^*(\mathbf{PQ}) = \Lambda^*(\mathbf{QP}). \tag{21}$$

With $\mathbf{P} = \mathbf{U}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}$ and $\mathbf{Q} = \mathbf{R}^{-1}\mathbf{H}\mathbf{U}$, Equation (21) implies that $\mathbf{U}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{U}$ has at least $n - m$ eigenvalues equal to zero, and that

$$\Lambda^*\left(\mathbf{U}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{U}\right) = \Lambda^*\left(\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}\right).$$

Therefore, $\mathbf{S}$ has an eigenvalue of 1 with multiplicity of $n - m$ and the remaining m eigenvalues are the elements of $1 + \Lambda^*\left(\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}\right)$. Since $\mathbf{HBH}^{\mathsf{T}}$ and $\mathbf{R}^{-1}$ are circulant matrices (see Lemma 1), they are

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

diagonalizable in the same basis. Therefore, the eigenvalues of their product is the product of their respective eigenvalues:

$$\forall i \in [\![0, m-1]\!], \quad \lambda_i\left(\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}\right) = \lambda_i\left(\mathbf{R}^{-1}\right)\lambda_i\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}\right) = \frac{\lambda_i\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}\right)}{\lambda_i\left(\mathbf{R}\right)}.$$

■

We can now write the eigenvalues of $\mathbf{S}$ in terms of the constant parameters of the *diffusion-modeled covariance matrices* by using the results of Lemmas 1 and 2, and the expressions for the eigenvalues of $\mathbf{B}$ and $\mathbf{R}$ given by Equations (16) and (17), respectively.

Theorem 3. *Let* $\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathsf{T}} \in \mathbb{R}^{n\times n}$ *and* $\mathbf{R} \in \mathbb{R}^{m\times m}$ *be circulant matrices defined by Equations* (10) *and* (11) *where the shifted Laplacian matrices* $\mathbf{T}_{\mathrm{b}}$ *and* $\mathbf{T}_{\mathrm{o}}$ *are defined in Section* 3.1. *Let* $\mathbf{H} \in \mathbb{R}^{n\times m}$ *be a uniform selection operator where* $\zeta m = n$ *with* ζ *a positive integer. If* $\mathbf{S} = \mathbf{I}_n + \mathbf{U}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{U}$ *then*

$$\lambda_i(\mathbf{S}) = \begin{cases} 1 + \alpha \displaystyle\sum_{r=0}^{\zeta-1} \frac{\left[1 + 4\widetilde{L}_{\mathrm{o}}^2 \sin^2\left(\pi \frac{i}{m}\right)\right]^{M_{\mathrm{o}}}}{\left[1 + 4\widetilde{L}_{\mathrm{b}}^2 \sin^2\left(\pi \frac{i+rm}{\zeta m}\right)\right]^{M_{\mathrm{b}}}} & \text{if } i \in [\![0, m-1]\!], \\ 1 & \text{otherwise,} \end{cases} \tag{22}$$

with $\widetilde{L}_{\mathrm{o}} = L_{\mathrm{o}}/h_{\mathrm{o}}$, $\widetilde{L}_{\mathrm{b}} = L_{\mathrm{b}}/h_{\mathrm{b}}$ *and*

$$\alpha = \frac{\sigma_{\mathrm{b}}^2 \nu_{\mathrm{b}} L_{\mathrm{b}}}{\sigma_{\mathrm{o}}^2 \nu_{\mathrm{o}} L_{\mathrm{o}}}.$$

Proof. Lemma 2 provides an expression for the eigenvalues of $\mathbf{S}$ in terms of the eigenvalues of $\mathbf{R}$ and $\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}$. The eigenvalues of $\mathbf{R}$ are known from Equation (17), and the eigenvalues of $\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}$ are obtained by applying the result of Lemma 1 to the eigenvalues of $\mathbf{B}$ from Equation (16):

$$\lambda_i\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}\right) = \frac{\sigma_{\mathrm{b}}^2 \nu_{\mathrm{b}} L_{\mathrm{b}}}{h_{\mathrm{o}}} \sum_{r=0}^{\zeta-1} \left[1 + 4\widetilde{L}_{\mathrm{b}}^2 \sin^2\left(\pi \frac{i+rm}{\zeta m}\right)\right]^{-M_{\mathrm{b}}}, \quad i \in [\![0, m-1]\!],$$

as $h_{\mathrm{o}} = \zeta h_{\mathrm{b}}$ and $n = \zeta m$. ■

From Theorem 3, it is clear that the minimum eigenvalue of $\mathbf{S}$, $\lambda_{\min}(\mathbf{S})$, is equal to one when $m < n$ (fewer observations than background variables), and is bounded below by 1 when $m = n$. The condition number of $\mathbf{S}$ is thus bounded above by the maximum eigenvalue of $\mathbf{S}$, $\lambda_{\max}(\mathbf{S})$. There is no simple analytical expression for $\lambda_{\max}(\mathbf{S})$ that can be deduced from Theorem 3. However, we can already notice that $\lambda_{\max}(\mathbf{S})$ increases with increasing ratio between the background- and observation-error variances, $\sigma_{\mathrm{b}}^2/\sigma_{\mathrm{o}}^2$. This basic dependency of the condition number on the relative variances is well known.[28,30,52]

3.3.1 | Spectral properties of the simplified $\mathbf{B}$-preconditioned Hessian matrix

Analysing the sensitivity of $\lambda_{\max}(\mathbf{S})$ with respect to the diffusion parameters is not straightforward from the expression given in Theorem 3. We can obtain a better understanding by approximating the effect of the matrix $\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}$ with a diffusion operator discretized directly on the observation grid. Specifically, let $\mathbf{B}_{\mathrm{o}} \in \mathbb{R}^{m\times m}$ be a diffusion operator with the same covariance parameters as $\mathbf{B} \in \mathbb{R}^{n\times n}$ but discretized on the observation grid:

$$\mathbf{B}_{\mathrm{o}} = \frac{\sigma_{\mathrm{b}}^2 \nu_{\mathrm{b}} L_{\mathrm{b}}}{h_{\mathrm{o}}} \left(\mathbf{I}_m - L_{\mathrm{b}}^2 \boldsymbol{\Delta}_{h_{\mathrm{o}}}\right)^{-M_{\mathrm{b}}}.$$

As $\mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}$ and $\mathbf{B}_{\mathrm{o}}$ are two spatial discretizations of the same continuous diffusion operator, the difference between the two is solely due to the error associated with the spatial discretization. If there are direct observations at each grid point

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

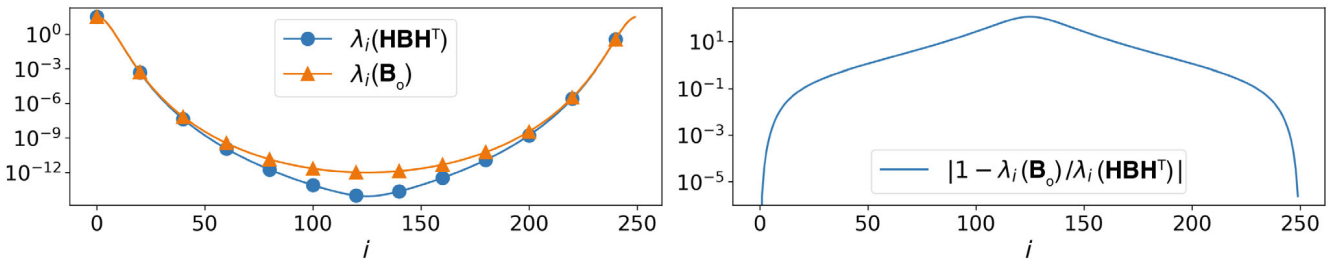


FIGURE 3 Eigenvalues of $\mathbf{HBH}^\mathsf{T}$ and $\mathbf{B}_o$ (left panel) and their relative difference (right panel), computed for a domain of 2000 km, with a model grid of $n = 500$ points and an observation at every other model grid point ($m = 250$). The correlation model represented by $\mathbf{B}$ is an AR function of order $M_b = 8$ with a Daley length-scale of $D_b = L_b\sqrt{2M_b - 3} = 100$ km.

($\mathbf{H} = \mathbf{I}_n$), both $\mathbf{B}_o$ and $\mathbf{HBH}^\mathsf{T}$ are equal to $\mathbf{B}$ and there is no approximation. If there are less observations than grid points, $\mathbf{B}_o$ and $\mathbf{HBH}^\mathsf{T}$ still share the same eigenvectors as they are both circulant. However, they have slightly different eigenvalues due to the different spatial discretizations. As $\mathbf{B}_o$ is a diffusion operator, its eigenvalues can be deduced from the results of Section 3.1:

$$\lambda_i(\mathbf{B}_o) = \sigma_b^2 \nu_b \widetilde{L}_{b/o}\left[1 + 4\widetilde{L}_{b/o}^2 \sin^2\left(\pi \frac{i}{m}\right)\right]^{-M_b}, \tag{23}$$

where $\widetilde{L}_{b/o} = L_b/h_o$. The eigenvalues of $\mathbf{B}_o$ tend to overestimate the eigenvalues of $\mathbf{HBH}^\mathsf{T}$, with maximum relative error occurring for the smallest eigenvalues, as illustrated in Figure 3 for the case where $h_o/h_b = 2$.

Let us recall that $\mathbf{S}$ has an eigenvalue of 1 with multiplicity of $n - m$ and that the remaining m eigenvalues are the elements of $1 + \Lambda^*\left(\mathbf{R}^{-1}\mathbf{HBH}^\mathsf{T}\right)$ (see Lemma 2). Approximating the matrix $\mathbf{HBH}^\mathsf{T}$ by $\mathbf{B}_o$, we are now interested in determining the eigenvalues of the matrix

$$\mathbf{S}_o = \mathbf{I}_m + \mathbf{R}^{-1}\mathbf{B}_o.$$

Theorem 4. *Let $\mathbf{B}_o \in \mathbb{R}^{m\times m}$ be the circulant matrix defined by Equation (10) with $h_b = h_o$, and let $\mathbf{R} \in \mathbb{R}^{m\times m}$ be the circulant matrix defined by Equation (11). The shifted Laplacian matrices $\mathbf{T}_b$ and $\mathbf{T}_o$ are defined in Section 3.1. If $\mathbf{S}_o = \mathbf{I}_m + \mathbf{R}^{-1}\mathbf{B}_o$ then*

$$\forall i \in [\![0, m-1]\!], \quad \lambda_i(\mathbf{S}_o) = 1 + \alpha \frac{\left[1 + 4\widetilde{L}_o^2 \sin^2\left(\pi \frac{i}{m}\right)\right]^{M_o}}{\left[1 + 4\widetilde{L}_{b/o}^2 \sin^2\left(\pi \frac{i}{m}\right)\right]^{M_b}},$$

with

$$\alpha = \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o} = \frac{\sigma_b^2 \nu_b L_b}{\sigma_o^2 \nu_o L_o}. \tag{24}$$

Proof. The eigenvalues of $\mathbf{S}_o$ are given by $\lambda_i(\mathbf{S}_o) = 1 + \lambda_i\left(\mathbf{R}^{-1}\mathbf{B}_o\right)$. As $\mathbf{B}_o$ and $\mathbf{R}^{-1}$ are both circulant matrices, they share the same eigenvectors. Therefore, we have

$$\lambda_i(\mathbf{S}_o) = 1 + \frac{\lambda_i(\mathbf{B}_o)}{\lambda_i(\mathbf{R})}.$$

Replacing the eigenvalues of $\mathbf{B}_o$ and $\mathbf{R}$ by their expressions provided by Equations (23) and (17), respectively, yields the expression for the eigenvalues of $\mathbf{S}_o$. ∎

The next theorem provides a bound on the condition number of $\mathbf{S}_o$ by using the expression for the eigenvalues of $\mathbf{S}_o$.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

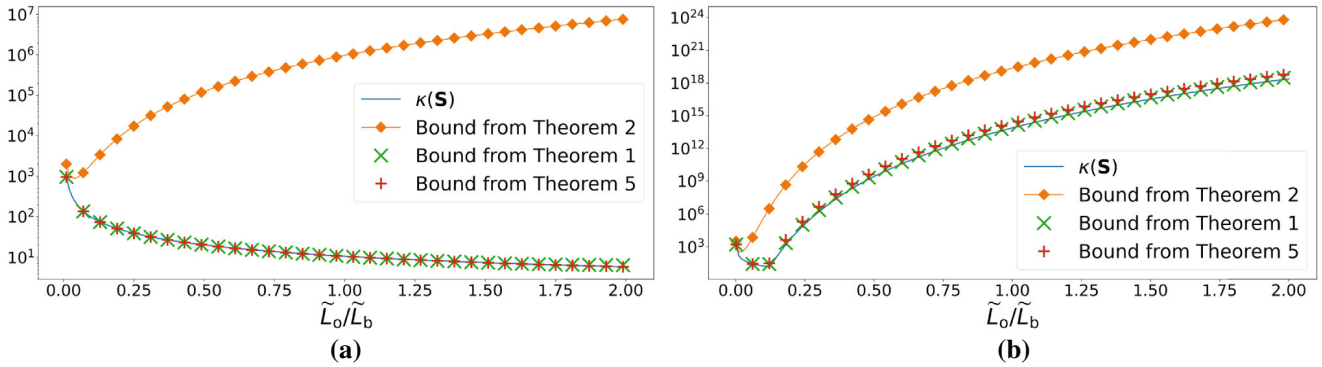


FIGURE 4 As in Figure 1 but with an additional curve for the new bound from Theorem 5. (a) $M_o = 2, M_b = 8$. (b) $M_o = 8, M_b = 2$.

Theorem 5. *Let* $\mathbf{S}_o$ *be defined as in Theorem 4, and let* α *be given by Equation (24). Then,* $\kappa(\mathbf{S}_o) \leq \eta$ *where*

$$\eta = \begin{cases} 1 + \alpha\left(\frac{\widetilde{L}_o^2}{M_b}\right)^{M_b}\left(\frac{M_o}{\widetilde{L}_{b/o}^2}\right)^{M_o}\left(\frac{M_b - M_o}{\widetilde{L}_o^2 - \widetilde{L}_{b/o}^2}\right)^{M_b - M_o} & \text{if (i) } \widetilde{L}_o^2 M_o > \widetilde{L}_{b/o}^2 M_b; \text{(ii) } M_o < M_b; \text{ and} \\ & \text{(iii) } \widetilde{L}_{b/o}^2 M_b - \widetilde{L}_o^2 M_o > 4\widetilde{L}_{b/o}^2 \widetilde{L}_o^2 (M_o - M_b) \\ 1 + \alpha \max\left\{\frac{(1+4\widetilde{L}_o^2)^{M_o}}{(1+4\widetilde{L}_{b/o}^2)^{M_b}}, 1\right\} & \text{otherwise.} \end{cases}$$

Proof. See Appendix B. ∎

As explained further in this section, Theorem 5 describes the sensitivity of the condition number of $\mathbf{S}_o$ to the diffusion parameters *while keeping the bound sharp*. The sharpness of the bound is illustrated in Figure 4, where it is compared with the exact condition number of $\mathbf{S}$ and the bounds given in Theorems 1 and 2. The exact condition number has been evaluated using the extreme eigenvalues taken from the full spectrum of exact eigenvalues provided by the expression in Theorem 3. These results show that taking into account the specific structure of the covariance matrices improves the bound relative to the one given in Theorem 2.

We can further simplify the result in Theorem 5 by considering η as a function of $\widetilde{L}_o$ only. Corollaries 1 and 2 below characterize the variations of η with respect to $\widetilde{L}_o$ when $M_o \geq M_b$ and $M_o < M_b$, respectively.

Corollary 1. *Consider that* η *defined in Theorem 5 is a function of* $\widetilde{L}_o$; $\eta = f(\widetilde{L}_o)$. *Assume that* $\widetilde{L}_o$ *has a lower bound such that* $\widetilde{L}_o\sqrt{2M_o - 1} > 1/2$ *and that condition (ii) from Theorem 5 is not met; that is,* $M_o \geq M_b$. *Then, the minimum of the function* f *is unique and reached when*

$$\left(1 + 4\widetilde{L}_o^2\right)^{M_o} = \left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}. \tag{25}$$

Proof. See Appendix C. ∎

Corollary 2. *Consider that* η *defined in Theorem 5 is a function of* $\widetilde{L}_o$; $\eta = f(\widetilde{L}_o)$. *Assume that condition (ii) from Theorem 5 is met* ($M_o < M_b$). *Assume further that condition (iii) holds when condition (i) is satisfied. Then, the minimum of the function* f *is unique and reached when*

$$L_o\sqrt{2M_o - 1} = L_b\sqrt{2M_b - 1} \tag{26}$$

Proof. See Appendix D. ∎

Although the assumptions in Corollaries 1 and 2 appear restrictive, they exclude cases that are of limited practical interest. In particular, values of $\widetilde{L}_o$ that are smaller than the observation grid resolution h_o correspond to observation errors that are effectively uncorrelated. To avoid this case, we focus on values of $\widetilde{L}_o \geq 1$, which automatically fulfils the

RIGHTSLINK
10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

condition on the lower bound in Corollary 1. Alternatively, the condition on the lower bound can be seen as restricting the Stein length-scale $\rho_o = L_o\sqrt{2M_o - 1}$ (see Equation 8 with $w = 1$) to be greater than half the grid resolution. In Corollary 2, we assume that condition (iii) of Theorem 5 holds if condition (i) is satisfied. To simplify condition (iii), we can impose a practical bound on the value of $\widetilde{L}_{b/o}$. For instance, we are not interested in cases where the length-scale L_b is smaller than h_o, that is, $\widetilde{L}_{b/o} \leq 1$. More generally, we can assume that $\widetilde{L}_{b/o}$ is bounded below by a positive scalar $\widetilde{L}_{min}$, which leads to the next corollary.

Corollary 3. *Assume that there exists a positive scalar* $\widetilde{L}_{min}$ *such that*

$$\widetilde{L}_{b/o} \geq \widetilde{L}_{min},$$

and that condition (ii) of Theorem 5 holds. Then, condition (iii) of Theorem 5 is met if

$$M_b \leq 2\left(1 + 4\widetilde{L}^2_{min}\right). \tag{27}$$

Proof. Let us define $r_M = M_o/M_b$. Using the assumption on the length-scale, condition (iii) of Theorem 5 can be rewritten as

$$\widetilde{L}^2_{min} + 4\widetilde{L}^2_{min}\widetilde{L}^2_o(1 - r_M) - \widetilde{L}^2_o r_M \geq 0.$$

As M_o and M_b are assumed to be even integers and as condition (ii) is met ($M_o < M_b$), we know that $r_M \leq (M_b - 2)/M_b$. Using this relation, we obtain that

$$M_b\widetilde{L}^2_{min} + 8\widetilde{L}^2_{min}\widetilde{L}^2_o - \widetilde{L}^2_o(M_b - 2) \geq 0,$$

which can be rearranged to give

$$M_b\left(1 - \widetilde{L}^2_{min}/\widetilde{L}^2_o\right) \leq 2\left(1 + 4\widetilde{L}^2_{min}\right).$$

Since $\widetilde{L}^2_{min}/\widetilde{L}^2_o$ is positive, we obtain the inequality (27). ∎

Taking $\widetilde{L}_{min} = 1$ in Equation (27) results in $M_b \leq 10$. Increasing M_b beyond 10 has little practical value as the correlation function is already approximately Gaussian with this value.

For the case where **B** and **R** are modeled with SOAR functions ($M_b = M_o = 2$),[30] point out that, for fixed L_b, the minimum of their upper bound for the condition number of the **B**-preconditioned Hessian matrix is found by setting $L_o = L_b$. Corollaries 1 and 2 confirm this result and extend it to other AR functions ($M_o = M_b > 2$). They also cover cases where the order of the AR functions differs between **B** and **R** ($M_b \neq M_o$), in which case the function defining the upper bound on the condition number, $\eta = f(\widetilde{L}_o)$, does not reach its minimum value when $L_o = L_b$.

If $M_o > M_b$ then $\widetilde{L}_o$ can be much smaller than $\widetilde{L}_{b/o}$ to attain the minimum of the function $f(\widetilde{L}_o)$. For example, if $M_b = 2$, $M_o = 10$, and $\widetilde{L}_o = 1.5$, then $\widetilde{L}_{b/o}$ needs to be 158 to satisfy Equation (25) of Corollary 1. The correlation functions with fixed values of $\left(1 + 4\widetilde{L}^2\right)^M$ have very different range for low values of M as illustrated in Figure 5a. On the other hand, if $M_o < M_b$ then Corollary 2 states that the minimum value is attained when the Stein length-scales $\rho_o = L_o\sqrt{2M_o - 1}$ and $\rho_b = L_b\sqrt{2M_b - 1}$ are equal. Note that, unlike condition (25), condition (26) is independent of h_o. As shown in Figure 5b, the correlation functions with fixed values of ρ are very similar for different values of M.

For an alternative interpretation of Corollaries 1 and 2, we can recast Equations (25) and (26) in terms of the Daley length-scales D_o and D_b (Equation 7), which is the length-scale parameter we will use to interpret the numerical experiments in the following sections. Assuming $M_b > 1$ and $M_o > 1$, we have

$$L_b = \frac{D_b}{\sqrt{2M_b - 3}} \quad \text{and} \quad L_o = \frac{D_o}{\sqrt{2M_o - 3}}. \tag{28}$$

If $M_o = M_b$ then the minimum is reached when $D_o = D_b$.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

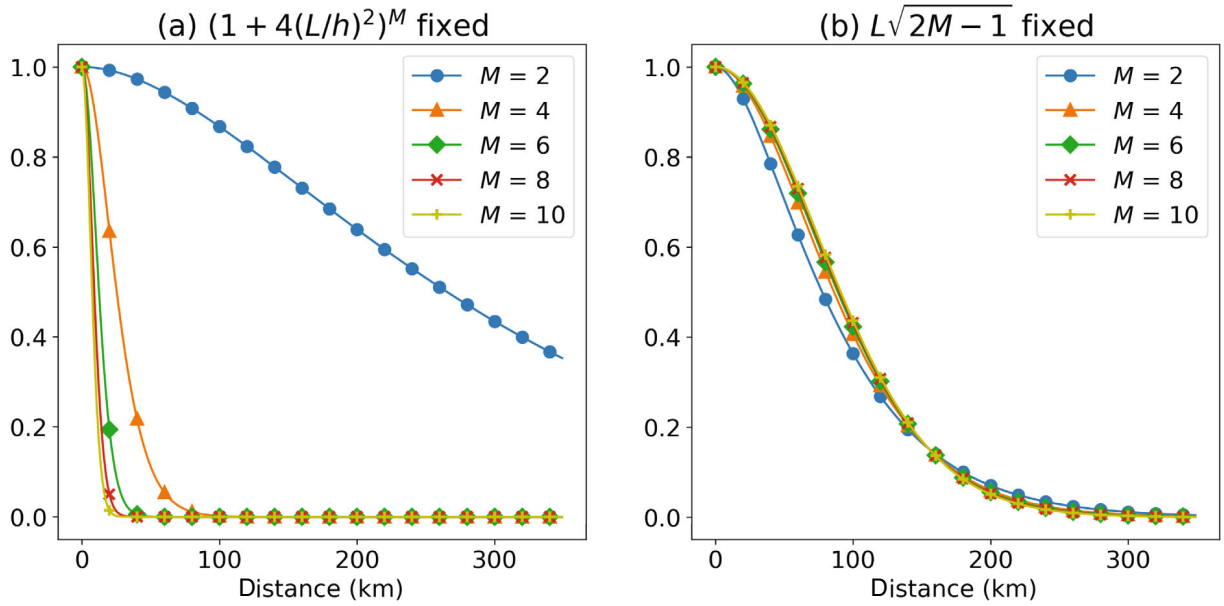


FIGURE 5 AR correlation functions (Equation A2) displayed for different values of M. For each M, the value of $\widetilde{L}$ is chosen such that the quantities (a) $\left(1+4\widetilde{L}^2\right)^M$ and (b) $L\sqrt{2M-1}$ from Corollaries 1 and 2, respectively, are kept constant. The corresponding values of L, the Stein length-scale $\rho = L\sqrt{2M-1}$, and the Daley length-scale $D = L\sqrt{2M-3}$ can be found in Table 1.

TABLE 1 Values of the length-scale parameter $L = \widetilde{L}h$ where $h = 1$ km, the Stein length-scale $\rho = L\sqrt{2M-1}$, and the Daley length-scale $D = L\sqrt{2M-3}$ associated with the curves in Figure 5.

	$(1+4\widetilde{L}^2)^M$ fixed			$L\sqrt{2M-1}$ fixed		
M	L (km)	ρ (km)	D (km)	L (km)	ρ (km)	D (km)
2	158.1	273.8	158.1	46.2	80.0	46.2
4	8.9	23.5	19.9	30.2	80.0	67.6
6	3.4	11.2	10.1	24.1	80.0	72.4
8	2.0	7.7	7.4	20.7	80.0	74.5
10	1.5	6.5	6.2	18.3	80.0	75.7

Note: The fixed values of $\left(1+4\widetilde{L}^2\right)^M$ and $L\sqrt{2M-1}$ are 10^{10} and 80 km, respectively.

If $M_o > M_b$ then Equation (25) translates as

$$\left(1+\frac{4D_o^2}{h_o^2(2M_o-3)}\right)^{M_o} = \left(1+\frac{4D_b^2}{h_o^2(2M_b-3)}\right)^{M_b}. \tag{29}$$

The location of the minima is very sensitive to M_o and M_b since they appear as exponents in Equation (29). While a small change of M_o from 8 to 10 would have limited effect on the correlation function, it can have a drastic effect on the quantities in Equation (29). In turn, this can significantly affect the condition number (as seen from Theorem 5) as well as the criteria in Corollary 1. This property can be detrimental if ignored, but can also be exploited to our advantage to improve the conditioning without significantly altering the correlation shape, as will be illustrated in Section 4.

If $M_o < M_b$ then Equation (26) translates as

$$D_o^2\left(\frac{2M_o-1}{2M_o-3}\right) = D_b^2\left(\frac{2M_b-1}{2M_b-3}\right), \tag{30}$$

RIGHTSLINK
10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from which we can deduce that the minimum is reached when $D_o < D_b$ (cf. $L_o > L_b$ and $\rho_o = \rho_b$). This is evident from the last column of Table 1, which shows D increasing with increasing M. The ratio between D_b and D_o reaches at most 1.6 for the limiting values of $M_o = 2$ and $M_b = 10$.

Equations (25) and (26) (respectively, Equations 29 and 30) provide simple criteria that can be used to adjust the value of L_o (respectively, D_o) to minimize the condition number of the **B**-preconditioned Hessian matrix. From this perspective, we can use Corollaries 1 and 2 as the basis of a method for *reconditioning* observation-error covariance matrices that account for spatial correlations with parametric functions from the Matérn family. This would be complementary to existing methods for reconditioning sample covariance matrices, for example, to represent inter-channel error correlations in satellite observations.[53,54] For more complex problems, where the assumptions of these corollaries are not perfectly satisfied, we can still use criteria (25) and (26) (or 29 and 30) as a guideline for adjusting covariance parameters in **B** and **R** to improve the conditioning of the **B**-preconditioned Hessian matrix.

3.3.2 | Condition number estimates with correlated and uncorrelated observation errors

In this section, we compare the condition number of **S** for different values of the correlation parameter pairs (M_o, D_o) and (M_b, D_b). The condition number $\kappa(\mathbf{S})$ is computed using the (exact) analytical expression of the eigenvalues of **S** given in Theorem 3. In addition, we compute the exact 'optimal' parameter pairs (i.e., those that minimize the condition number) and compare them with those predicted by the optimality criteria in Corollaries 1 and 2. As this theorem applies to the matrix $\mathbf{S}_o$, and not **S**, these optimality criteria are only exact when there is a direct observation at each grid point.

In presenting the results, we choose to normalize $\kappa(\mathbf{S})$ by $\kappa(\mathbf{S}_u)$ where $\mathbf{S}_u$ is given by Equation (6) with $\mathbf{R} = \sigma_o^2 \mathbf{I}_m$; that is, with observation-error correlations neglected. An analytical expression for the eigenvalues of $\mathbf{S}_u$ can be derived directly from Equation (22) of Theorem 3 by setting $M_o = 0$ (no diffusion) and $\nu_o \widetilde{L}_o = \gamma^2 / h_o = 1$ (exact normalization):

$$\forall i \in [\![0, n-1]\!], \quad \lambda_i(\mathbf{S}_u) = \begin{cases} 1 + \alpha_u \sum_{r=0}^{\zeta-1} \left[1 + 4\widetilde{L}_b^2 \sin^2\left(\pi \frac{i+rm}{\zeta m} \right) \right]^{-M_b} & \text{if } i \in [\![0, m-1]\!], \\ 1 & \text{otherwise,} \end{cases}$$

where

$$\alpha_u = \frac{\sigma_b^2 \nu_b L_b}{\sigma_o^2 h_o}.$$

As we are considering the case where there are fewer observations than grid points ($n > m$), the minimum eigenvalue of $\mathbf{S}_u$ is one. The condition number of $\mathbf{S}_u$ is thus computed as

$$\kappa(\mathbf{S}_u) = 1 + \alpha_u \max_{i \in [\![0, m-1]\!]} \sum_{r=0}^{\zeta-1} \left[1 + 4\widetilde{L}_b^2 \sin^2\left(\pi \frac{i+rm}{\zeta m} \right) \right]^{-M_b}. \tag{31}$$

Note that the maximum of the sum in Equation (31) is larger than one (as it is larger than one at least for $i = 0$) and approximately equal to one for parameter values of interest; that is, for $M_b \geq 2$ and $\widetilde{L}_b \geq 1$, its maximum is less than 1.04. The condition number of $\mathbf{S}_u$ is thus dominated by α_u.

We denote χ the ratio of condition numbers:

$$\chi = \frac{\kappa(\mathbf{S})}{\kappa(\mathbf{S}_u)}. \tag{32}$$

As M_o and D_o have no effect on $\kappa(\mathbf{S}_u)$, variations of χ with respect to these parameters will reflect variations of $\kappa(\mathbf{S})$. If $\chi < 1$ then accounting for correlated observation error will improve the conditioning of **S** and thus we can expect the convergence rate of CG to be improved. Conversely, if $\chi > 1$ then accounting for correlated observation error will degrade the conditioning of **S** and we can expect the convergence rate of CG to be degraded.

In the following, we will compute the condition numbers as a function of the Daley length-scales defined in Equation (28). Furthermore, since we are mainly interested in the sensitivity of the condition number to the correlation model parameters, we will assume that the background- and observation-error variances are equal ($\sigma_b^2 / \sigma_o^2 = 1$). We

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

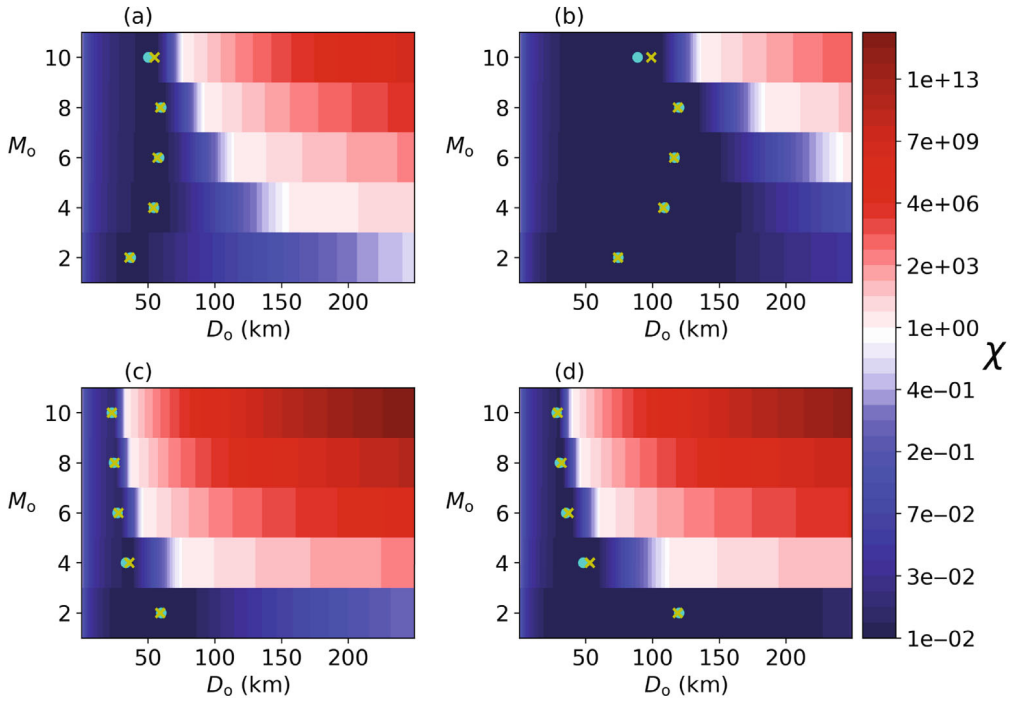


FIGURE 6 The ratio χ (Equation 32) is plotted for a fixed parameter pair (M_b, D_b) per panel (indicated in the title), and values of M_o and D_o that vary along the axes. The cyan circles mark the minima predicted by Corollaries 1 and 2. The yellow crosses mark the true minima. Note that the color palette uses a logarithmic scale with a different range below and above $\chi = 1$. (a) $D_b = 60$ km; $M_b = 8$. (b) $D_b = 120$ km; $M_b = 8$. (c) $D_b = 60$ km; $M_b = 2$. (d) $D_b = 120$ km; $M_b = 2$.

consider a domain of length 2000 km, composed of $n = 500$ points that are equally spaced every $h_b = 4$ km. We assume that a direct observation is available every other grid point ($\zeta = 2$, $m = 250$, $h_o = 8$ km).

Figure 6 shows χ as a function of D_o (abscissa) and M_o (ordinate) for different parameter pairs (M_b, D_b) indicated in the title of each panel. The zones in blue (respectively, red) correspond to parameter pairs (M_o, D_o) that improve (respectively, degrade) the condition number. When $M_o \leq M_b$ and $D_o \leq D_b$, the conditioning is systematically improved. An improvement is also possible when $D_o \geq D_b$ if M_o is small enough. However, when M_o becomes too large compared to M_b or when D_o becomes too large compared to D_b, the conditioning is degraded and can become significantly degraded even for modest changes in the parameter values. For example, when $M_b = 8$ and $D_b = 60$ km (Figure 6a), and $M_o = 10$, χ (and thus $\kappa(\mathbf{S})$) increases by several orders of magnitude when the value of D_o is increased to less than double D_b. When D_o is approximately four times D_b, χ reaches 10^{10} (top right corner of Figure 6a). In these cases, we can expect the convergence of CG to be significantly affected, as illustrated later in Section 4.

As predicted by Corollaries 1 and 2, if D_b, M_b, and M_o are fixed, then $\kappa(\mathbf{S})$ admits a unique minimum. When $M_o > M_b$, the minima predicted by Corollary 1 are visibly shifted towards lower values of D_o (cf. circles and crosses in Figure 6). This shift corresponds to an increase of the condition number of up to 5%. As the variations of the condition number studied here cover a range of more than 10 orders of magnitude, this increase is acceptable. When $M_o < M_b$, there is no significant difference in the position of the minima predicted by Corollary 2 and the exact minima; there is an increase of the condition number between the predicted minima and exact minima that is smaller than 0.1%. The pattern is similar with each fixed settings for (M_b, D_b) (i.e., each panel of Figure 6). If D_b increases (decreases) then the 'optimal' values of D_o are shifted to the right (left) towards larger (smaller) values of D_o (cf. Figure 6a–d).

4 | NUMERICAL EXPERIMENTS

In this section, we illustrate how different covariance parameter settings influence the performance of the CG minimization. We evaluate the convergence rate in relation to the condition number diagnostic χ presented in Section 3.3.2 (see Figure 6) and the results of Corollaries 1 and 2.

RIGHTSLINK
10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.1 | Experimental framework

As in Section 3.3.2, we define our baseline 1D-Var experiment as one in which the background- and observation-error variances are taken to be equal, with their actual values set to one unit ($\sigma_b^2 = \sigma_o^2 = 1$). The domain is periodic with length 2000 km and there are $n = 500$ grid points ($h_b = 4$ km). We define $\mathbf{H}$ as a selection operator where direct observations are assumed to be available at every other grid point ($m = 250$, $h_o = 8$ km).

We consider different 'scenarios' where observations with perfectly known error correlations are assimilated together with a background state that also has perfectly known error correlations. We start by defining a 'true state', $\mathbf{x}_t$, which is specified by an analytical function. As $\mathbf{H}$ is linear in our framework, the actual choice of the true state has no impact on the performance of the CG minimization as it is subtracts out from the innovation vector. The background state and observations are then generated by adding to the true state, unbiased random perturbations of covariance matrices $\mathbf{B}$ and $\mathbf{R}$, respectively. Specifically, let $\hat{\boldsymbol{\epsilon}}_b$ and $\hat{\boldsymbol{\epsilon}}_o$ be normally-distributed vectors with zero mean and covariance matrix equal to the identity matrix. We can generate many realizations of $\hat{\boldsymbol{\epsilon}}_b \sim N(\mathbf{0}, \mathbf{I}_n)$ and $\hat{\boldsymbol{\epsilon}}_o \sim N(\mathbf{0}, \mathbf{I}_m)$ using a random number generator. Then, using the factored covariance matrices $\mathbf{B} = \mathbf{U}\mathbf{U}^T$ and $\mathbf{R} = \mathbf{V}\mathbf{V}^T$, we define

$$\begin{aligned} \mathbf{x}_b &= \mathbf{x}_t + \boldsymbol{\epsilon}_b, \\ \mathbf{y}_o &= \mathbf{H}\mathbf{x}_t + \boldsymbol{\epsilon}_o, \end{aligned} \tag{33}$$

where $\boldsymbol{\epsilon}_b = \mathbf{U}\hat{\boldsymbol{\epsilon}}_b$ and $\boldsymbol{\epsilon}_o = \mathbf{V}\hat{\boldsymbol{\epsilon}}_o$. By construction, $\mathbb{E}[\boldsymbol{\epsilon}_b\boldsymbol{\epsilon}_b^T] = \mathbf{B}$ and $\mathbb{E}[\boldsymbol{\epsilon}_o\boldsymbol{\epsilon}_o^T] = \mathbf{R}$ where $\mathbb{E}[\ ]$ is the expectation operator.

To assess the convergence rate of the CG algorithm at each iteration, it is common to monitor the reduction of the cost function or, equivalently, the reduction of the $\mathbf{A}$-norm of the analysis (solution) error. However, if we want to compare the convergence rate of CG with different $\mathbf{R}$, the $\mathbf{A}$-norm is not appropriate since it depends on $\mathbf{R}$ and thus does not represent the same quantity in all cases. Since we are working with an idealized system for which the true state $\mathbf{x}_t$ is known, we have access to alternative metrics that would not be available in a realistic system.

At the ℓ-th iteration of the CG algorithm, an increment $\delta\mathbf{x}_\ell$ is produced. We can deduce from this increment the analysis error that would result if the CG algorithm was stopped at the ℓ-th iteration:

$$\boldsymbol{\epsilon}_a^{(\ell)} = \mathbf{x}_b + \delta\mathbf{x}_\ell - \mathbf{x}_t.$$

In each experiment, there is a random component in the generation of the background and observations, which will affect $\boldsymbol{\epsilon}_a^{(\ell)}$. By performing multiple experiments with different right-hand sides ($\mathbf{b}$ in Equation 5), we can obtain multiple realizations of $\boldsymbol{\epsilon}_a^{(\ell)}$ from which analysis-error statistics can be deduced. In particular, we can estimate at each iteration the total analysis-error variance or, equivalently, the trace of the analysis-error covariance matrix. This is the quantity that is minimized explicitly in a statistical analysis based on the best linear unbiased estimator (BLUE). It is well known that, when the constraints are linear and when the background and observation errors are normally distributed, the minimizing solution of the cost function of variational data assimilation is equivalent to the BLUE when both are formulated under the same assumptions.[55]

As a diagnostic, we compute the square root of the average of the total analysis-error variance:

$$\sigma_a^{(\ell)} = \sqrt{\frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left((\boldsymbol{\epsilon}_a^{(\ell)})(\boldsymbol{\epsilon}_a^{(\ell)})^T\right)\right]} = \sqrt{\frac{1}{n}\mathbb{E}\left[(\boldsymbol{\epsilon}_a^{(\ell)})^T(\boldsymbol{\epsilon}_a^{(\ell)})\right]}, \tag{34}$$

where $\mathbb{E}$ denotes the expectation operator and Tr denotes the trace operator. We approximate the expectation operator as an average of 1000 realizations with random right-hand sides.

This metric can be used to assess not only the convergence rate of the minimization on which we focused in the previous sections, but also the accuracy of the solution at each iteration. We expect the solution of the minimization at full convergence to be more accurate when the actual observation-error correlations are accounted for. If the condition number is reduced by a non-diagonal $\mathbf{R}$ (i.e., $\chi < 1$ as in the 'blue zone' of Figure 6) then the minimization should converge faster. In this situation, we can expect the solution to be more accurate no matter when the minimization is stopped. On the other hand, if the condition number is increased by a non-diagonal $\mathbf{R}$ (i.e., $\chi > 1$ as in the 'red zone' of Figure 6) then we can expect the convergence rate to be slower. In this situation, it is not clear whether a non-diagonal $\mathbf{R}$ is beneficial to the analysis or not, as there is a trade-off between the convergence rate and the expected accuracy at full convergence. Monitoring the analysis error at each iteration allows us to visualize this trade-off as it indicates, at each iteration, how accurate the analysis would be (on average) if the convergence was stopped at this point.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

A natural choice of normalization for $\sigma_a^{(\ell)}$ is its initial value $\sigma_a^{(0)}$, which is equal to σ_b in the experiments as the initial increment $\delta\mathbf{x}_0$ is zero. The quantity $\sigma_a^{(\ell)}/\sigma_a^{(0)}$ thus indicates the relative error reduction on each iteration of CG. We denote σ_a^* the value of $\sigma_a^{(\ell)}$ at full convergence of CG. If the specifications of $\mathbf{B}$ and $\mathbf{R}$ used to compute the analysis match the actual error statistics, this quantity should become equal to its theoretical minimum, σ_a^{opt}, which can be computed directly from the trace of the theoretical analysis-error covariance matrix:

$$\sigma_a^{\text{opt}} = \sqrt{\frac{1}{n}\text{Tr}\left[\left(\mathbf{B}^{-1} + \mathbf{H}^{\text{T}}\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\right]}, \tag{35}$$

where $\mathbf{B}$ and $\mathbf{R}$ are the same as those used to generate the random errors in Equation (33).

In the experiments, $\mathbf{R}$ denotes the 'true' observation-error covariance matrix used to generate the spatially-correlated random errors that are added to the observations. The matrix $\widetilde{\mathbf{R}}_1 = \sigma_o^2\mathbf{I}_m$ is a diagonal approximation where σ_o^2 is the same constant variance used in the 'true' $\mathbf{R}$. This corresponds to the common case where spatially-correlated observation errors are ignored in the weighting matrix in the cost function, which is inconsistent with the statistical properties of the observations that are assimilated. The third scenario also uses a diagonal matrix, $\widetilde{\mathbf{R}}_2 = \upsilon\ \sigma_o^2\mathbf{I}_m$, but the variances are multiplied by an inflation factor (υ) to mitigate the effect of neglecting the error correlations. This procedure is common in real-data assimilation problems, to avoid overfitting observations at large spatial scales while retaining a simple covariance matrix. In practice, the inflation factor is usually estimated empirically. In our experiments, we can determine the best-possible inflation factor by minimizing σ_a^* with respect to υ. As σ_a^* behaves approximately as a convex function of υ, this can be achieved by computing σ_a^* for increasing values of υ until it stops decreasing (i.e., until the observations are no longer overfit). This method cannot be applied in an operational context as it requires access to the true state. Even with a performance metric that uses a proxy for the true state, the cost of the procedure would be prohibitive as thousands of realizations of σ_a^* are required. The experiments using $\widetilde{\mathbf{R}}_2$ thus represent the best inflation can offer rather than what could be achieved in practice.

4.2 | Results

In the first set of experiments, we consider the case where the background- and observation-error correlation parameters are in the regime $M_o < M_b$ and $D_o < D_b$. The observation-error correlation parameters are set to $M_o = 2$ and $D_o = 30$ km, which corresponds to a SOAR function as used in Reference 30. These values are roughly similar to those proposed by Reference 18, where the parameter settings were determined to provide a suitable fit of a diffusion-model to error correlation estimates of certain satellite radiance observations.[7] The background-error correlation parameters are set to $M_b = 8$ and $D_b = 60$ km, which makes the correlation function more Gaussian-like than that of $\mathbf{R}$. The correlation length-scale of $\mathbf{B}$ is double the correlation length-scale of $\mathbf{R}$. These are the same $\mathbf{B}$ parameters that were used in Figure 6. With these parameters, we know that $\chi < 1$ (Figure 7a), which means that the condition number is reduced when observation-error correlations are accounted for.

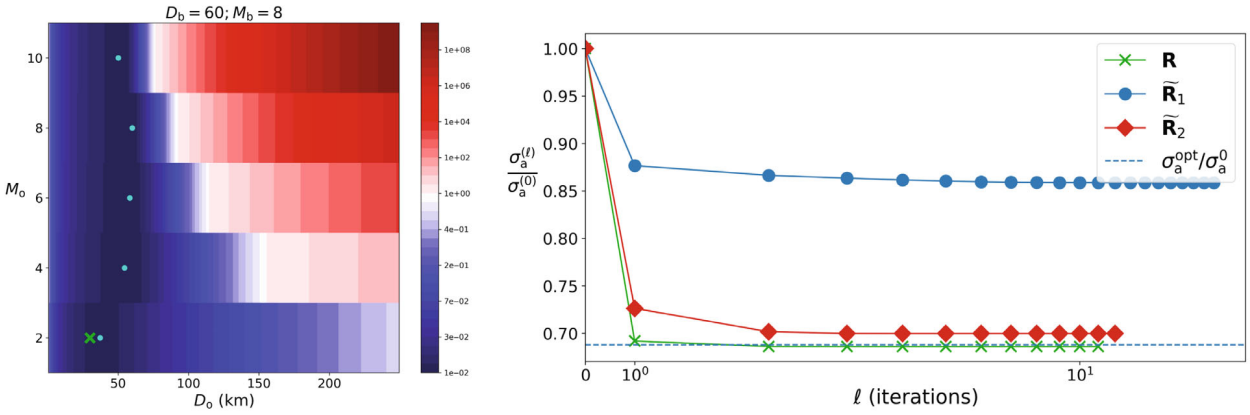
The average error-convergence curves from the 1D-Var experiments with $\mathbf{R}$, $\widetilde{\mathbf{R}}_1$, and $\widetilde{\mathbf{R}}_2$ are shown in Figure 7b. For each experiment, minimizations are performed in parallel for all 1000 realizations of the random right-hand side and are stopped when the 2-norm of the residual normalized by its initial value reaches 10^{-6}. Convergence is achieved rather quickly, taking about 20 iterations with $\widetilde{\mathbf{R}}_1$ and about 10 iterations with $\mathbf{R}$ and $\widetilde{\mathbf{R}}_2$.

For the experiment with $\widetilde{\mathbf{R}}_1$, the analysis-error standard deviation is only reduced by about 15% at full convergence, compared to the theoretical limit of 32%. In this set-up, the optimal variance inflation factor is approximately equal to 10.5. Inflating the error variances significantly improves the error reduction (30%). However, the theoretical minimum error cannot be reached, even with an inflation factor that has been optimized for this specific experiment. It is also important to remark that the experiment with inflated variances converges faster than the experiment with the original variances. This is consistent with Equation (31), which shows that, for large α_u, the condition number of $\mathbf{S}_u$ is approximately inversely proportional to the observation-error variance and is thus divided by 10.5 in this case. Best results are obtained with the true $\mathbf{R}$. First, the convergence rate is the fastest of the three experiments. Second, on each iteration, the solution is more accurate than the solutions from either $\widetilde{\mathbf{R}}_1$ or $\widetilde{\mathbf{R}}_2$. At full convergence, the solution attains the theoretical minimum error.

We now consider a case where the parameter values of the observation-error correlations suggest that the convergence rate of the minimization will be degraded (i.e., $\chi > 1$). Using Figure 6, we can select parameter values that will increase the condition number. In particular, we set $M_o = 10$ and $D_o = 120$ km, while keeping the background-error parameter values

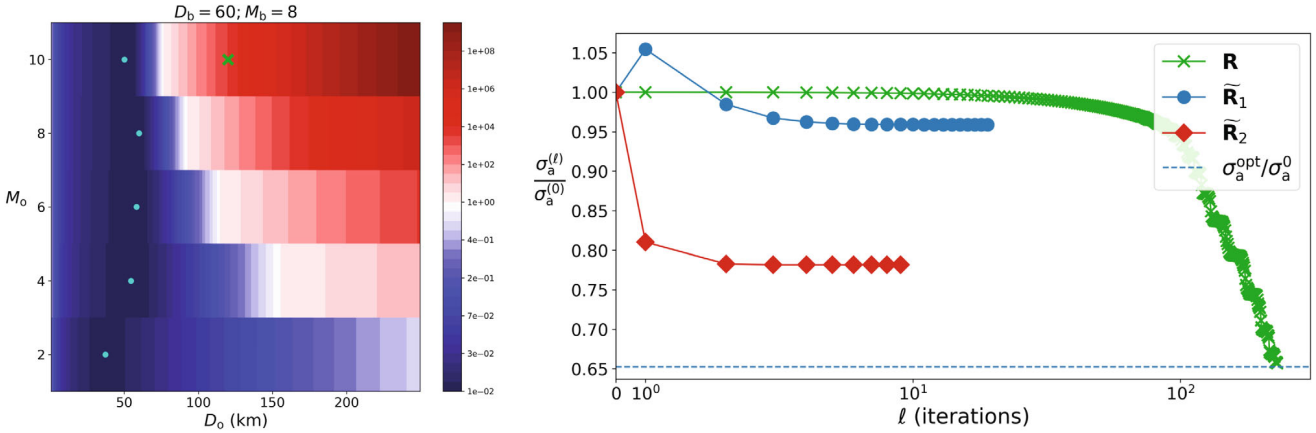
RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(a) χ for different covariance parameters **(b)** CG minimisation

FIGURE 7 (a) Same as Figure 6 but with a cross sign added to indicate the parameter pair (M_o, D_o) used for the 1D-Var experiments in panel (b). (b) $\sigma_a^{(\ell)}/\sigma_a^{(0)}$ (Equation 34) as a function of CG iteration count ℓ for three 1D-Var experiments with the same covariance parameters for $\mathbf{B}$ ($\sigma_b^2 = 1$, $M_b = 8$, $D_b = 60$ km) but different covariance parameters for $\mathbf{R}$: (1) $\mathbf{R}$ with the 'true' correlation parameters ($\sigma_o^2 = 1$, $M_o = 2$, $D_o = 30$ km); (2) a diagonal approximation, $\widetilde{\mathbf{R}}_1 = \sigma_o^2 \mathbf{I}_m$ with $\sigma_o^2 = 1$; (3) a diagonal approximation with inflated variances, $\widetilde{\mathbf{R}}_2 = \upsilon\, \sigma_o^2 \mathbf{I}_m$ where $\upsilon = 10.5$ is an optimally-estimated inflation factor. The theoretical minimum analysis-error ratio $\sigma_a^{opt}/\sigma_a^{(0)}$ (Equation 35) is marked by a horizontal dashed line.



(a) χ for different covariance parameters **(b)** CG minimisation

FIGURE 8 Same as Figure 7 but with a different parameter pair ($M_o = 10, D_o = 120$ km), indicated by the cross sign in panel (a). The optimal inflation factor in panel (b) is $\upsilon = 17$.

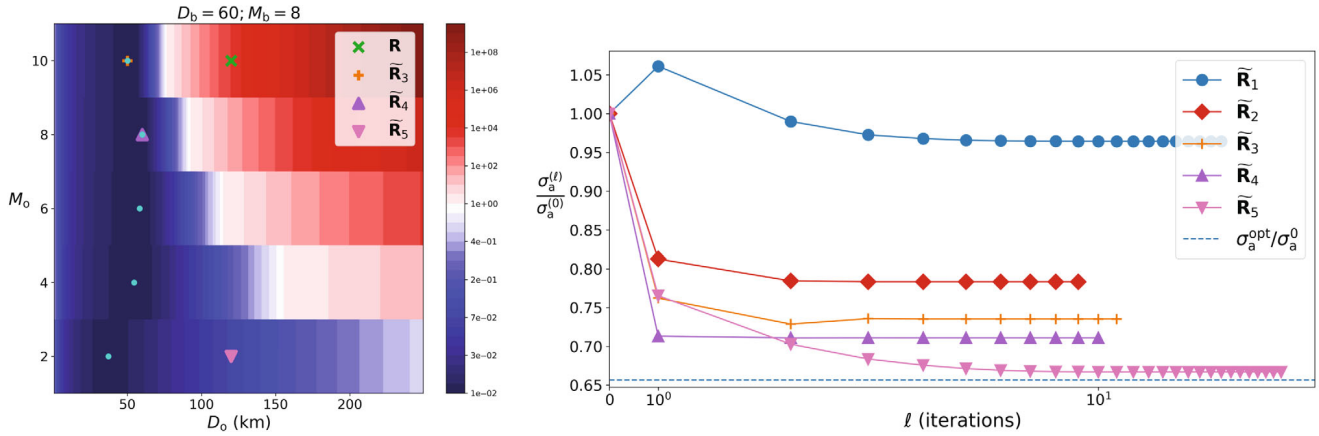
unchanged. In this set-up, the observation-error length-scale D_o is double the background-error correlation length-scale. With this set of parameter values, the condition number is increased by a factor of 10^4. In this set-up, the theoretical minimum error σ_a^{opt}/σ_b is lower than in the previous experiment: 65% instead of 68%. This decrease means that observations with highly correlated errors 'complement' the background better than those of the previous experiment.

As shown in Figure 8b, while the minimization with $\mathbf{R}$ does reach the theoretical minimum, it requires about 200 iterations to converge. If the minimization was terminated in its early iterations (< 50) then the analysis would be hardly better than that of the background and not as accurate as the solutions from either the $\widetilde{\mathbf{R}}_1$ or $\widetilde{\mathbf{R}}_2$ experiments. In this case, it would be clearly detrimental to account for the observation-error correlations instead of ignoring them.

It is interesting to note that the convergence curve for $\widetilde{\mathbf{R}}_1$ in Figure 8b differs from the one in Figure 7b even though the Hessian matrix $\mathbf{S}$ for this case is the same in both experiments. On the other hand, the assimilated observations are different in each experiment as they have different correlated errors. The difference in the convergence curves in the two experiments thus highlights the role of the right-hand side (which depends on the observations) of the system on the convergence rate of CG. While this is an important issue, we have not attempted to address it in this article.

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

RIGHTSLINK



(a) χ for different covariance parameters **(b)** CG minimisation

FIGURE 9 Same as Figure 8 but with different parameter pairs ($M_o = 10$, $D_o = 50$ km), ($M_o = 8$, $D_o = 60$ km), and ($M_o = 2$, $D_o = 120$ km) for $\mathbf{R}$ in panel (b) (experiments labeled $\widetilde{\mathbf{R}}_3$, $\widetilde{\mathbf{R}}_4$, and $\widetilde{\mathbf{R}}_5$, respectively) compared to the parameter pair ($M_o = 10$, $D_o = 120$ km) used to generate the observations. The different pairs are indicated by the different colored symbols in panel (a). The optimal inflation factor in panel (b) is $\upsilon = 17$.

The experiment $\widetilde{\mathbf{R}}_1$ results in an error reduction of only 5%, compared to the theoretical minimum of 35%. Moreover, the error reduction is non-monotonic, which is symptomatic of a more concerning issue: as $\widetilde{\mathbf{R}}_1$ is an approximation of the actual error covariances, there is no guarantee that the analysis will be more accurate than the background (even at full convergence). In this case, due to the neglected correlations in $\widetilde{\mathbf{R}}_1$, the contribution of the observations to the analysis is larger than it should be. This has a more detrimental impact on the analysis when the observations are less accurate than the background: repeating the experiment with larger observation error results in a monotonically increasing analysis error (not shown).

In the current scenario, the optimal variance inflation factor is approximately equal to 17, and the experiment with $\widetilde{\mathbf{R}}_2$ gives the best results when using a modest number of iterations (< 150). It produces a similar, rapid convergence rate as in the previous scenario (Figure 7b) and produces an accurate analysis, with a 23% error reduction compared to the theoretical minimum error reduction of 35%.

Rather than adopting a diagonal approximation, an alternative approach would be to adjust the correlation parameters to accelerate the convergence rate while trying to retain the salient features of the original correlation function, which in principle should correspond to our best available estimate of the actual correlation function. As discussed in Section 3.3, certain adjustments to the parameter settings can have a significant impact on the condition number, while inducing relatively minor changes to the correlation function and hence to σ_a^*. Moreover, previous studies have shown that even an approximate correlation structure in $\mathbf{R}$ can yield higher quality analyses than ones obtained with wrongly assuming uncorrelated observation errors (e.g., Reference 12).

Corollaries 1,2 and Figure 6 can be used as a guideline to find suitable parameters. In this scenario, we would like to pick values of M_o and D_o that are 'close enough' to our target parameter values of $M_o = 10$ and $D_o = 120$ km so as not to increase the analysis error by too much at full convergence (σ_a^*), but which produce a much smaller condition number for $\mathbf{S}$. In Figure 9, we consider three additional experiments (labeled $\widetilde{\mathbf{R}}_3$, $\widetilde{\mathbf{R}}_4$, and $\widetilde{\mathbf{R}}_5$), which all use correlation models that are approximate compared to the actual one used to generate the observation error ($M_o = 10$ and $D_o = 120$ km) but which lead to improved convergence rates. Figure 9 shows that all three experiments outperform the diagonal $\mathbf{R}$ experiments $\widetilde{\mathbf{R}}_1$ and $\widetilde{\mathbf{R}}_2$ at every iteration. We now discuss the choice of the parameter values for these experiments in relation to Corollaries 1,2 and Figure 6.

In the situation where using the accurate correlation model would degrade the condition number, for given values of M_o, M_b, and D_b, we can modify D_o to approximate the theoretical minimum condition number predicted by Corollaries 1 or 2. The experiment with $\widetilde{\mathbf{R}}_3$ corresponds to the 'extreme' case where D_o is modified using Equation (29) ($M_o > M_b$ in this experiment) so its value coincides exactly with the minimum. To do so, we retain the true value of $M_o = 10$ but use an approximate value of $D_o = 50$ km to compute the analysis, instead of the length-scale of 120 km that was used to generate the correlated observation errors. The consequence of reducing D_o and not M_o is that the correlation function

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

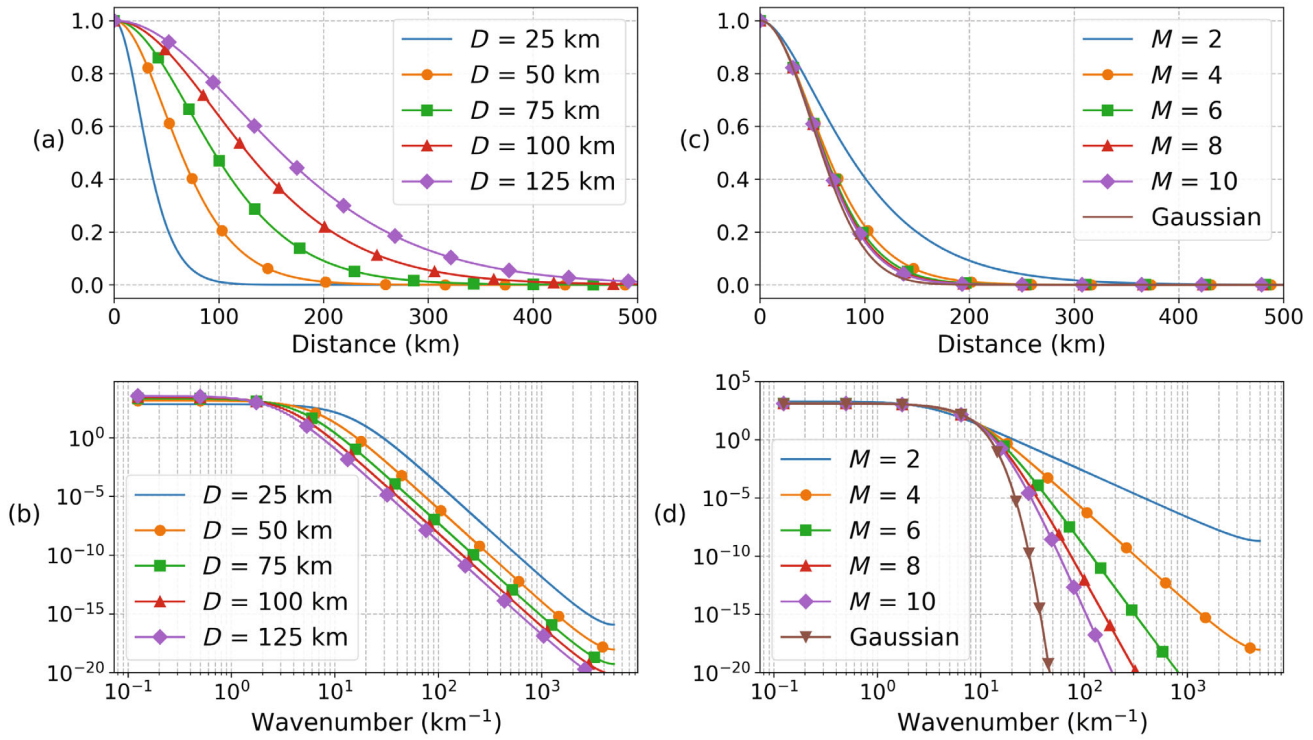


FIGURE A1 (a) Equation (A2) plotted as a function of distance r, and (b) Equation (A4) plotted as a function of wavenumber. The curves are displayed for different values of D and fixed value of $M = 4$. Panels (c) and (d) show corresponding plots where M is varied for fixed value of $D = 50$ km.

associated with $\widetilde{\mathbf{R}}_3$ has the same shape as that of $\mathbf{R}$ but covers a much smaller range (see Figure A1a). Decreasing D_o to this value reduces the condition number by a factor of 10^6. With these new parameters, $\chi = 10^{-2}$, and the condition number obtained with $\widetilde{\mathbf{R}}_3$ is lower than the one obtained with either $\widetilde{\mathbf{R}}_1$ or $\widetilde{\mathbf{R}}_2$ (the condition number with $\widetilde{\mathbf{R}}_2$ is only 17 times lower than the condition number with $\widetilde{\mathbf{R}}_1$). Figure 9b shows that the experiment with $\widetilde{\mathbf{R}}_3$ outperforms both diagonal approximations at every iteration. With this modified value of D_o, the error reduction is 27% compared to 35% with the actual value of D_o, but allows a much faster convergence.

Another possibility is to modify M_o, so that a smaller modification on D_o is required to approximate a minimum condition number predicted by Corollaries 1 or 2. When $M_b = 8$ and $D_b = 60$ km, Equation (29) associated with Corollary 1 predicts a minimum with the parameter pairs ($M_o = 10, D_o = 50$ km) and ($M_o = 8, D_o = 60$ km). It is thus possible to reach a minimum condition number with a smaller decrease of D_o if M_o is reduced from 10 to 8. The experiment with $\widetilde{\mathbf{R}}_4$ uses $M_o = M_b = 8$ and $D_b = D_o = 60$ km, which means that the correlation functions associated with $\mathbf{B}$ and $\mathbf{R}$ are very close (although $\mathbf{B}$ and $\mathbf{R}$ are not equal as there are fewer observations than grid points). Consequently, the condition number with $\widetilde{\mathbf{R}}_4$ is slightly lower than with $\widetilde{\mathbf{R}}_3$ (and thus also lower than with $\widetilde{\mathbf{R}}_1$ and $\widetilde{\mathbf{R}}_2$). This correlation model allows a slightly better error reduction than with $\widetilde{\mathbf{R}}_3$ (30% compared to 27%), while also converging slightly faster. As for the experiment with $\widetilde{\mathbf{R}}_3$, intermediate values of D_o which approach the local minima while staying closer to the true parameters might offer a better compromise between a fast convergence and good error reduction at full convergence.

In the two previous experiments, the parameter values were chosen in order to reach a local condition number minimum predicted by Corollaries 1 or 2. If the estimated parameter values have a severe impact on convergence as in Figure 8 then Figure 6 suggests that, to lower the condition number, we should reduce the values of M_o and/or D_o. In particular, if $M_o < M_b$ and $D_o < D_b$ then $\chi < 1$. Generally speaking, lower values of D_o or M_o, relative to the corresponding values of D_b and M_b, reduce the risk of the condition number being increased compared to the condition number with a diagonal $\mathbf{R}$ (i.e., of being in the red area of Figure 6 where $\chi > 1$). The parameters D_o and M_o can be reduced progressively through trial-and-error to determine a convergence rate at least as good as the one obtained with $\widetilde{\mathbf{R}}_1$. In particular, reducing M_o does not necessarily induce a visible change in the shape of the correlation function (Figure A1b) but has a large effect on its Fourier transform at high wavenumbers (Figure A1d). This correlation function is associated with the first row of the correlation matrix, which is circulant. The Fourier transform of the correlation function thus corresponds

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

RIGHTSLINK

to the eigenvalues of the correlation matrix (see Section 3.1). The Fourier transform of the correlation function can be understood as the power spectrum associated with the observation-error correlation matrix (which does not necessarily correspond to the true observation-error correlation matrix used to simulate the observations). As a consequence, decreasing M_o means that the observation-error power spectrum that we specify through the correlation matrix has larger variance at high wavenumbers than that of the actual observation-error power spectrum.

For example, in the experiment with $\widetilde{\mathbf{R}}_5$, we set $M_o = 2$ instead of 10 while keeping the correct value of $D_o = 120$ km. Although this does not match the true observation-error correlations, the modified observation-error spectrum does not induce a large increase of the analysis error: the experiment with $\widetilde{\mathbf{R}}_5$ achieves an error reduction of 33%, which is close to the theoretical minimum of 35% and better than what is obtained with either $\widetilde{\mathbf{R}}_3$ or $\widetilde{\mathbf{R}}_4$. The low sensitivity of the analysis error to this large change in the specification of the observation error can be explained by the fact that this change mostly affects the high wavenumbers where both the observation error and background error (and hence analysis error) are negligible. A misspecification of the observation error at these wavenumbers is thus unlikely to contribute significantly to the total analysis error. However, changing the observation-error spectrum at high wavenumbers by reducing M_o has a drastic effect on the conditioning of the Hessian matrix: with $\widetilde{\mathbf{R}}_5$, the condition number is approximately 20 times smaller than with $\widetilde{\mathbf{R}}_1$ (and thus slightly smaller than with $\widetilde{\mathbf{R}}_2$). This large difference in the condition number can be explained by the fact that with M_o equal to 10, the minimum eigenvalue of $\mathbf{R}$ is extremely small, which can significantly increase the amplitude of the maximum eigenvalue of the Hessian matrix according to Lemma 2. Reducing M_o can increase the amplitude of the minimum eigenvalue by several orders of magnitude (see the end of the spectrum in Figure A1d, or Equation 17 with $i = m - 1$), and thus reduce the condition number without altering significantly the observation-error spectrum at low wavenumbers. If a case of poor convergence were to be observed in a real application, then this result suggests that reducing M_o until an acceptable convergence rate is reached would be a reasonable remedy.

5 | SUMMARY AND CONCLUSIONS

Data assimilation concerns the problem of determining the optimal state of a system given observations, a background (prior) estimate of the state, and constraints that link the system state to the observations. Mathematically, the problem can be cast as one of nonlinear weighted least-squares. The technique of variational data assimilation, which is commonly used in atmospheric and ocean applications, seeks an approximate solution by using a truncated Gauss–Newton (TGN) algorithm to minimize the cost function of the nonlinear weighted least-squares problem. The TGN algorithm approximates the nonlinear problem by a connected sequence of linear sub-problems where each sub-problem is solved using a conjugate gradient (CG) algorithm.

In this article, we have studied the convergence properties of the CG algorithm with respect to parameter specifications in the background-error covariance matrix ($\mathbf{B}$) and observation-error covariance matrix ($\mathbf{R}$) whose respective inverse matrices are used to define the weights for the background and observations in the cost function. In line with common practice in variational data assimilation, we considered a CG algorithm that uses $\mathbf{B}$ as a preconditioner, which we referred to as the $\mathbf{B}$-preconditioned conjugate gradient ($\mathbf{B}$-PCG) algorithm. Our results have shown that the convergence rate of $\mathbf{B}$-PCG (and thus of the whole TGN minimization) is highly sensitive to parameters controlling the shape of typical covariance functions used to model $\mathbf{B}$ and $\mathbf{R}$. In particular, the number of CG iterations needed to reach a given tolerance can change by a few orders of magnitude depending on the relative parameter specifications between $\mathbf{B}$ and $\mathbf{R}$. This underlines the importance of including convergence impact as an additional constraint when adjusting covariance model parameters to fit covariances estimated from statistics.

We began by recalling the general upper bounds on the condition number derived by Reference 30. These upper bounds do not depend on the type of covariance matrices used, and thus can be overly pessimistic in specific cases. In order to derive more accurate bounds, we need to consider specific covariance matrices. In this article, we have focussed on covariance matrices that can be modeled as a matrix-vector product using a diffusion operator. Diffusion operators are commonly used for modeling spatial covariances in ocean data assimilation[24] and are closely related to other techniques for modeling spatial covariances in atmospheric data assimilation,[45] geostatistics,[43] inverse problems,[56] and uncertainty quantification.[57] They are well suited for problems that have large state and observation vectors, they provide convenient access to an inverse covariance operator (specifically $\mathbf{R}^{-1}$ as required by $\mathbf{B}$-PCG), and they have useful flexibility for specifying covariance functions with different characteristics.

In order to simplify the theoretical analysis, we assumed that the basic parameters of the diffusion-based covariance models for both $\mathbf{B}$ and $\mathbf{R}$ were constant. These parameters consist of the standard deviations (σ_b and σ_o), as well as

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

parameters that control the degree of smoothness (integers M_b and M_o) and spatial range (length-scales L_b and L_o) of the underlying correlation functions. (The quantities with subscripts 'b' and 'o' refer to the background and observation quantities, respectively.) With these assumptions, the covariance functions implied by the diffusion model are of Matérn type. In $\mathbb{R}$, they are auto-regressive (AR) functions of order M, which are the functions relevant for our one-dimensional (1D) analysis. Furthermore, we assumed that the grids supporting the background state and observations have uniform resolution h_b and $h_o = \zeta h_b$ where ζ is a positive integer, which implies that there are fewer observations than background grid points. Under these assumptions, we derived an analytical expression for the eigenvalues of the **B**-preconditioned Hessian matrix (**S**). By further assuming that $\mathbf{HBH}^T$ can be approximated by a diffusion operator $\mathbf{B}_o$ that is discretized directly on the observation grid, it has been possible to derive criteria that the parameter pairs (M_b, L_b) and (M_o, L_o) must jointly satisfy to obtain a minimum upper bound for the condition number of **S**. These constraints are exact when the observation and background grids coincide ($\zeta = 1$), but are affected by a minor discretization error when the observation grid is coarser than the background grid ($\zeta < 1$). We used these analytical results to interpret numerical results from experiments with a 1D variational data assimilation system (1D-Var).

First, our analytical expressions expose the already well-known result that increasing (decreasing) the ratio σ_b/σ_o leads to an increase (decrease) of the condition number of **S**. Furthermore, when $M_o = M_b$, our results show that the condition number is minimized when the same correlation model is used for both background and observation errors (i.e., $L_b = L_o$). This is consistent with the results of Tabeart et al.[30] who considered only the special case when $M_o = M_b = 2$ (i.e., when the correlation functions are second-order AR functions). However, the main contribution of our work has been to extend the analysis to the more realistic case where the background and observation errors are modeled with different smoothness parameters ($M_o \neq M_b$).

While we have derived the analytical results in terms of the parameter pairs (M_b, L_b) and (M_o, L_o), we have mainly interpreted them and the results of the numerical experiments in terms of the parameter pairs (M_b, D_b) and (M_o, D_o) where D_b and D_o are alternative ('Daley') length-scale parameters commonly used for differentiable correlation functions in data assimilation.[39] Specifically, for the 1D problem under consideration, $D_b = L_b\sqrt{2M_b - 3}$ and $D_o = L_o\sqrt{2M_o - 3}$ where $M_b > 1$ and $M_o > 1$. In terms of fixed values of D_b and D_o, the AR functions have the convenient property that they converge to Gaussian functions for large M_b and M_o. Our results have also exposed a direct relationship with closely-related ('Stein') length-scale parameters, $\rho_b = L_b\sqrt{2M_b - 1}$ and $\rho_o = L_o\sqrt{2M_o - 1}$, used in geostatistics.[41] In terms of fixed values of ρ_b and ρ_o, the AR functions also converge to Gaussian functions for large M_b and M_o, and are defined for both the differentiable ($M_b > 1$ and $M_o > 1$) and non-differentiable AR functions ($M_b = M_o = 1$).

The condition number is markedly more sensitive to the parameter specifications for the case $M_o > M_b$ than $M_o < M_b$. This has been illustrated in the numerical experiments and is evident from the analytical expression (Equation 29) that describes the relationship between (M_b, D_b) and (M_o, D_o) required to achieve the minimum upper bound of the condition number when $M_o > M_b$. In general, D_o needs to be much smaller than D_b for this optimality condition to be met because of the presence of M_b and M_o as exponents in the expressions. For this case, the small eigenvalues of **S** are amplified by $\mathbf{R}^{-1}$ more than they are damped by **B**, which can result in a drastic increase of the condition number and thus a significant risk that the convergence rate of **B**-PCG will be substantially degraded. For the case $M_o < M_b$, the minimum upper bound of the condition number is attained when the 'Stein' length-scales ρ_b and ρ_o are equal (Equation 26). In contrast with the case $M_o > M_b$, this means that the minimum upper bound is obtained when the background- and observation-error correlation functions have similar spatial range. When $M_o < M_b$ and $D_o \leq D_b$, accounting for observation-error correlations systematically improves the conditioning of **S** compared to the case when a diagonal **R** is used.

While M_o and M_b are intended as free parameters of the correlation model, to be adjusted to achieve the best possible fit to statistical estimates of correlated error, they also provide valuable leverage for controlling the conditioning of **S** when a non-diagonal **R** is used. Whereas using a non-diagonal **R** is likely to degrade significantly the convergence rate when $M_o > M_b$, it can accelerate the convergence rate compared to the case where a diagonal **R** is used when $M_o < M_b$. In practice, this situation would correspond to choosing a Gaussian-like correlation function for background error (e.g., $M_b \approx 10$) and a correlation function with fatter tails (more power at smaller scales) for observation error (e.g., $M_o = 2$). Interestingly, there is evidence in the atmospheric data assimilation literature that suggests that error correlations for certain observation types do exhibit fat tails. In this case, the interest in using a non-diagonal **R** is twofold: it can accelerate the **B**-PCG convergence rate as well as providing a more accurate correlation model. For the case where statistical estimates of the correlation parameters result in unfavourable values in terms of conditioning ($M_o > M_b$ and/or $D_o \gg D_b$), **S** can be 'reconditioned' by adjusting the values of M_o and D_o to enable faster convergence. We have shown in our 1D-Var experiments that approximate correlation models can be used to reduce the condition number without causing a significant loss of solution accuracy at full convergence. This corroborates the conclusion of several previous studies (e.g.,

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Reference 12) that even a very crude approximation of the observation-error correlations can give a better solution than one obtained by ignoring them altogether.

The analysis in this article has exposed important sensitivities of the convergence rate of **B**-PCG to fundamental parameters of diffusion model representations of **B** and **R**, and has lead to conditions for adjusting the parameters to improve the conditioning of **S**. We can expect similar results in higher dimensions where diffusion kernels have similar (Matérn-like) functional forms as those in our 1D study. However, more work is required to extend these results to account for more sophisticated diffusion models, such as ones that include diffusion tensors that are anisotropic and spatially varying, or multiple-scale and hybrid formulations that are built from linear combinations of diffusion operators. In this study, we considered a simple observation network. The convergence properties need to be revisited in an operational-like framework using a full network of diverse observations for which only a subset may be affected by spatially correlated errors in **R**, and where observation operators will be much more complex. The results from this study are a first step towards understanding and controlling the convergence properties in this more challenging framework.

ACKNOWLEDGMENTS

This work has benefitted from funding from the Copernicus Climate Change Service (C3S), one of six services of the European Union's Copernicus Programme, which is implemented by ECMWF on behalf of the European Commission. It is a contribution to the contract C3S_601_INRIA: "*Advancing ocean data assimilation methodology for climate applications*". Additional support has been provided by the French National Programme LEFE/INSU.

CONFLICT OF INTEREST STATEMENT

This study does not have any conflicts to disclose.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ENDNOTES

[1]In the current context, the time coordinate in the diffusion equation does not represent physical time but should be interpreted as a pseudo-time coordinate that controls the smoothing properties of the diffusion kernel. This explains why 'time' has been written within quotation marks.

[2]$w = \sqrt{2}$ in Reference 41 [pp. 48–50], $w = 2$ in Reference 43, and $w = 1$ in Reference 42 [Chap. 4.2].

[3]The exact condition number has been computed using results from Theorem 3 described later in Section 3.

[4]Each column (row) of a circulant matrix is a cyclic permutation of the previous column (row).

ORCID

Olivier Goux https://orcid.org/0000-0002-2848-7442
Youssef Diouane https://orcid.org/0000-0002-6609-7330

REFERENCES

1. Gratton S, Lawless AS, Nichols NK. Approximate Gauss–Newton methods for nonlinear least squares problems. SIAM J Optim. 2007;18:106–32.
2. Courtier P, Thépaut JN, Hollingsworth A. A strategy for operational implementation of 4D-Var, using an incremental approach. Q J Roy Meteorol Soc. 1994;120:1367–87.
3. Gürol S, Weaver AT, Moore AM, Piacentini A, Arango HG, Gratton S. **B**-preconditioned minimization algorithms for variational data assimilation with the dual formulation. Q J Roy Meteorol Soc. 2013;140:539–56.
4. Derber J, Rosati A. A global oceanic data assimilation system. J Phys Oceanogr. 1989;19:1333–47.
5. Lorenc AC. Development of an operational variational assimilation scheme. J Meteorol Soc Japan. 1997;75:339–46.
6. Bormann N, Bauer P. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: methods and application to ATOVS data. Q J Roy Meteorol Soc. 2010;136:1036–50.
7. Waller JA, Ballard S, Dance S, Kelly G, Nichols NK, Simonin D. Diagnosing horizontal and inter-channel observation error correlations for SEVIRI observations using observation-minus-background and observation-minus-analysis statistics. Remote Sens (Basel). 2016;8:581.
8. Liu ZQ, Rabier F. The interaction between model resolution, observation resolution and observation density in data assimilation: a one-dimensional study. Q J Roy Meteorol Soc. 2002;128:1367–86.
9. Reid R, Good S, Martin MJ. Use of uncertainty inflation in OSTIA to account for correlated errors in satellite-retrieved sea surface temperature data. Remote Sens (Basel). 2020;12:1083.

RIGHTSLINK
10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10. Rainwater S, Bishop CH, Campbell WF. The benefits of correlated observation errors for small scales. Q J Roy Meteorol Soc. 2015;141:3439–45.
11. Healy SB, White AA. Use of discrete Fourier transforms in the 1D-Var retrieval problem. Q J Roy Meteorol Soc. 2005;131:63–72.
12. Stewart LM, Dance SL, Nichols NK. Data assimilation with correlated observation errors: experiments with a 1-D shallow water model. Tellus Ser A. 2013;65:19546.
13. Ruggiero GA, Cosme E, Brankart JM, Le Sommer J. An efficient way to account for observation error correlations in the assimilation of data from the future SWOT high-resolution altimeter mission. J Atmos Oceanic Tech. 2016;33:2755–68.
14. Pinnington EM, Casella E, Dance SL, et al. Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using four-dimensional variational data assimilation. Agric Meteorol. 2016;228-229:299–314.
15. Brankart JM, Ubelmann C, Testut CE, Cosme E, Brasseur P, Verron J. Efficient parameterization of the observation error covariance matrix for square root or ensemble Kalman filters: application to ocean altimetry. Mon Weather Rev. 2009;137:1908–27.
16. Michel Y. Revisiting Fisher's approach to the handling of horizontal spatial correlations of observation errors in a variational framework. Q J Roy Meteorol Soc. 2018;144:2011–25.
17. Bédard J, Buehner M. A practical assimilation approach to extract smaller-scale information from observations with spatially correlated errors: an idealized study. Q J Roy Meteorol Soc. 2019;146:468–82.
18. Guillet O, Weaver AT, Vasseur X, Michel Y, Gratton S, Gürol S. Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. Q J Roy Meteorol Soc. 2019;145:1947–67.
19. Hu G, Dance SL. Efficient computation of matrix-vector products with full observation weighting matrices in data assimilation. Q J Roy Meteorol Soc. 2021;147:4101–21.
20. Guttorp P, Gneiting T. Studies in the history of probability and statistics XLIX on the Matérn correlation family. Biometrika. 2006;93:989–95.
21. Mirouze I, Weaver AT. Representation of correlation functions in variational assimilation using an implicit diffusion operator. Q J Roy Meteorol Soc. 2010;136:1421–43.
22. Weaver AT, Mirouze I. On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. Q J Roy Meteorol Soc. 2013;139:242–60.
23. Egbert G, Bennett A, Foreman M. Topex/Poseidon tides estimated using a global inverse model. J Geophys Res. 1994;99:24821–52.
24. Weaver AT, Courtier P. Correlation modelling on the sphere using a generalized diffusion equation. Q J Roy Meteorol Soc. 2001;127:1815–46.
25. Weaver AT, Tshimanga J, Piacentini A. Correlation operators based on an implicitly formulated diffusion equation solved with the Chebyshev iteration. Q J Roy Meteorol Soc. 2015;142:455–71.
26. Weaver AT, Chrust M, Ménétrier B, Piacentini A. An evaluation of methods for normalizing diffusion-based covariance operators in variational data assimilation. Q J Roy Meteorol Soc. 2020;147:289–320.
27. Axelsson O, Kaporin I. On the sublinear and superlinear rate of convergence of conjugate gradient methods. Numer Algorithms. 2000;25:1–22.
28. Haben SA, Lawless AS, Nichols NK. Conditioning of incremental variational data assimilation, with application to the met office system. Tellus Ser A. 2011;63:782–92.
29. Tabeart JM, Dance SL, Haben SA, Lawless AS, Nichols NK, Waller JA. The conditioning of least-squares problems in variational data assimilation. Numer Linear Algebra Appl. 2018;25:e2165.
30. Tabeart JM, Dance SL, Lawless AS, Nichols NK, Waller JA. New bounds on the condition number of the hessian of the preconditioned variational data assimilation problem. Numer. Linear Algebra Appl. 2021;29:e2405.
31. Ide K, Courtier P, Ghil M, Lorenc AC. Unified notation for data assimilation: operational, sequential and variational. J. Meteorol. Soc. Japan. 1997;75B:181–9.
32. Golub GH, Van Loan CF. Matrix computations. Baltymore, MD: J. Hopkins Uni. Press; 2013.
33. Saad Y. Iterative methods for sparse linear systems. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2003.
34. Nocedal J, Wright S. Numerical optimization. New York: Springer-Verlag GmbH; 2006.
35. Kelley CT. Iterative methods for linear and nonlinear equations. Cambridge, UK: Cambridge University Press; 1987.
36. Axelsson O. Iterative solution methods. Cambridge, UK: Cambridge University Press; 1994.
37. Lorenc AC. Optimal nonlinear objective analysis. Q J Roy Meteorol Soc. 1988;114:205–40.
38. Whittle P. Stochastic processes in several dimensions. Bull Inst Int Stat. 1963;40:974–94.
39. Daley R. Atmospheric data analysis. Cambridge, UK: Cambridge University Press; 1991.
40. Pannekoucke O, Berre L, Desroziers G. Background-error correlation length-scale estimates and their sampling statistics. Q J Roy Meteorol Soc. 2008;134:487–508.
41. Stein LM. Interpolation of spatial data. New York, NY: Springer-Verlag; 1999.
42. Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge, MA: MIT Press; 2006.
43. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J Roy Stat Soc: Series B Stat Method. 2011;73:423–98.
44. Gaspari G, Cohn SE. Construction of correlation functions in two and three dimensions. Q J Roy Meteorol Soc. 1999;125:723–57.
45. Purser RJ, Wu WS, Parrish DF, Roberts NM. Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: spatially homogeneous and isotropic Gaussian covariances. Mon Weather Rev. 2003;131:1524–35.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

46. Derber J, Bouttier F. A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus A: Dyn Meteorol Oceanography. 1999;51:195–221.
47. Weaver AT, Deltel C, Machu E, Ricci S, Daget N. A multivariate balance operator for variational ocean data assimilation. Q J Roy Meteorol Soc. 2005;131:3605–25.
48. Waller JA, Simonin D, Dance SL, Nichols NK, Ballard SP. Diagnosing observation error correlations for Doppler radar radial winds in the met Office UKV model using observation-minus-background and observation-minus-analysis statistics. Mon Weather Rev. 2016;444:3534–51.
49. Tarantola A. Inverse problem theory: methods for data fitting and model parameter estimation. Amsterdam: Elsevier; 1987.
50. Gray RM. Toeplitz and circulant matrices: a review. Found Trends Commun Inf Theory. 2005;2:155–239.
51. Harville DA. Matrix algebra from a statistician's perspective. New York: Springer; 1997.
52. Andersson E, Fisher M, Munro R, McNally A. Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. Q J Roy Meteorol Soc. 2000;126:1455–72.
53. Weston PP, Bell W, Eyre JR. Accounting for correlated error in the assimilation of high-resolution sounder data. Q J Roy Meteorol Soc. 2014;140:2420–9.
54. Tabeart JM, Dance SL, Lawless AS, et al. The impact of using reconditioned correlated observation-error covariance matrices in the met office 1D-Var system. Q J Roy Meteorol Soc. 2020;146:1372–90.
55. Gelb A. Applied optimal estimation. Cambridge: The MIT Press; 1974.
56. Bui-Thanh T, Ghattas O, Martin J, Stadler G. A computational framework for infinite-dimensional Bayesian inverse problems. Part I: the linearized case, with application to global seismic inversion. SIAM J. Sci. Comput. 2013;35:A2494–523.
57. Gmeiner B, Drzisga D, Rüde U, Scheichl R, Wohlmuth B. Scheduling massively parallel multigrid for multilevel Monte Carlo methods. SIAM J Sci Comput. 2017;39:S873–97.

How to cite this article: Goux O, Gürol S, Weaver AT, Diouane Y, Guillet O. Impact of correlated observation errors on the conditioning of variational data assimilation problems. Numer Linear Algebra Appl. 2024;31(1):e2529. https://doi.org/10.1002/nla.2529

APPENDIX A. MATéRN FUNCTIONS AND DIFFUSION OPERATORS ON $\mathbb{R}$ AND $\mathbb{S}$

A.1 Diffusion on $\mathbb{R}$

Matérn random fields on $\mathbb{R}^d$ can be derived by solving a general stochastic fractional partial differential equation (PDE).[20,38] Here, we are interested in the correlation functions of a subset of Matérn fields on $\mathbb{R}$ ($d = 1$) where parameters are chosen such that the generating PDE has a simplified form for numerical computations.

Let $\chi : z \mapsto \chi(z)$ and $\eta : z \mapsto \eta(z)$ be square-integrable functions ($\chi, \eta \in L^2(\mathbb{R})$) of the spatial coordinate $z \in \mathbb{R}$. We consider solutions of the following elliptic equation on $\mathbb{R}$:

$$\frac{1}{\gamma^2}\left(I - L^2\frac{\partial^2}{\partial z^2}\right)^M \eta(z) = \chi(z), \tag{A1}$$

where I is the identity operator, M is a positive integer, L is a length-scale parameter, and γ^2 is a normalization constant. Equation (A1) can be interpreted as the *inverse* of a diffusion operator, $\eta \mapsto \mathcal{L}^{-1}\eta$, which is formed by discretizing the time derivative of the diffusion equation with an Euler backward (implicit) scheme and by applying the resulting operator over M time steps.[21] With this interpretation, $L^2 = \mu\Delta t$ where μ is the diffusion coefficient and Δt is the time step. The integral solution of Equation (A1) is thus a diffusion operator, $\chi \mapsto \mathcal{L}\chi$. The solution, which is straightforward to derive using the Fourier transform, is a convolution operator, $\mathcal{L}\chi \equiv c \times \chi$, where $c = c(r)$ is an Mth-order AR function (a polynomial times the exponential function) given by

$$c(r) = \sum_{j=0}^{M-1} \beta_j \left(\frac{r}{L}\right)^j e^{-r/L}, \tag{A2}$$

$r = |z - z'|$ is the Euclidean distance between points z and z', and

$$\beta_j = \frac{2^j (M-1)!\,(2M-j-2)!}{j!\,(M-j-1)!\,(2M-2)!}.$$

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Setting the normalization constant to

$$\gamma^2 = \nu\, L,$$

where

$$\nu = \frac{2^{2M-1}[(M-1)!]^2}{(2M-2)!}, \tag{A3}$$

ensures that $c(0) = 1$.[21] The power spectrum of c, which is given by the Fourier transform $\hat{c}$ of c, describes the smoothness properties of c as a function of spectral scale:

$$\hat{c}(\hat{z}) = \frac{\gamma^2}{\left(1 + L^2\hat{z}^2\right)^M}, \tag{A4}$$

where $\hat{z}$ is the spectral wavenumber.

We focus on the differentiable AR functions that correspond to $M > 1$. For these functions, we use a standard parameter[39]

$$D = \sqrt{-\frac{1}{\partial^2 c/\partial z^2|_{z=z'}}},$$

to characterize the length-scale of the correlation function. The parameter D, which we call the Daley length-scale, corresponds to the distance between $z = z'$ and the mid-amplitude point of a parabola that osculates the AR function at $z = z'$. Using Equation (A2), it is straightforward to show that $D = L\sqrt{2M-3}$ (Equation 7), which is a function of both L and M. An important property of AR functions is that, for fixed D, they converge to the Gaussian function $c_g(r)$ as $M \to \infty$:

$$c_g(r) = \exp\left(-r^2/2D^2\right).$$

Figure A1 shows the effect on c and $\hat{c}$ of varying D for a fixed value of M, and vice versa. Increasing D with M held fixed increases the spatial reach of the correlation functions but does not affect their spectral decay rate at small wavelengths. On the other hand, increasing M with D held fixed results in correlation functions with thinner tails and sharper spectral decay rates at small wavelengths.

A.2 Diffusion on $\mathbb{S}$

References 29 and 30 use a SOAR function, which is equal to Equation (A2) with $M = 2$ and hence $D = L$ from Equation (7). Furthermore, they restrict the SOAR function to the circular domain ($\mathbb{S}$) of radius a by using chordal distance $r = 2a\sin(\theta/2)$ where θ is the angle between points z and z' on the circle. This ensures that $c(r)$ is positive definite on $\mathbb{S}$.[44] Taking a as the radius of the Earth, the domain $\mathbb{S}$ can be interpreted as a latitude circle at the Equator.

In this article, we have also considered a circular domain of radius a. For length-scales $L \ll a$, the correlation functions associated with the diffusion operator applied on $\mathbb{S}$ are approximately Matérn since the influence of curvature is minor. It is instructive nevertheless to present the exact correlation functions on $\mathbb{S}$, which can be derived by considering the solution of the elliptic equation

$$\frac{1}{\gamma^2}\left(I - \frac{L^2}{a^2}\frac{\partial^2}{\partial\phi^2}\right)^M \eta(\phi) = \mu(\phi), \tag{A5}$$

subject to periodic boundary conditions on the solution and its derivative:

$$\eta(-\pi) = \eta(\pi),$$
$$\frac{\partial\eta}{\partial\phi}\Big|_{\phi=-\pi} = \frac{\partial\eta}{\partial\phi}\Big|_{\phi=\pi}.$$

Solving Equation (A5) is equivalent to solving Equation (A1) on the periodic domain $-Z \le z \le Z$ with $z = a\phi$ and $Z = a\pi$.

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The solutions that satisfy the boundary conditions are of the general form

$$\eta(\phi) = \sum_{m=0}^{\infty} A_m \cos(m\phi) + B_m \sin(m\phi). \tag{A6}$$

The coefficients A_m and B_m can be determined using the orthogonality relations of the sine and cosine functions:

$$\begin{aligned}\int_{-\pi}^{\pi} \sin(m\phi)\sin(n\phi)\mathrm{d}\phi &= \int_{-\pi}^{\pi} \cos(m\phi)\cos(n\phi)\mathrm{d}\phi = \pi\ \delta_{mn},\\ \int_{-\pi}^{\pi} \sin(m\phi)\cos(n\phi)\mathrm{d}\phi &= 0\ \ \forall\ m, n,\end{aligned}$$

where δ_{mn} is the Kronecker delta. To determine A_m, we substitute Equations (A6) in (A5), multiply the resulting equation by $\cos(n\phi)$, integrate from $-\pi$ to π, and use the orthogonality relations above. This yields

$$A_m = \frac{\gamma^2}{\pi}\left(1 + \frac{L^2}{a^2}\ m^2\right)^{-M} \int_{-\pi}^{\pi} \mu(\phi)\cos(m\phi)\ \mathrm{d}\phi. \tag{A7}$$

To determine B_m, we follow the same procedure but multiply by $\sin(n\phi)$. This yields

$$B_m = \frac{\gamma^2}{\pi}\left(1 + \frac{L^2}{a^2}\ m^2\right)^{-M} \int_{-\pi}^{\pi} \mu(\phi)\sin(m\phi)\ \mathrm{d}\phi. \tag{A8}$$

Substituting Equations (A7) and (A8) into (A6), and using the trigonometric identity

$$\cos(m(\phi - \phi')) = \cos(m\phi)\ \cos(m\phi') + \sin(m\phi)\ \sin(m\phi'),$$

yields the solution

$$\eta(\phi) = \int_{-\pi}^{\pi} c(\theta)\ \mu(\phi')\ \mathrm{d}\phi',$$

where $\theta = \phi - \phi'$,

$$c(\theta) = \gamma^2 \sum_{m=0}^{\infty} c_m \cos(m\theta), \tag{A9}$$

and

$$c_m = \frac{1}{\pi}\left(1 + \frac{L^2}{a^2}\ m^2\right)^{-M}. \tag{A10}$$

The normalization factor and Daley length-scale are, respectively,

$$\gamma^2 = \frac{1}{\sum_{m=0}^{\infty} c_m},$$

and

$$D = a\sqrt{-\frac{1}{\partial^2 c/\partial\phi^2|_{\theta=0}}} = a\ \sqrt{\frac{1}{\sum_{m=0}^{\infty} m^2 c_m}}.$$

All valid continuous isotropic correlation functions on $\mathbb{S}$ can be represented by a Fourier cosine series expansion with non-negative Fourier coefficients (see Theorem 2.11 in Reference 44), which is clearly satisfied by Equations (A9)

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and (A10). The smoothness properties of the correlation function are determined by the Fourier coefficients c_m in Equation (A10). They can be seen to have a similar dependence on L and M as $\hat{c}$ in Equation (A4) where we can associate $\hat{z}$ on $\mathbb{R}$ with m^2/a^2 on $\mathbb{S}$.

APPENDIX B. PROOF OF THEOREM 5

The eigenvalues of $\mathbf{S}_o$ are bounded below by 1 (see Theorem 4), which implies that

$$\kappa(\mathbf{S}_o) \leq \max_{i \in [\![0, m-1]\!]} \lambda_i(\mathbf{S}_o) = \lambda_{\max}(\mathbf{S}_o).$$

Since $0 \leq \sin^2(y) \leq 1$ for any $y \in [0, \pi]$, $\lambda_{\max}(\mathbf{S}_o)$ is bounded by

$$\lambda_{\max}(\mathbf{S}_o) \leq \max_{x \in [0,1]} \phi(x),$$

where ϕ is a continuously differentiable function given by

$$\phi(x) = 1 + \alpha \frac{\left[1 + 4\widetilde{L}_o^2 x\right]^{M_o}}{\left[1 + 4\widetilde{L}_{b/o}^2 x\right]^{M_b}}. \tag{B1}$$

We seek a solution to the following bound-constraint problem:

$$\max_{x \in [0,1]} \phi(x). \tag{B2}$$

Let $x^* \in [0, 1]$ be a stationary point for problem (B2) and let us first assume that such point is inside the domain; that is, $\phi'(x^*) = 0$. The derivative of the function ϕ can be expressed as

$$\phi'(x) = 4\alpha \; v(x) \; w(x),$$

where

$$v(x) = \left(\frac{\widetilde{L}_o^2 M_o}{1 + 4\widetilde{L}_o^2 x} - \frac{\widetilde{L}_{b/o}^2 M_b}{1 + 4\widetilde{L}_{b/o}^2 x} \right) \quad \text{and} \quad w(x) = \frac{\left[1 + 4\widetilde{L}_o^2 x\right]^{M_o}}{\left[1 + 4\widetilde{L}_{b/o}^2 x\right]^{M_b}}.$$

Since $w(x)$ is strictly positive, and $\alpha > 0$, a stationary point inside the domain satisfies $v(x^*) = 0$. This yields

$$x^* = \frac{\widetilde{L}_{b/o}^2 M_b - \widetilde{L}_o^2 M_o}{4\widetilde{L}_o^2 \widetilde{L}_{b/o}^2 (M_o - M_b)}. \tag{B3}$$

The second derivative of f is

$$\phi''(x) = 16\alpha \; w(x) \left(\frac{\widetilde{L}_{b/o}^4 M_b}{\left(1 + 4\widetilde{L}_{b/o}^2 x\right)^2} - \frac{\widetilde{L}_o^4 M_o}{\left(1 + 4\widetilde{L}_o^2 x\right)^2} \right) + 16\alpha \; v(x)^2 \; w(x). \tag{B4}$$

Substituting (B3) into (B4) gives

$$\phi''(x^*) = 16\alpha (M_o - M_b) \; w(x^*) \left(\frac{\widetilde{L}_{b/o}^4 \widetilde{L}_o^4 (M_o - M_b)^2}{\left(\widetilde{L}_{b/o}^2 - \widetilde{L}_o^2\right)^2 M_o M_b} \right).$$

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Therefore, the stationary point x^* can be a maximum point if and only if $\phi''(x^*) < 0$; that is, if $M_o < M_b$. In addition, for x^* to be a feasible point then $0 < x^* < 1$ and from Equation (B3) the following conditions must be satisfied:

$$\widetilde{L}_{b/o}^2 M_b - \widetilde{L}_o^2 M_o < 0,$$

and

$$\widetilde{L}_{b/o}^2 M_b - \widetilde{L}_o^2 M_o > 4\widetilde{L}_o^2 \widetilde{L}_{b/o}^2 (M_o - M_b).$$

For the other cases, x^* is equal to either the lower bound ($x^* = 0$) or the upper bound ($x^* = 1$), with function values of

$$\phi(0) = 1 + \alpha,$$
$$\phi(1) = 1 + \alpha \frac{\left[1 + 4\widetilde{L}_o^2\right]^{M_o}}{\left[1 + 4\widetilde{L}_{b/o}^2\right]^{M_b}}.$$

Finally, substituting (B3) into (B1), we obtain that

$$\phi(x^*) = 1 + \alpha \left(\frac{\widetilde{L}_o^2}{M_b}\right)^{M_b} \left(\frac{M_o}{\widetilde{L}_{b/o}^2}\right)^{M_o} \left(\frac{M_b - M_o}{\widetilde{L}_o^2 - \widetilde{L}_{b/o}^2}\right)^{M_b - M_o}.$$

APPENDIX C. PROOF OF COROLLARY 1

We consider η as a function of $\widetilde{L}_o$, denoted $f(\widetilde{L}_o)$. Hereafter, the conditions (i), (ii), and (iii) will refer to the conditions stated in Theorem 5. We consider the case where condition (ii) does not hold, that is, $M_o \geq M_b$. In this case, Theorem 5 states that

$$f(\widetilde{L}_o) = 1 + \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o} \max\left\{\frac{\left(1 + 4\widetilde{L}_o^2\right)^{M_o}}{\left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}}; 1\right\}. \tag{C1}$$

Let us first assume that $\left(1 + 4\widetilde{L}_o^2\right)^{M_o} < \left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}$. Then, Equation (C1) simplifies to

$$f(\widetilde{L}_o) = 1 + \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o},$$

which is a decreasing function of $\widetilde{L}_o$. Let us now assume that $\left(1 + 4\widetilde{L}_o^2\right)^{M_o} > \left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}$. In this case, Equation (C1) becomes

$$f(\widetilde{L}_o) = 1 + \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o} \frac{\left(1 + 4\widetilde{L}_o^2\right)^{M_o}}{\left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}},$$

whose derivative is given by

$$f'(\widetilde{L}_o) = \underbrace{\frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o^2} \frac{\left(1 + 4\widetilde{L}_o^2\right)^{M_o - 1}}{\left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}}}_{>0} \left(4\widetilde{L}_o^2 (2M_o - 1) - 1\right).$$

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Under the assumption that $\widetilde{L}_o > 1/\left(2\sqrt{2M_o - 1}\right)$, $f(\widetilde{L}_o)$ is an increasing function of $\widetilde{L}_o$. As we already showed that $f(\widetilde{L}_o)$ is decreasing when $\left(1 + 4\widetilde{L}_o^2\right)^{M_o} < \left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}$, the function f then reaches its unique minimum when

$$\left(1 + 4\widetilde{L}_o^2\right)^{M_o} = \left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}.$$

APPENDIX D. PROOF OF COROLLARY 2

We consider η as a function of $\widetilde{L}_o$, denoted by $f(\widetilde{L}_o)$. Hereafter, the conditions (i), (ii), and (iii) will refer to the conditions stated in Theorem 5. Let us first assume that condition (ii) holds, that is, $M_o < M_b$ whereas condition (i) does not hold, that is,

$$\widetilde{L}_o^2 M_o \leq \widetilde{L}_{b/o} M_b. \tag{D1}$$

From (D1), we first consider the variations of $f(\widetilde{L}_o)$ when $\widetilde{L}_o$ is in the interval $\left[0, \widetilde{L}_{b/o}\sqrt{M_b/M_o}\right]$. In this case, Theorem 5 states that

$$f(\widetilde{L}_o) = 1 + \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o} \max\left\{\frac{(1 + 4\widetilde{L}_o^2)^{M_o}}{(1 + 4\widetilde{L}_{b/o}^2)^{M_b}};\ 1\right\}. \tag{D2}$$

From condition (D1), we have

$$\widetilde{L}_o^2 \leq \widetilde{L}_{b/o}^2 \frac{M_b}{M_o},$$

which implies that

$$\left(1 + 4\widetilde{L}_o^2\right)^{\frac{M_o}{M_b}} \leq \left(1 + 4\widetilde{L}_{b/o}^2 \frac{M_b}{M_o}\right)^{\frac{M_o}{M_b}},$$

and hence

$$\left(1 + 4\widetilde{L}_o^2\right)^{\frac{M_o}{M_b}} - \left(1 + 4\widetilde{L}_{b/o}^2\right) \leq \left(1 + 4\widetilde{L}_{b/o}^2 \frac{M_b}{M_o}\right)^{\frac{M_o}{M_b}} - \left(1 + 4\widetilde{L}_{b/o}^2\right) = g(\widetilde{L}_{b/o}^2), \tag{D3}$$

where the function g is given by $g(x) = (1 + 4xM_b/M_o)^{M_o/M_b} - (1 + 4x)$. Taking the first derivative of g gives

$$g'(x) = 4\left(\left(1 + 4x\frac{M_b}{M_o}\right)^{\frac{M_o}{M_b} - 1} - 1\right).$$

Since $M_o/M_b < 1$, for all $x \geq 0$, we have $(1 + 4xM_b/M_o)^{M_o/M_b - 1} < 1$ and hence $g'(x) < 0$; that is, g is a decreasing function on $[0, +\infty)$. Consequently, $g(\widetilde{L}_{b/o}^2) \leq g(0) = 0$ and thus inequality (D3) implies that

$$\left(1 + 4\widetilde{L}_o^2\right)^{\frac{M_o}{M_b}} - \left(1 + 4\widetilde{L}_{b/o}^2\right) \leq 0,$$

or, equivalently,

$$\left(1 + 4\widetilde{L}_o^2\right)^{M_o} \leq \left(1 + 4\widetilde{L}_{b/o}^2\right)^{M_b}.$$

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

By using this inequality in Equation (D2), we obtain that

$$f(\widetilde{L}_o) = 1 + \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o}.$$

In this expression, $f(\widetilde{L}_o)$ is inversely proportional to $\widetilde{L}_o$. Therefore, f decreases on the interval $\left[0, \widetilde{L}_{b/o}\sqrt{M_b/M_o}\right]$. As a consequence, the minimum of $f(\widetilde{L}_o)$ must be located on the interval $\left[\widetilde{L}_{b/o}\sqrt{M_b/M_o}, +\infty\right)$. We now study the variations of f on this interval, which means that condition (i) holds:

$$\widetilde{L}_o^2 M_o > \widetilde{L}_{b/o} M_b.$$

We assumed that condition (iii) holds when condition (i) is satisfied, which means that $f(\widetilde{L}_o)$ takes the form

$$f(\widetilde{L}_o) = 1 + \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o} \left(\frac{\widetilde{L}_o^2}{M_b}\right)^{M_b} \left(\frac{M_o}{\widetilde{L}_{b/o}^2}\right)^{M_o} \left(\frac{M_b - M_o}{\widetilde{L}_o^2 - \widetilde{L}_{b/o}^2}\right)^{M_b - M_o}.$$

The derivative of $f(\widetilde{L}_o)$ can be expressed as

$$\frac{\partial f}{\partial \widetilde{L}_o}(\widetilde{L}_o) = \frac{\sigma_b^2 \nu_b \widetilde{L}_{b/o}}{\sigma_o^2 \nu_o \widetilde{L}_o^2 \left(\widetilde{L}_o^2 - \widetilde{L}_{b/o}^2\right)} \left(\frac{\widetilde{L}_o^2}{M_b}\right)^{M_b} \left(\frac{M_o}{\widetilde{L}_{b/o}^2}\right)^{M_o} \left(\frac{M_b - M_o}{\widetilde{L}_o^2 - \widetilde{L}_{b/o}^2}\right)^{M_b - M_o} \left[\widetilde{L}_o^2 (2M_o - 1) - \widetilde{L}_{b/o}^2 (2M_b - 1)\right]. \tag{D4}$$

If conditions (i) and (ii) are met, we have $\widetilde{L}_o^2 - \widetilde{L}_{b/o}^2 > 0$. Therefore, the stationary point for f satisfies

$$\begin{aligned} \frac{\partial f}{\partial \widetilde{L}_o}(\widetilde{L}_o) = 0 &\Leftrightarrow \widetilde{L}_o^2 (2M_o - 1) = \widetilde{L}_{b/o}^2 (2M_b - 1) \\ &\Leftrightarrow \widetilde{L}_o = \widetilde{L}_{b/o}\sqrt{\frac{2M_b - 1}{2M_o - 1}}. \end{aligned} \tag{D5}$$

Since $M_o < M_b$, it follows that

$$\widetilde{L}_{b/o}\sqrt{\frac{2M_b - 1}{2M_o - 1}} > \widetilde{L}_{b/o}\sqrt{\frac{M_b}{M_o}}.$$

We are now interested in examining the behavior of f on the intervals $\left[\widetilde{L}_{b/o}\sqrt{M_b/M_o}, \widetilde{L}_{b/o}\sqrt{(2M_b - 1)/(2M_o - 1)}\right]$ and $\left[\widetilde{L}_{b/o}\sqrt{(2M_b - 1)/(2M_o - 1)}, +\infty\right)$. For the first interval, we can show that f is decreasing since $\partial f(\widetilde{L}_o)/\partial \widetilde{L}_o < 0$ from Equation (D4) if

$$\widetilde{L}_o^2 (2M_o - 1) < \widetilde{L}_{b/o}^2 (2M_b - 1) \;\Leftrightarrow\; \widetilde{L}_o < \widetilde{L}_{b/o}\sqrt{\frac{2M_b - 1}{2M_o - 1}}.$$

Similarly, for the second interval, we can show that f is increasing since $\partial f(\widetilde{L}_o)/\partial \widetilde{L}_o > 0$ from Equation (D4) if

$$\widetilde{L}_o^2 (2M_o - 1) > \widetilde{L}_{b/o}^2 (2M_b - 1) \;\Leftrightarrow\; \widetilde{L}_o > \widetilde{L}_{b/o}\sqrt{\frac{2M_b - 1}{2M_o - 1}}.$$

Therefore, the stationary point (D5) is the unique minimum of f. Finally, multiplying both sides of Equation (D5) by $h_o\sqrt{2M_o - 1}$ yields Equation (26).

RIGHTSLINK

10991506, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/nla.2529 by Cochrane France, Wiley Online Library on [23/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Chapter 5

Article 2: On the impact of observation-error correlations in data assimilation, with application to along-track altimeter data

One of the key messages of Chapter 4 is that choosing an appropriate observation-error correlation model for DA is not just a matter of finding the best fit to the actual observation error-correlations: accounting for numerical efficiency when selecting a correlation error model is also critical. In order to better understand how to find this balance between realism and efficiency, we first need to understand what makes a 'good' observation-error correlation model. We have learned in Chapter 4 how to alter an observation error correlation model to reduce the condition number, but we need to understand the cost of this alteration in terms of the accuracy of $\mathbf{x}^*$, the solution of the minimisation at full convergence, with respect to the true state $\mathbf{x}^{(t)}$.

We attempt in Chapter 5 to offer a broader view of the response of a DA system to observation-error correlations. This is achieved by analysing an idealised DA system from a scale-dependent perspective, which offers a clearer view of the role that observation-error correlations play in DA. In order to assess the practical relevance of these results, we compare the behaviour predicted in this idealized framework to the actual response of a more realistic system (NEMOVAR) in which we have implemented an observation error correlation operator based on a diffusion operator.

These results are described in an article to be submitted to the *Quarterly Journal of the Royal Meteorological Society* and included hereafter.

ORIGINAL ARTICLE

On the impact of observation error correlations in data assimilation, with application to along-track altimeter data

Olivier Goux[1] | **Anthony T. Weaver**[1] | **Selime Gürol**[1] | **Oliver Guillet**[2] | **Andrea Piacentini**[1]

[1]CECI UMR 5318, CERFACS and CNRS, Toulouse, France

[2]CNRM UMR 3589, Météo-France and CNRS, Toulouse, France

Correspondence
Olivier Goux, CERFACS, Toulouse, France
Email: goux@cerfacs.fr

Funding information
Copernicus Climate Change Service, Grant/Award Number: C3S_601_INRIA; French National Programme LEFE/INSU

Data assimilation involves estimating the state of a system by combining observations from various sources with a background estimate of the state. The weights given to the observations and background state depend on their specified error covariance matrices. Observation errors are often assumed to be uncorrelated even though this assumption is inaccurate for many modern data-sets such as those from satellite observing systems. As methods allowing for a more realistic representation of observation-error correlations are emerging, our aim in this article is to provide insight on their expected impact in data assimilation. First, we use a simple idealised system to analyse the effect of observation-error correlations on the spectral characteristics of the solution. Next, we assess the relevance of these results in a more realistic setting in which simulated along-track (nadir) altimeter observations with correlated errors are assimilated in a global ocean model using a three-dimensional variational assimilation (3D-Var) method. Correlated observation errors are modelled in the 3D-Var system using a diffusion operator.

When the correlation length scale of observation error is small compared to that of background error, inflating the observation-error variances can mitigate most of the nega-

tive effects from neglecting the observation-error correlations. Accounting for observation-error correlations in this situation still outperforms variance inflation since it allows small-scale information in the observations to be more effectively extracted and does not affect the convergence of the minimization. Conversely, when the correlation length scale of observation error is large compared to that of background error, the effect of observation-error correlations cannot be properly approximated with variance inflation. However, the correlation model needs to be constructed carefully to ensure the minimization problem is adequately conditioned so that a robust solution can be obtained. Practical ways to achieve this are discussed.

KEYWORDS
data assimilation, observation error, correlations, diffusion operator, altimeter, variance inflation

1 | INTRODUCTION

The background- and observation-error covariance matrices (**B** and **R**, respectively) are fundamental components of data assimilation systems for the atmosphere and ocean. Much research has been devoted to the specification of **B**, to the extent that quite sophisticated representations of **B** have been developed for operational systems (*e.g.*, for a review see Bannister, 2008). Comparatively, **R** has received much less attention, especially with regard to the specification of spatial correlations (off-diagonal elements in **R**), which are generally neglected. Neglecting spatial observation-error correlations, however, is not a good assumption for many modern observation datasets. Multiple studies describe significant observation-error correlations in, for example, radiances from the Infrared Atmospheric Sounding Interferometer (Bormann and Bauer, 2010) and the Spinning Enhanced Visible and InfraRed Imager (Waller et al., 2016a; Michel, 2018), sea-level anomalies from satellite altimeters (Storto et al., 2019), Doppler radial winds (Waller et al., 2016b), and satellite-retrieved sea surface temperature (Reid et al., 2020). Moreover, particularly large spatial error correlations are to be expected in a new-generation of high-resolution satellite observations, such as wide-swath altimeter data (Esteban-Fernandez and Peral, 2018; King and Martin, 2021) and total surface current velocity measurements (Waters et al., 2024). The benefits of accounting for correlated observation error, particularly for the analysis of small spatial scales, have been highlighted in several studies (Healy and White, 2005; Stewart et al., 2013; Rainwater et al., 2015; Pinnington et al., 2016).

Inter-channel error correlations in satellite radiance observations are now accounted for in multiple Numerical Weather Prediction systems (Weston et al., 2014; Bormann et al., 2016; Campbell et al., 2017; Geer, 2019), with a positive impact on forecast scores. Accounting for spatial correlations, however, has proved to be more challenging, since, unlike observations from different satellite instrument channels, observations at different spatial locations cannot be easily separated into manageable, independent subsets of observations for computing correlations. Though still in early development, practical methods are emerging to account for horizontal correlations in **R**, particularly in

the context of variational data assimilation.

In variational data assimilation, a cost function is minimised to find the state of the system that best fits the background state and observations in a weighted least-squares sense, where the weighting matrices are defined by $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$. The minimization strategy is typically based on an iterative Gauss-Newton algorithm (Gratton et al., 2007), which in the data assimilation literature is usually referred to as an incremental algorithm (Courtier et al., 1994). Central to the Gauss-Newton algorithm is a so-called inner loop that involves minimising a quadratic cost function using a conjugate gradient (CG) method. An important feature of the algorithm is that matrices do not need to be defined explicitly; only operators representing matrix-vector products are required. Furthermore, it is customary to employ $\mathbf{B}$ as a preconditioning matrix for CG (*e.g.*, Derber and Rosati, 1989; Gürol et al., 2014). In a $\mathbf{B}$-preconditioned CG algorithm, an operator needs to be specified for $\mathbf{B}$ (or a factored form of $\mathbf{B}$) as well as for the inverse covariance matrix, $\mathbf{R}^{-1}$. The fact that $\mathbf{R}^{-1}$, and not $\mathbf{R}$, is required by the minimization algorithm brings additional computational challenges when accounting for spatially correlated observation errors.

To compute a matrix-vector product with $\mathbf{R}^{-1}$, Hu and Dance (2024) propose the local Singular Value Decomposition (SVD) - Fast Multipole Method (FMM), which is a variant of the original SVD-FMM method of Hu and Dance (2021) but designed to be more computationally efficient and to require a reduced number of parallel processor communications. The main purpose of their method is to reduce the cost of computing the product of a vector with $\mathbf{R}^{-1}$, where they assume the latter to be available in explicit matrix form. To do so, the domain is partitioned into 'boxes' containing subsets of observations. The block-diagonal components of $\mathbf{R}^{-1}$ that correspond to observations in the same box are stored and applied directly to sub-vectors associated with those observations. The remaining block components of $\mathbf{R}^{-1}$ are approximated, by using truncated SVD for blocks that are associated with observations within a prescribed 'interaction' area, or by explicitly setting the blocks to zero if the observations are located beyond the interaction area. Crucially, the method relies on the availability of $\mathbf{R}^{-1}$, which in Hu and Dance (2024) is obtained by inverting a covariance matrix built explicitly from a parametric covariance function.

Guillet et al. (2019) consider a different approach, which uses knowledge of $\mathbf{R}$ to construct a meaningful $\mathbf{R}^{-1}$ operator, but does not require $\mathbf{R}$ itself to be built explicitly (and then to be inverted), which for large data-sets is challenging. This is achieved by considering a specific set of covariance functions from the Matérn class for which analytical expressions of the inverse covariance operator are known and straightforward to represent numerically as a product of sparse matrices. As discussed in Guillet et al. (2019) and references therein, the inverse covariance operator associated with these functions can be interpreted as the inverse of an iterative diffusion process. The diffusion tensor and number of diffusion iterations are adjustable parameters that control the spatial-range and shape of the associated Matérn covariance functions. To use this approach to model $\mathbf{R}^{-1}$, the (inverse) diffusion operator is discretised on a triangular mesh with nodes at the observation locations. The Finite Element Method (FEM) with standard $\mathcal{P}_1$ elements is used for spatially discretizing the inverse diffusion operator. Since each node on the mesh depends only on its direct neighbours, $\mathbf{R}^{-1}$ is sparse and scalable with large volumes of data on memory-distributed computers. With knowledge of the underlying covariance function, the parameters of the diffusion model can be estimated by fitting the function to samples of observation-error covariances obtained, for example, from Desroziers diagnostics (Desroziers et al., 2005).

An indirect way to account for spatially correlated observation errors is to augment the observation vector with spatial derivatives of the observations in addition to the observations themselves and to use a diagonal approximation for the error covariance matrix of the augmented observation vector (Brankart et al., 2009; Chabot et al., 2015; Ruggiero et al., 2016; Yaremchuk et al., 2018; Bédard and Buehner, 2019). In a continuous framework, Guillet et al. (2019, Section 4.1) show that, for specific parameter settings, this is equivalent to the diffusion approach for modelling a non-diagonal $\mathbf{R}^{-1}$ for the non-augmented observations. Assuming that an estimate of $\mathbf{R}$ is available

and can be stored in matrix form, Ruggiero et al. (2016) and Yaremchuk et al. (2018) describe an approach for estimating the weights (the inverse error variances) of the augmented observation-error covariance matrix that best fits the estimate of **R** in a squared Frobenius norm sense. Although the method does not require a direct implementation of a non-diagonal **R**, it substantially increases the effective size of the observation vector (by almost a factor of five in the example of Ruggiero et al. (2016)) and is primarily designed to handle observations with structured distributions for which simple finite-difference formulae can be used to compute spatial derivatives. Adapting the method to sparse and heterogeneously distributed observations requires more sophisticated numerical techniques such as those presented in Guillet et al. (2019, Section 4.2).

The purpose of the current article is threefold. First, we aim to complement other studies by providing further insight on the expected impact of correlated observation error in data assimilation. We address this initially by studying the optimal analysis equations in an idealised, one-dimensional (1D) framework where **B** and **R** are described by circulant matrices and hence diagonalizable in a Fourier basis. By analysing the equations in Fourier space, we are able to highlight the scale-dependent characteristics of the optimal solution, focussing on the role of **R**. We then use the same idealised framework to analyse the sensitivity of the solution to a misspecification of **R**. Second, we assess the practical relevancy of these results experimentally using a global-ocean three-dimensional variational data assimilation (3D-Var) configuration of the NEMOVAR system (Mogensen et al., 2009), where assimilated sea-level anomalies from (nadir) altimeters are taken to have correlated observation errors along the satellite track. While the configuration is realistic, we design the experiments in an idealised way to help with the interpretation of the results in relation with the theoretical analysis presented earlier. We also pay particular attention to the convergence properties of the (**B**-preconditioned) CG minimization, which previous studies, involving mainly theoretical analysis, have shown to be highly sensitive to a non-diagonal **R** (Haben et al., 2011; Tabeart et al., 2018, 2020; Goux et al., 2024). Third, with this work, we provide a practical demonstration of the method of Guillet et al. (2019) in an operationally-based data assimilation system. We view this as an important first step towards applying the method to account for spatially correlated error in other data-sets such as wide-swath altimeter and satellite-retrieved sea-surface temperature.

The structure of the article is as follows. In Section 2, we present the data assimilation problem and the idealised theoretical framework for studying the impact on the analysis of both correctly and incorrectly specified representations of **R**. In Section 3, we describe the ocean data assimilation system, focussing on the method for accounting for along-track correlated errors in altimeter observations. We provide more details on the method and its technical implementation in Appendix A. In Section 3, we also present the idealised experimental framework that we use for assessing the impact of the correlated observation-error model. We discuss the results of the experiments in Section 4. Finally, we provide a summary and conclusions in Section 5.

2 | A SIMPLIFIED THEORETICAL FRAMEWORK

In this section, we first recall the basic equations describing the optimal variational analysis in a weakly non-linear system. We then present the idealised framework that allows us to interpret the analysis equations in Fourier space and to investigate the sensitivity of the analysis to the observation-error covariances.

2.1 | The variational analysis equations

Data assimilation methods seek to combine a background state $\mathbf{x}_\mathrm{b} \in \mathbb{R}^n$ and observations $\mathbf{y}_\mathrm{o} \in \mathbb{R}^p$ in a statistically optimal manner (Fletcher, 2022). We assume the background state to be the sum of the true state $\mathbf{x}_\mathrm{t} \in \mathbb{R}^n$ that we

are attempting to estimate, and an additive Gaussian error $\boldsymbol{\epsilon}_\mathrm{b} \in \mathbb{R}^n$ with mean zero and covariance matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$:

$$\mathbf{x}_\mathrm{b} = \mathbf{x}_\mathrm{t} + \boldsymbol{\epsilon}_\mathrm{b}. \tag{1}$$

We define a generalised observation operator $\mathcal{G}$ such that, for any estimate of the state of the system $\mathbf{x} \in \mathbb{R}^n$, the vector $\mathcal{G}(\mathbf{x})$ represents how the state of the system would appear through the observation network. In 4D-Var, $\mathcal{G}$ will include the time-dependent model operator as well as the observation operator itself (Courtier et al., 1994). We assume the observations to be the sum of the counterpart of the true state in observation space and an additive Gaussian error $\boldsymbol{\epsilon}_\mathrm{o} \in \mathbb{R}^p$ with mean zero and covariance matrix $\mathbf{R} \in \mathbb{R}^{p \times p}$:

$$\mathbf{y}_\mathrm{o} = \mathcal{G}(\mathbf{x}_\mathrm{t}) + \boldsymbol{\epsilon}_\mathrm{o}. \tag{2}$$

In variational data assimilation, the analysis is obtained by minimizing a cost function $\mathcal{J}$ that measures the simultaneous fit of $\mathbf{x}$ to $\mathbf{x}_\mathrm{b}$ and $\mathbf{y}_\mathrm{o}$:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}_\mathrm{b}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{G}(\mathbf{x}) - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}},$$

which can be written in terms of an increment $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_\mathrm{b}$ as

$$\mathcal{J}(\mathbf{x}_\mathrm{b} + \delta\mathbf{x}) = \frac{1}{2}\|\delta\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{G}(\mathbf{x}_\mathrm{b} + \delta\mathbf{x}) - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}}. \tag{3}$$

The inverse error covariance matrices $\mathbf{B}^{-1}$ and $\mathbf{R}^{-1}$ define symmetric, positive-definite (SPD) weighting matrices for the squared norm $\|\mathbf{x}\|^2_{\mathbf{A}^{-1}} = \mathbf{x}^\mathsf{T}\mathbf{A}^{-1}\mathbf{x}$. The minimizing solution of this cost function with nonlinear constraints is generally obtained by minimizing successive quadratic cost functions involving linearised constraints (Courtier et al., 1994). For the sake of simplicity, we will focus on the first of these approximating quadratic cost functions, which is obtained by replacing $\mathcal{G}$ with a linear operator

$$\mathcal{G}(\mathbf{x}_\mathrm{b} + \delta\mathbf{x}) \simeq \mathcal{G}(\mathbf{x}_\mathrm{b}) + \mathbf{G}\delta\mathbf{x} \tag{4}$$

where $\mathbf{G}$ is the Jacobian matrix of $\mathcal{G}$ (or of a simplified operator that approximates $\mathcal{G}$) evaluated at $\mathbf{x} = \mathbf{x}_\mathrm{b}$. Substituting Equation (4) in Equation (3) yields the quadratic cost function

$$\begin{aligned} J(\delta\mathbf{x}) &= \frac{1}{2}\|\delta\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{G}(\mathbf{x}_\mathrm{b}) + \mathbf{G}\delta\mathbf{x} - \mathbf{y}_\mathrm{o}\|^2_{\mathbf{R}^{-1}} \\ &= \frac{1}{2}\|\delta\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathbf{G}\delta\mathbf{x} - \mathbf{d}\|^2_{\mathbf{R}^{-1}}, \end{aligned}$$

where

$$\mathbf{d} = \mathbf{y}_\mathrm{o} - \mathcal{G}(\mathbf{x}_\mathrm{b}) \tag{5}$$

is the innovation vector. Under these assumptions, the expression for the analysis increment, $\delta\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{a} - \mathbf{x}_\mathrm{b}$, that

minimizes the cost function J (associated with the first outer loop) is well known (*e.g.*, Fletcher, 2022):

$$\delta \mathbf{x}_\mathrm{a} = \mathbf{K}\mathbf{d}$$

where

$$\mathbf{K} = \mathbf{B}\mathbf{G}^\mathsf{T}\left(\mathbf{G}\mathbf{B}\mathbf{G}^\mathsf{T} + \mathbf{R}\right)^{-1} = \mathbf{P}_\mathrm{a}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1},$$

is the Kalman gain matrix, and

$$\mathbf{P}_\mathrm{a} = \left(\mathbf{B}^{-1} + \mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}\right)^{-1}$$

is the analysis-error covariance matrix (in the linear approximation). In observation space, the analysis increment is given by

$$\delta \mathbf{y}_\mathrm{a} = \mathbf{G}\delta \mathbf{x}_\mathrm{a} = \mathbf{S}\mathbf{d} \tag{6}$$

where

$$\mathbf{S} = \mathbf{G}\mathbf{K} = \mathbf{G}\mathbf{B}\mathbf{G}^\mathsf{T}\left(\mathbf{G}\mathbf{B}\mathbf{G}^\mathsf{T} + \mathbf{R}\right)^{-1} = \mathbf{G}\mathbf{P}_\mathrm{a}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}. \tag{7}$$

Equation (7) is called the *sensitivity* matrix since $\mathbf{S} = \partial \delta \mathbf{y}_\mathrm{a}/\partial \mathbf{d}$ and hence its elements quantify the sensitivity of the analysis increment (in observation space) with respect to the innovations (Cardinali et al., 2004; Fowler et al., 2018).

2.2 | A Fourier decomposition of the analysis equations

To gain insight on the effect of the Kalman gain with a non-diagonal $\mathbf{R}$, we consider a simplified framework where instructive theoretical results can be derived. In particular, we assume the following:

1. the domain is 1D and periodic;
2. the state space consists of a single variable defined at discrete locations on a uniform grid;
3. observations of the state variable are direct and available at uniformly distributed locations (*i.e.*, $\mathbf{G}$ is a selection operator);
4. the background- and observation-error covariances are homogeneous.

From the last assumption, we can define

$$\mathbf{R} = \sigma_\mathrm{o}^2\mathbf{C}_\mathrm{o}$$
$$\text{and} \qquad \mathbf{G}\mathbf{B}\mathbf{G}^\mathsf{T} = \sigma_\mathrm{b}^2\mathbf{C}_\mathrm{b},$$

where σ_o^2 is the observation-error variance, σ_b^2 is background-error variance in observation space, $\mathbf{C}_\mathrm{o}$ is the observation-error correlation matrix, and $\mathbf{C}_\mathrm{b}$ is the background-error correlation matrix in observation space. The main interest of this idealised framework is that $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{G}\mathbf{B}\mathbf{G}^\mathsf{T}$, and their associated correlation matrices, are circulant matrices,

meaning that each of their rows can be obtained as a cyclic permutation of the previous row (see Lemma 1 in Goux et al., 2024).

All circulant matrices are diagonal in a Fourier basis (Gray, 2005). Let $\mathbf{F}_p$ be the $p \times p$ matrix whose element on the k-th row and ℓ-th column is defined by the Fourier basis function

$$[\mathbf{F}_p]_{k,\ell} = \frac{1}{\sqrt{p}} e^{j2\pi \frac{k\ell}{p}}, \tag{8}$$

where j denotes the unit imaginary number ($j^2 = -1$). We can define the Fourier transform of any vector $\mathbf{x}$ as

$$\widehat{\mathbf{x}} = \mathbf{F}_p^{\mathsf{H}} \mathbf{x},$$

where the superscript 'H' denotes conjugate (Hermitian) transpose. The matrix $\mathbf{F}_p$ is unitary; *i.e.*, $\mathbf{F}_p^{\mathsf{H}} = \mathbf{F}_p^{-1}$, so that $\mathbf{F}_p^{\mathsf{H}}\mathbf{F}_p = \mathbf{F}_p\mathbf{F}_p^{\mathsf{H}} = \mathbf{I}$.

The matrices $\mathbf{R}$ and $\mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}}$ are both diagonal in the Fourier basis defined by $\mathbf{F}_p$, and can thus be expressed as

$$\begin{aligned} \mathbf{R} &= \mathbf{F}_p \mathbf{\Lambda}_{\mathrm{o}} \mathbf{F}_p^{\mathsf{H}} \\ \text{and} \qquad \mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}} &= \mathbf{F}_p \mathbf{\Lambda}_{\mathrm{b}} \mathbf{F}_p^{\mathsf{H}}, \end{aligned}$$

where the matrices $\mathbf{\Lambda}_{\mathrm{o}}$ and $\mathbf{\Lambda}_{\mathrm{b}}$ are diagonal. The k-th diagonal element of $\mathbf{\Lambda}_{\mathrm{o}}$ is $\sigma_{\mathrm{o}}^2 \, \lambda_{\mathrm{o}}^{(k)}$ where $\lambda_{\mathrm{o}}^{(k)}$ is the k-th eigenvalue of $\mathbf{C}_{\mathrm{o}}$. Similarly, the k-th diagonal element of $\mathbf{\Lambda}_{\mathrm{b}}$ is $\sigma_{\mathrm{b}}^2 \, \lambda_{\mathrm{b}}^{(k)}$ where $\lambda_{\mathrm{b}}^{(k)}$ is the k-th eigenvalue of $\mathbf{C}_{\mathrm{b}}$.

Each eigenvector of $\mathbf{R}$ (*i.e.*, each column of $\mathbf{F}_p$) can be associated with a different wavenumber p/k as seen in Equation (8). To facilitate physical interpretation, this non-dimensional wavenumber can be converted into a 'spatial scale' by multiplying it by h, the distance between consecutive observation locations. This spatial scale ph/k will be used as the horizontal axis in figures showing eigenvalue spectrums of circulant matrices throughout this article. An eigenvalue of $\mathbf{R}$ (*i.e.*, diagonal element of $\mathbf{\Lambda}_{\mathrm{o}}$) can be interpreted as the *scale-dependent* observation-error variance at the spatial scale defined by the associated eigenvector. Specifically, $\mathbf{\Lambda}_{\mathrm{o}}$ corresponds to the covariance matrix of $\widehat{\boldsymbol{\epsilon}}_{\mathrm{o}}$, the Fourier transform of the observation error:

$$\mathbf{\Lambda}_{\mathrm{o}} = \mathbf{F}_p^{\mathsf{H}} \mathbf{R} \mathbf{F}_p = \mathbf{F}_p^{\mathsf{H}} \mathbb{E}\left[\boldsymbol{\epsilon}_{\mathrm{o}} \boldsymbol{\epsilon}_{\mathrm{o}}^{\mathsf{T}}\right] \mathbf{F}_p = \mathbb{E}\left[\mathbf{F}_p^{\mathsf{H}} \boldsymbol{\epsilon}_{\mathrm{o}} \boldsymbol{\epsilon}_{\mathrm{o}}^{\mathsf{T}} \mathbf{F}_p\right] = \mathbb{E}\left[\widehat{\boldsymbol{\epsilon}}_{\mathrm{o}} \widehat{\boldsymbol{\epsilon}}_{\mathrm{o}}^{\mathsf{H}}\right],$$

where $\mathbb{E}[\,]$ denotes the mathematical expectation operator. Similarly, the eigenvalues of $\mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}}$ describe the scale-dependent variances of $\mathbf{G}\boldsymbol{\epsilon}_{\mathrm{b}}$ (*i.e.*, the variances of the Fourier transform of the background error mapped to observation space):

$$\mathbf{\Lambda}_{\mathrm{b}} = \mathbf{F}_p^{\mathsf{H}} \mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}} \mathbf{F}_p = \mathbf{F}_p^{\mathsf{H}} \mathbf{G} \mathbb{E}\left[\boldsymbol{\epsilon}_{\mathrm{b}} \boldsymbol{\epsilon}_{\mathrm{b}}^{\mathsf{T}}\right] \mathbf{G}^{\mathsf{T}} \mathbf{F}_p = \mathbb{E}\left[(\mathbf{F}_p^{\mathsf{H}} \mathbf{G} \boldsymbol{\epsilon}_{\mathrm{b}})(\mathbf{F}_p^{\mathsf{H}} \mathbf{G} \boldsymbol{\epsilon}_{\mathrm{b}})^{\mathsf{H}}\right].$$

As $\mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}}$ and $\mathbf{R}$ are both circulant, the sensitivity matrix $\mathbf{S}$ is also circulant[1], and thus diagonal in the Fourier basis defined by $\mathbf{F}_p$:

$$\mathbf{S} = \mathbf{F}_p \mathbf{\Lambda}_{\mathrm{s}} \mathbf{F}_p^{\mathsf{H}}, \tag{9}$$

[1] The sum, product and inverse of circulant matrices are also circulant matrices.

where $\mathbf{\Lambda}_\mathrm{s}$ is a diagonal matrix. We can obtain an explicit expression for the eigenvalues of $\mathbf{S}$ by combining Equations (7) and (9) and using the orthogonality property of $\mathbf{F}_p$ to yield

$$\mathbf{\Lambda}_\mathrm{s} = \mathbf{\Lambda}_\mathrm{b} \left(\mathbf{\Lambda}_\mathrm{b} + \mathbf{\Lambda}_\mathrm{o}\right)^{-1} = \left(\mathbf{\Lambda}_\mathrm{b}^{-1} + \mathbf{\Lambda}_\mathrm{o}^{-1}\right)^{-1} \mathbf{\Lambda}_\mathrm{o}^{-1}. \tag{10}$$

The k-th diagonal element of $\mathbf{\Lambda}_\mathrm{s}$, which corresponds to the k-th eigenvalue of $\mathbf{S}$, can then be expressed as

$$\lambda_\mathrm{s}^{(k)} = \frac{1}{1+\nu^{(k)}} \quad \text{where} \quad \nu^{(k)} = \frac{\sigma_\mathrm{o}^2 \lambda_\mathrm{o}^{(k)}}{\sigma_\mathrm{b}^2 \lambda_\mathrm{b}^{(k)}}. \tag{11}$$

Using Equation (9), the Fourier transform of Equation (6) has the simple form

$$\widehat{\delta \mathbf{y}}_\mathrm{a} = \mathbf{F}_p^\mathrm{H} \delta \mathbf{y}_\mathrm{a} = \mathbf{F}_p^\mathrm{H}\, \mathbf{S}\, \mathbf{d} = \mathbf{F}_p^\mathrm{H}\, \mathbf{S}\, \mathbf{F}_p \mathbf{F}_p^\mathrm{H}\, \mathbf{d} = \mathbf{\Lambda}_\mathrm{s} \widehat{\mathbf{d}}.$$

Let $\widehat{\delta y}_\mathrm{a}^{(k)}$ and $\widehat{d}^{(k)}$ denote the k-th elements of the vectors $\widehat{\delta \mathbf{y}}_\mathrm{a}$ and $\widehat{\mathbf{d}}$, respectively. The relationship between the analysis increment in observation space and the innovation is trivial in Fourier space:

$$\widehat{\delta y}_\mathrm{a}^{(k)} = \lambda_\mathrm{s}^{(k)} \widehat{d}^{(k)} = \left(\frac{1}{1+\nu^{(k)}}\right) \widehat{d}^{(k)}. \tag{12}$$

Note that Equation (12) has the same basic form as the expression of the analysis increment of a scalar quantity, obtained by combining a background estimate and a single direct observation (*e.g.*, Fisher, 2007):

$$\delta y_\mathrm{a} = \delta x_\mathrm{a} = \left(\frac{1}{1+\nu}\right) d, \quad \text{with} \quad \nu = \frac{\sigma_\mathrm{o}^2}{\sigma_\mathrm{b}^2}. \tag{13}$$

In the scalar case, σ_o^2 and σ_b^2 fully describe the observation- and background-error variance, respectively. In our case, however, the error correlations determine effective error variances at each spatial scale through the eigenvalues $\lambda_\mathrm{o}^{(k)}$ and $\lambda_\mathrm{b}^{(k)}$. In turn, this determines a different weighting between the background and observations at each spatial scale. The eigenvalue $\lambda_\mathrm{s}^{(k)}$ can be interpreted as the *sensitivity* of the analysis to the observations at the k-th spatial scale. Hereafter, we will refer to $\left(\lambda_\mathrm{s}^{(k)}\right)_{k \in [\![1,p]\!]}$ as the scale-dependent sensitivities.

The covariance matrix of the analysis error in observation space is also circulant and diagonalised by $\mathbf{F}_p$:

$$\mathbf{G}\mathbf{P}_\mathrm{a}\mathbf{G}^\mathsf{T} = \mathbf{F}_p \mathbf{\Lambda}_\mathrm{a} \mathbf{F}_p^\mathrm{H}, \tag{14}$$

where $\mathbf{\Lambda}_\mathrm{a}$ is a diagonal matrix. Its k-th element is denoted $\sigma_\mathrm{a}^2 \lambda_\mathrm{a}^{(k)}$ and can be interpreted as a variance at a specific spatial scale of the analysis error in observation space. From Equations (7), (9) and (10), it follows that

$$\mathbf{\Lambda}_\mathrm{a} = \left(\mathbf{\Lambda}_\mathrm{b}^{-1} + \mathbf{\Lambda}_\mathrm{o}^{-1}\right)^{-1}.$$

The k-th diagonal element of $\mathbf{\Lambda}_\mathrm{a}$ is thus given by

$$\sigma_\mathrm{a}^2 \lambda_\mathrm{a}^{(k)} = \left(\frac{1}{\sigma_\mathrm{b}^2 \lambda_\mathrm{b}^{(k)}} + \frac{1}{\sigma_\mathrm{o}^2 \lambda_\mathrm{o}^{(k)}}\right)^{-1}. \tag{15}$$

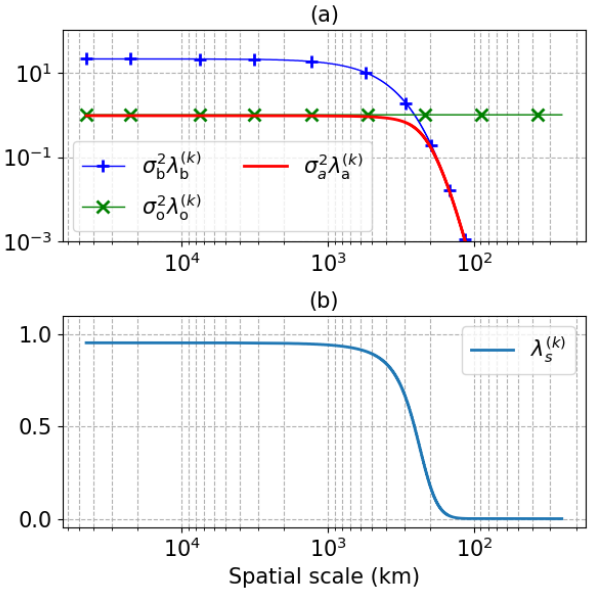


FIGURE 1 (a) Scale-dependent variance of the background error ($\sigma_b^2\lambda_b^{(k)}$), observation error ($\sigma_o^2\lambda_o^{(k)}$), and analysis error ($\sigma_a^2\lambda_a^{(k)}$; Equation (15)), as a function of spatial scale. (b) Scale-dependent sensitivity ($\lambda_s^{(k)}$; Equation (11)) as a function of spatial scale. In this idealised 1D problem, the domain size is 40000 km; the background-error correlation function is approximately Gaussian with a length scale of 200 km; the observation errors are uncorrelated; and the background- and observation-error variances are both equal to one (see text for further description). The spatial scale is computed as ph/k; *i.e.*, the inverse of the non-dimensional wavenumber k/p from Equation (8) multiplied by the distance between observations ($h = 25$ km). The scale on the vertical axis has been cut-off at 10^{-3} (the minimum background-error variance reaches 10^{-13} for the smallest resolved spatial scale).

Equation (15) is similar to the expression for the analysis-error variance for the scalar problem described earlier:

$$\sigma_a^2 = \left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right)^{-1}. \tag{16}$$

The equation pairs [(13), (16)] and [(12), (15)] highlight the similarity between the scalar example and what occurs at each spatial scale in the circulant framework.

We illustrate the properties of Equations (12) and (15) with a specific example. Consider a 1D domain of length 40000 km for which there is an observation every 25 km. The domain can be interpreted as a latitude circle at the Equator where observations are available every quarter of a degree in longitude. Both the observation- and background-error variances are set to one unit ($\sigma_o^2 = \sigma_b^2 = 1$). The background-error correlation function is taken to be a 10-th order auto-regressive (quasi-Gaussian) function with a length scale of 250 km (see Section 3.2). The eigenvalues $\lambda_b^{(k)}$ of the associated correlation matrix are computed using the analytical expression given by Equation (16) in Goux et al. (2024). The observation error is taken to be uncorrelated, implying that $\lambda_o^{(k)} = 1$ for all k.

In this idealised example, the Kalman gain – or more precisely the sensitivity matrix – acts as a spectral filter to determine how damped the innovation should be at each spatial scale in order to produce an analysis increment with minimum error variance. As such, the analysis contains the best compromise between background and observations *at each spatial scale*. Figure 1(a) shows the scale-dependent variances of observation error ($\sigma_o^2\lambda_o^{(k)}$), background error

($\sigma_b^2\lambda_b^{(k)}$), and analysis error ($\sigma^2\lambda_a^{(k)}$) as a function of spatial scale, while Figure 1(b) shows the scale-dependent sensitivity of the analysis to the observations ($\lambda_s^{(k)}$) as a function of spatial scale. From these figures, we can make the following basic observations. At *large* ($> 10^3$ km) spatial scales, $\sigma_a^2\lambda_a^{(k)} \simeq \sigma_o^2\lambda_o^{(k)} \ll \sigma_b^2\lambda_b^{(k)}$ and $\lambda_s^{(k)} \simeq 1$, so that $\widehat{\delta y}_a^{(k)} \simeq \widehat{d}^{(k)}$ from Equation (12), and thus the analysis at observation locations fits the observations. At *small* (< 150 km) spatial scales, $\sigma_a^2\lambda_a^{(k)} \simeq \sigma_b^2\lambda_b^{(k)} \ll \sigma_o^2\lambda_o^{(k)}$ and $\lambda_s^{(k)} \simeq 0$, so that $\widehat{\delta y}_a^{(k)} \simeq 0$, and thus the analysis at observation locations remains close to the background. At spatial scale 250 km ($k = 180$), where the background- and observation-error variance intersect, $\sigma_a^2\lambda_a^{(k)} = \frac{1}{2}\sigma_b^2\lambda_b^{(k)} = \frac{1}{2}\sigma_o^2\lambda_o^{(k)}$ and $\lambda_s^{(k)} = 1/2$, so that $\widehat{\delta y}_a^{(k)} = \frac{1}{2}\widehat{d}^{(k)}$, and thus the analysis at observation locations is half-way between the background and observations.

Equations (12) and (15) assume perfect knowledge of the background- and observation-error covariances as well as perfect practical capability to use these covariances to define the Kalman gain. In practice, the error covariance matrices applied in the Kalman gain (which is defined implicitly in variational data assimilation through the iterative minimization of the cost function) are not exactly representative of the (unknown) 'true' observation- and background-error covariance matrices, **R** and **B**. In the next section, we revisit the analysis above but considering the case where the observation-error covariances are misspecified.

2.3 | Misspecified observation-error covariances

In this section, we study the effect of misspecifying the observation-error covariances, drawing inspiration from Fisher (2007), who studied the effect of misspecifying the background-error variance in a scalar example. We introduce the notation $\widetilde{\mathbf{R}}$ to denote the *practical* observation-error covariance matrix used in the gain matrix, which can potentially differ from the *true* observation-error covariance matrix, **R**. When using $\widetilde{\mathbf{R}}$ instead of **R**, the resulting gain, sensitivity and analysis-error covariance matrices are suboptimal, and will be denoted by $\widetilde{\mathbf{K}}$, $\widetilde{\mathbf{S}}$, and $\widetilde{\mathbf{P}}_a$, respectively, to distinguish them from their 'true' counterparts **K**, **S**, and $\mathbf{P}_a$ that are defined in terms of **R**. Eigenvalues and variances associated with the suboptimal (practical) matrices $\widetilde{\mathbf{R}}$, $\widetilde{\mathbf{S}}$ and $\widetilde{\mathbf{P}}_a$ are also denoted with a tilde. As the focus here is on observation error, we assume that the practical and true background-error covariance matrices are equal.

With potentially misspecified observation-error covariances, we can define the practical gain matrix, the sensitivity matrix and the eigenvalues of the latter using expressions analogous to those in the optimal case:

$$\widetilde{\mathbf{K}} = \mathbf{B}\mathbf{G}^{\mathsf{T}}\left(\mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}} + \widetilde{\mathbf{R}}\right)^{-1}, \quad \widetilde{\mathbf{S}} = \mathbf{G}\widetilde{\mathbf{K}}$$

and

$$\widetilde{\lambda}_s^{(k)} = \left(1 + \frac{\widetilde{\sigma}_o^2\widetilde{\lambda}_o^{(k)}}{\sigma_b^2\lambda_b^{(k)}}\right)^{-1}. \tag{17}$$

The expression for the practical analysis-error covariance matrix, however, becomes more complex as it involves both the practical and true error covariance matrices. For any $\widetilde{\mathbf{K}}$, potentially different from the optimal **K**, the practical analysis-error covariance matrix $\widetilde{\mathbf{P}}_a$ can be defined as (*e.g.*, Equation 4.2.12 in Gelb, 1974)

$$\widetilde{\mathbf{P}}_a = (\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{G})\mathbf{B}(\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{G})^{\mathsf{T}} + \widetilde{\mathbf{K}}\,\mathbf{R}\,\widetilde{\mathbf{K}}^{\mathsf{T}}. \tag{18}$$

As in the optimal case, we focus on the analysis-error covariance matrix in observation space, which is given by

$$\mathbf{G}\widetilde{\mathbf{P}}_{\mathrm{a}}\mathbf{G}^{\mathsf{T}} = (\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{G}\mathbf{B}\mathbf{G}^{\mathsf{T}}(\mathbf{I} - \widetilde{\mathbf{S}}) + \widetilde{\mathbf{S}}\,\mathbf{R}\,\widetilde{\mathbf{S}}^{\mathsf{T}}.$$

As in the optimal case (Equation (14)), we can use $\mathbf{F}_p$ to diagonalize the left-hand side of Equation (18):

$$\mathbf{F}_p\widetilde{\mathbf{\Lambda}}_{\mathrm{a}}\mathbf{F}_p^{\mathsf{H}} = \mathbf{G}\widetilde{\mathbf{P}}_{\mathrm{a}}\mathbf{G}^{\mathsf{T}} \tag{19}$$

where $\widetilde{\mathbf{\Lambda}}_{\mathrm{a}}$ is a diagonal matrix. Combining Equations (18) and (19), and using the fact that $\mathbf{F}_p^{\mathsf{H}}\mathbf{F}_p = \mathbf{I}$, yields

$$\widetilde{\mathbf{\Lambda}}_{\mathrm{a}} = (\mathbf{I} - \widetilde{\mathbf{\Lambda}}_{\mathrm{s}})\mathbf{\Lambda}_{\mathrm{b}}(\mathbf{I} - \widetilde{\mathbf{\Lambda}}_{\mathrm{s}}) + \widetilde{\mathbf{\Lambda}}_{\mathrm{s}}\mathbf{\Lambda}_{\mathrm{o}}\widetilde{\mathbf{\Lambda}}_{\mathrm{s}}.$$

From the expression for the elements of $\widetilde{\mathbf{\Lambda}}_{\mathrm{s}}$ in Equation (17), the diagonal elements of $\widetilde{\mathbf{\Lambda}}_{\mathrm{a}}$ can be expressed as

$$\widetilde{\sigma}_{\mathrm{a}}^2\widetilde{\lambda}_{\mathrm{a}}^{(k)} = \frac{\sigma_{\mathrm{b}}^2\lambda_{\mathrm{b}}^{(k)}\left(\widetilde{\sigma}_{\mathrm{o}}^2\widetilde{\lambda}_{\mathrm{o}}^{(k)}\right)^2 + \sigma_{\mathrm{o}}^2\lambda_{\mathrm{o}}^{(k)}\left(\sigma_{\mathrm{b}}^2\lambda_{\mathrm{b}}^{(k)}\right)^2}{\left(\widetilde{\sigma}_{\mathrm{o}}^2\widetilde{\lambda}_{\mathrm{o}}^{(k)} + \sigma_{\mathrm{b}}^2\lambda_{\mathrm{b}}^{(k)}\right)^2}. \tag{20}$$

The elements $\widetilde{\sigma}_{\mathrm{a}}^2\widetilde{\lambda}_{\mathrm{a}}^{(k)}$ are the scale-dependent analysis-error variances resulting from the use of a suboptimal representation of the observation-error covariances. To help interpret Equation (20), we first normalize it by $\sigma_{\mathrm{a}}^2\lambda_{\mathrm{a}}^{(k)}$ in Equation (15). This yields the expression

$$\eta_{\widetilde{\mathrm{a}},\mathrm{a}}^{(k)} = \frac{1 + \left(\eta_{\mathrm{o,b}}^{(k)}\right)^2\left(\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)}\right)^2 + \left[1 + \left(\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)}\right)^2\right]\eta_{\mathrm{o,b}}^{(k)}}{1 + \left(\eta_{\mathrm{o,b}}^{(k)}\right)^2\left(\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)}\right)^2 + 2\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)}\eta_{\mathrm{o,b}}^{(k)}} \tag{21}$$

where $\eta_{\widetilde{\mathrm{a}},\mathrm{a}}^{(k)} = \widetilde{\sigma}_{\mathrm{a}}^2\widetilde{\lambda}_{\mathrm{a}}^{(k)}/\sigma_{\mathrm{a}}^2\lambda_{\mathrm{a}}^{(k)}$, $\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)} = \widetilde{\sigma}_{\mathrm{o}}^2\widetilde{\lambda}_{\mathrm{o}}^{(k)}/\sigma_{\mathrm{o}}^2\lambda_{\mathrm{o}}^{(k)}$ and $\eta_{\mathrm{o,b}}^{(k)} = \sigma_{\mathrm{o}}^2\lambda_{\mathrm{o}}^{(k)}/\sigma_{\mathrm{b}}^2\lambda_{\mathrm{b}}^{(k)}$. It is easy to see from Equation (21) that, as expected, $\widetilde{\sigma}_{\mathrm{a}}^2\widetilde{\lambda}_{\mathrm{a}}^{(k)} \geq \sigma_{\mathrm{a}}^2\lambda_{\mathrm{a}}^{(k)}$ ($\eta_{\widetilde{\mathrm{a}},\mathrm{a}}^{(k)} \geq 1$) and $\widetilde{\sigma}_{\mathrm{a}}^2\widetilde{\lambda}_{\mathrm{a}}^{(k)} = \sigma_{\mathrm{a}}^2\lambda_{\mathrm{a}}^{(k)}$ if $\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)} = 1$.

Figure 2 shows $\eta_{\widetilde{\mathrm{a}},\mathrm{a}}^{(k)}$ plotted as a function of $\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)}$ (horizontal axis) and $\eta_{\mathrm{o,b}}^{(k)}$ (vertical axis). Values of $\eta_{\widetilde{\mathrm{o}},\mathrm{o}}^{(k)}$ less (greater) than one correspond to an underestimation (overestimation) of the observation-error variance at spatial scale k, while values of $\eta_{\mathrm{o,b}}^{(k)}$ less (greater) than one correspond to a more (less) accurate background than observation at spatial scale k. The most striking feature of Figure 2 is the asymmetry in magnitude between the top left/bottom right corners and the top right/bottom left corners. If the observations are more accurate than the background at a given spatial scale, overestimating $\widetilde{\sigma}_{\mathrm{o}}^2\widetilde{\lambda}_{\mathrm{o}}^{(k)}$ is worse (in terms of increasing the analysis-error variance) than underestimating it; *i.e.*, underfitting the observations is worse than overfitting them. If the observations are less accurate than the background at a given spatial scale, overestimating $\widetilde{\sigma}_{\mathrm{o}}^2\widetilde{\lambda}_{\mathrm{o}}^{(k)}$ is better than underestimating it; *i.e.*, underfitting the observations is better than overfitting them.

In other words, at any spatial scale, it is better to be too close to the most accurate source of information than too far from it. Note that this last conclusion would also stand if the scale-dependent background-error variance was misspecified instead of the scale-dependent observation-error variance. This can be inferred from Figure 6 in Fisher (2007) which, although derived for a misspecification of the background-error variance in the scalar example, also translates to the scale-dependent background-error variance in our idealised circulant framework.

A less visually impressive but perhaps more impactful feature is that the iso-contours of the normalised analysis-

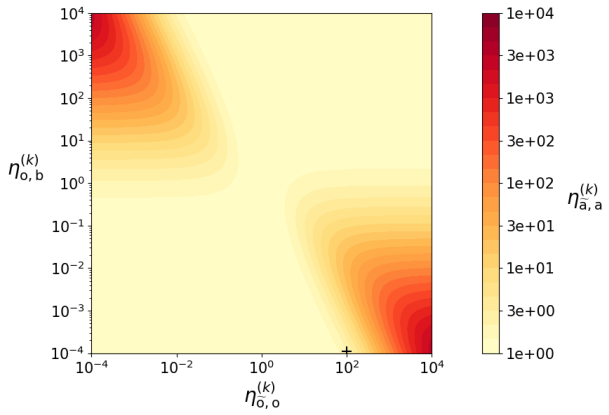


FIGURE 2 Ratio of the practical scale-dependent analysis-error variance to the optimal scale-dependent analysis-error variance ($\eta^{(k)}_{\mathrm{a,a}}$; see Equation (21)) where the former is associated with a misspecification of the observation-error variance at a given spatial scale ($\eta^{(k)}_{\mathrm{o,o}} \neq 1$). The black cross marks the point where $\eta^{(k)}_{\mathrm{o,b}} = 10^{-4}$ and $\eta^{(k)}_{\mathrm{o,o}} = 10^2$ (referenced later in the main text).

error variance recede close to the top and bottom of the figure. This shows that a misspecification of $\widetilde{\sigma}^2_{\mathrm{o}}\widetilde{\lambda}^{(k)}_{\mathrm{o}}$ — whether it is overestimated or underestimated — can only have a significant impact on the analysis error if $\sigma^2_{\mathrm{o}}\lambda^{(k)}_{\mathrm{o}}$ and $\sigma^2_{\mathrm{b}}\lambda^{(k)}_{\mathrm{b}}$ are relatively close to each other at this spatial scale. For example, if $\sigma^2_{\mathrm{o}}\lambda^{(k)}_{\mathrm{o}} < \sigma^2_{\mathrm{b}}\lambda^{(k)}_{\mathrm{b}}$ by a factor of 10000, then overestimating $\widetilde{\sigma}^2_{\mathrm{o}}\widetilde{\lambda}^{(k)}_{\mathrm{o}}$ by a factor as large as 100 has no significant impact on the analysis error (point indicated by the cross symbol in the figure). In such a case, the analysis will remain extremely close to the observations at this spatial scale.

3 | DATA ASSIMILATION SYSTEM AND EXPERIMENTAL FRAMEWORK

The insight from the idealised problem in the previous section will be helpful for interpreting the results from the ocean data assimilation experiments presented in Section 4. In this section, we describe the ocean data assimilation system, focussing on the method for accounting for correlated error in along-track altimeter observations, which is a new development to the system. We also detail the experimental framework, which involves the assimilation of sea-surface height (SSH) innovations that are constructed from simulated error samples.

3.1 | The ocean data assimilation system

The experiments are performed with NEMOVAR (Mogensen et al., 2009, 2012), an incremental variational ocean data assimilation system that has been developed for the ocean model NEMO (Madec and the NEMO System Team, 2023). NEMOVAR is used operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) and the UK Meteorological Office (Met Office; *e.g.*, see Chrust et al. (2025) and Mignac et al. (2024) for up-to-date descriptions of the ECMWF and Met Office ocean data assimilation systems, respectively). For the purposes of this study where the innovation vector is simulated, only the numerical code for the quadratic cost function minimization (inner loop) of NEMOVAR is required. The inner-loop configuration is global and similar to that of Balmaseda et al. (2013). The global grid is the tri-polar (ORCA) grid of NEMO, with approximately 1° horizontal resolution in the Extratropics and

refined meridional resolution in the Tropics, with a minimum value of 0.3° directly at the Equator. The model has 42 vertical levels, 18 of which are in the upper 200 m. Assimilation increments are computed for the ocean model state variables consisting of potential temperature (T), practical salinity (S), the two components of horizontal velocity (U and V) and SSH.

The data assimilation method is based on 3D-Var with the First-Guess at Appropriate Time (FGAT) approach. The control variables in the 3D-Var are T, and the so-called *unbalanced* parts of S and SSH. The control variables are linked to the state variables (T, S, U, V and SSH) via a non-linear balance operator (Weaver et al., 2005). The background-error covariance matrix is taken to be block-diagonal with respect to the control variables. Each covariance block is factored into a symmetric product of a correlation matrix and a diagonal matrix of standard deviations. The correlation matrices are modelled using a diffusion operator (Weaver and Courtier, 2001), with diffusion parameters chosen such that the smoothing kernel (correlation function) is approximately Gaussian. The diffusion-based operator approach for modelling correlation functions will be presented in more detail in the next section in the context of **R**. For all three control variables, the background-error standard deviations and correlation length scales are parameterised in terms of either the background state or geographical position, as described in Appendix B of Balmaseda et al. (2013). The specific values of the covariance parameters are given in Table B1 of Balmaseda et al. (2013).

Only SSH innovations are assimilated in the experiments since our main objective in this study is to evaluate the impact of the newly developed correlated-error model for sea-level anomaly (SLA) observations derived from along-track nadir altimeters. For generating the synthetic SSH innovations (see Section 3.3), we only require the geographical locations of the SLA observations. For this, we use a realistic network of SLA observations from ENVISAT and Jason-2 altimeter observations on January 1st, 2009. Locations close to land and poleward of 50°S and 50°N have been rejected. The total number of SLA observation locations retained is 15460, which provides a dense coverage of the ocean surface. The SSH observation-error standard deviations are set to 5 cm and are inflated near coastlines as described in Mogensen et al. (2012).

Model estimates of SLA (*i.e.*, the model-estimated SSH minus an estimate of the Mean Dynamic Topography) are mapped to the observation locations using bi-linear interpolation. The balance relations are linearised around a realistic background state from January 1st, 2009. The composition of the spatial interpolation operator and the linearised balance operator constitutes the linearised generalised observation operator (**G**) that links the control variable increments to the innovation vector in Equation (4), and that will be used to define the innovation itself in Equation (32). The quadratic cost function is minimised iteratively using the Restricted **B**-preconditioned Conjugate Gradient (RBCG) algorithm described in Gürol et al. (2014) (see their Algorithm 3). On each iteration of RBCG, methods are required to compute a matrix-vector product with $\mathbf{B}$ and $\mathbf{R}^{-1}$ (methods to apply $\mathbf{B}^{-1}$ and $\mathbf{R}$ are not required, however).

3.2 | Accounting for along-track error correlations in altimeter observations

In this section, we describe the method that has been implemented in NEMOVAR to account for along-track error correlations in nadir[2] altimeter data. The choice of nadir altimeter data is motivated by several factors. First, they are widely available and already assimilated in NEMOVAR; second, they are considered to be affected by errors correlated along track, which could be due to representativeness error associated with spatial scales observed but not resolved by the model (*e.g.*, Storto et al., 2019), or due to instrumental error associated with, for example, imperfect geophysical corrections or contamination between satellite footprints (*e.g.*, see Section 1.8 of Stammer and Cazenave, 2018); third, their 1D nature vastly simplifies the implementation of a correlated error model in NEMOVAR, even though much of the code development is generic and applicable to other observation types as well; and fourth, the interpretation of

[2]As opposed to wide-swath altimeters such as SWOT.

the experimental results in relation to the theoretical 1D analysis is made much easier.

We assume that errors affecting observations from different nadir instruments are not cross-correlated. The matrix $\mathbf{R}$ is thus block diagonal where each block ($\mathbf{R}_{\text{nadir}}$) corresponds to a specific nadir instrument. This matrix and its inverse can be written in the standard factored form

$$\begin{aligned} \mathbf{R}_{\text{nadir}} &= \boldsymbol{\Sigma}\mathbf{C}\boldsymbol{\Sigma} \\ \text{and} \qquad \mathbf{R}_{\text{nadir}}^{-1} &= \boldsymbol{\Sigma}^{-1}\mathbf{C}^{-1}\boldsymbol{\Sigma}^{-1} \end{aligned} \tag{22}$$

where $\boldsymbol{\Sigma}$ is a diagonal matrix of standard deviations and $\mathbf{C}$ is a correlation matrix. To define the correlation matrix and its inverse, we use the implicit diffusion operator approach adapted to unstructured meshes described in Guillet et al. (2019). As with the implicit diffusion operator applied to structured meshes (*e.g.*, for modelling $\mathbf{B}$ in NEMOVAR as described in Weaver et al. (2016)), the correlation matrix and its inverse have the general form

$$\begin{aligned} \mathbf{C} &= \boldsymbol{\Gamma}\mathbf{D}\boldsymbol{\Gamma} \\ \text{and} \qquad \mathbf{C}^{-1} &= \boldsymbol{\Gamma}^{-1}\mathbf{D}^{-1}\boldsymbol{\Gamma}^{-1} \end{aligned} \tag{23}$$

where $\mathbf{D}$ ($\mathbf{D}^{-1}$) is the matrix representing a diffusion (inverse diffusion) operator, and $\boldsymbol{\Gamma}$ is a diagonal matrix of normalisation factors that ensures that the diagonal elements of $\mathbf{C}$ are approximately equal to 1. The main requirement for applying the diffusion-operator approach to $\mathbf{R}_{\text{nadir}}$ is to adopt an appropriate spatial discretization method for defining $\mathbf{D}$ and $\mathbf{D}^{-1}$ on a 1D mesh that can become unstructured along the satellite tracks due to, for example, missing observations or quality control decisions that result in observations being excluded from the assimilation.

3.2.1 | Finite-Element Method formulation

Following Guillet et al. (2019), we discretize the diffusion operator using a finite-element method (FEM) with $\mathcal{P}_1$-FEM basis functions (*e.g.* see Ciarlet (2002)). We focus on the main points in this section and relegate the details on the numerical implementation to Appendix A. The symmetric expressions for $\mathbf{D}$ and $\mathbf{D}^{-1}$ with the FEM formulation are (cf. Equations (42) and (43) in Guillet et al. (2019))

$$\begin{aligned} \mathbf{D} &= \left[\left(\mathbf{M} + \mathbf{A} \right)^{-1} \mathbf{M} \right]^{m} \mathbf{M}^{-1} \\ \text{and} \qquad \mathbf{D}^{-1} &= \mathbf{M} \left[\mathbf{M}^{-1} \left(\mathbf{M} + \mathbf{A} \right) \right]^{m}, \end{aligned} \tag{24}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{A}$ is the stiffness matrix[3] and m is the number of implicit diffusion iterations. The mass and stiffness matrices are SPD and sparse (see Figure 4 of Guillet et al., 2019) as the only non-zero off-diagonal elements are those associated with consecutive observation locations along the satellite track. In the case of nadir altimeter data, there are thus only three non-zero elements on each row of $\mathbf{A}$. Matrix-vector products with $\mathbf{M}$ and $\mathbf{A}$ can thus be implemented efficiently even with a large volume of observations. Moreover, as observations only interact with their direct neighbours on the mesh, the matrix-vector products are straightforward to parallelize on multiple processors. In NEMOVAR, vectors in observation space can be distributed following the tiled model domain decomposition. Nodes located within each tile (processor) are considered local nodes. In addition to these local nodes, a halo is created

[3] To denote the stiffness matrix, we use the notation $\mathbf{A}$ instead of the standard notation $\mathbf{K}$ (*e.g.*, as presented in Guillet et al. (2019)) to avoid confusion with the Kalman gain matrix.

that contains those nodes on other tiles which are connected to the local nodes by an edge of the mesh. The value of vectors inside this halo are updated once before each matrix-vector product with $\mathbf{M}+\mathbf{A}$ using Message Passing Interface (MPI) exchanges.

Let Ω denote the domain of the observations (in the case of nadir altimeters, the domain is the satellite track). Let the vector z denote a position on this domain. Let ψ_i denote a $\mathcal{P}_1$-FEM basis function of the i-th observation, *i.e.*; the function is equal to 1 at the position of the i-th observation, equal to 0 at the position of other observations, and is linear in-between (see Figure 3 of Guillet et al., 2019). The kl-th matrix element of $\mathbf{M}$ and $\mathbf{A}$ is given by the integral equations (cf. Equations (27) and (28) in Guillet et al. (2019))

$$\mathbf{M}_{k,l} = \int_\Omega \psi_k(z)\psi_l(z)dz \tag{25}$$

$$\text{and} \qquad \mathbf{A}_{k,l} = \int_\Omega [\boldsymbol{\kappa}(z)\boldsymbol{\nabla}\psi_k(z)]^\mathsf{T}\boldsymbol{\nabla}\psi_l(z)dz, \tag{26}$$

where $\boldsymbol{\nabla}$ denotes the gradient operator, and $\boldsymbol{\kappa}$ is the diffusion tensor. Equations (25) and (26) show that each element of the matrices of $\mathbf{M}$ and $\mathbf{A}$, respectively, can be expressed as an integral over the whole domain. Computing the value of these integrals separately for each matrix element is not computationally efficient. Instead, we use the partition of the domain into sub-domains by the mesh where the sub-domains correspond to the segments between observation locations in 1D (triangles in 2D). We cycle through each sub-domain, and increment all integrals that are non-zero on it before moving on to the next sub-domain. This process is usually referred to as the assembly process in the FEM literature (see Appendix A for more details).

To limit the cost of applying $\mathbf{M}^{-1}$, we adopt here the classical mass-lumping approach where $\mathbf{M}$ is approximated by a diagonal matrix (see Appendix A). While this approximation increases the spatial discretization error, we did not observe any significant differences in the solution when compared to that obtained with the un-approximated mass matrix. With mass lumping, applying $\mathbf{D}^{-1}$ to a vector simplifies to m products with a tri-diagonal matrix, and $m+1$ products with diagonal matrices. Applying $\mathbf{R}^{-1}$ then requires 4 additional products with the diagonal matrices $\boldsymbol{\Gamma}^{-1}$ and $\boldsymbol{\Sigma}^{-1}$.

For the specific case of 1D observations from nadir altimeter data, $\boldsymbol{\kappa}$ becomes a 1×1 tensor, and thus a scalar diffusion coefficient. The diffusion coefficient controls the spatial extent of the along-track correlations. While it is straightforward to make the diffusion coefficient spatially varying, for the experiments in this article we used a constant diffusion coefficient denoted by the scalar κ. The diffusion coefficient and the number of diffusion iterations (m) are parameters that can be used to fit the correlation model to actual estimates of observation-error correlations. The effect of these parameters on determining the shape and spectral characteristics of the underlying correlation function is discussed in the next section.

3.2.2 | Specification of the diffusion-model parameters

In the Euclidean space $\mathbb{R}$, the smoothing kernel of the implicit diffusion operator with a constant diffusion coefficient is a covariance function from the Matérn class (Guttorp and Gneiting, 2006; Weaver and Mirouze, 2013). Specifically, with m implicit diffusion iterations, the covariance function is an Auto-Regressive (AR) function of order m; *i.e.*, a function characterised by a polynomial of order $m-1$ times the exponential function (*e.g.*, see Mirouze and Weaver (2010)). This analytical result provides the basis for controlling and interpreting the spatial and spectral smoothness properties of the implicit diffusion-based correlation model.

The number of iterations m controls the smoothness properties of the associated correlation function, as illus-

trated in Figure 3. With m large (values beyond about 10), the correlation function is approximately Gaussian (brown curve in Figure 3(a)). Small values of m correspond to correlation functions with a sharper (slower) decay at short (long) range. This is most clearly visible from the blue curve in Figure 3(a), which corresponds to $m = 1$. For larger values of m, the difference between the curves is difficult to distinguish in physical space, but the effect of m is clearer in spectral space (Figure 3(b)). In spectral space, m is seen to control the steepness of the slope (or roll-off) of the spectrum at small spatial scales, with large values of m (quasi-Gaussian functions) providing steepest roll-off. Background-error correlation functions are often specified with large values of m; *e.g.*, m is set to 10 in the diffusion-based correlation models used for **B** in the operational implementations of NEMOVAR at ECMWF and Met Office. One of the reasons motivating this choice is that the regularity of quasi-Gaussian functions allows them to be differentiated, which is typically required in the specification of multivariate cross-covariances (Daley, 1991; Derber and Bouttier, 1999; Weaver et al., 2005). The constraint of regularity is less stringent for observation-error correlations. In fact, diagnosed observation-error correlations are often quite sharp near their peak (Michel, 2018; Waller et al., 2016a,b), which is a feature that AR functions can capture with small values of m.

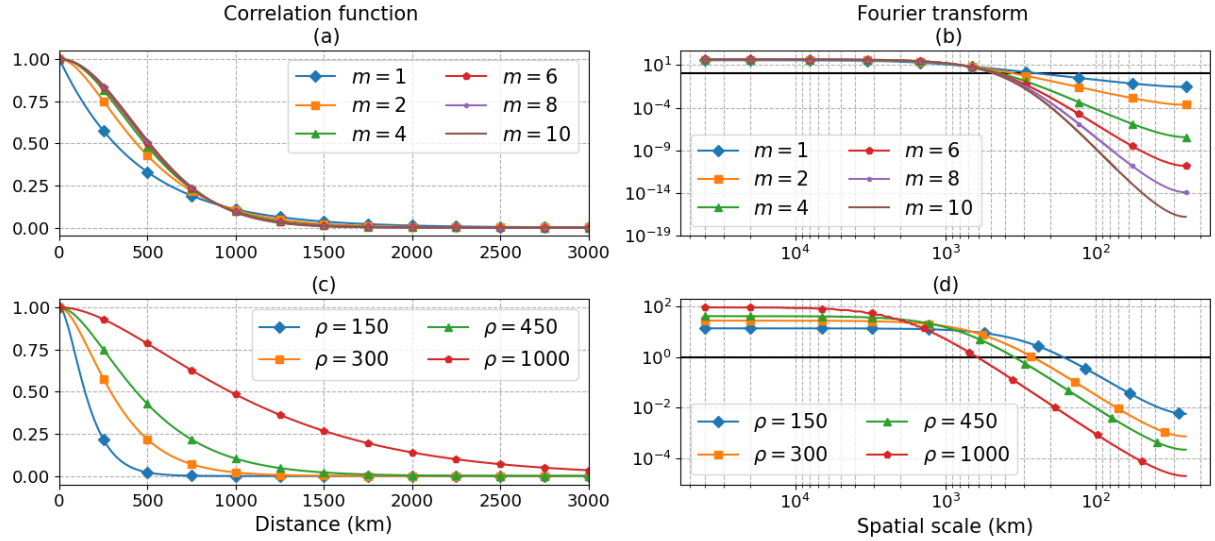


FIGURE 3 AR correlation functions (panels (a) and (c)) and their respective Fourier transforms (panels (b) and (d)). In panels (a) and(b), ρ is fixed to a value of 450 km and curves are displayed for different values of values of m. In panels (c) and (d), m is fixed to a value of 2 and curves are displayed for different values of ρ.

In practice, the diffusion coefficient is specified indirectly through a more convenient length-scale parameter. Different definitions of correlation length scale have been proposed in the literature. A commonly used definition for background-error correlations is the Daley length scale (Daley, 1991; Pannekoucke et al., 2008; Weaver and Mirouze, 2013), which is related to the local curvature of the correlation function at zero separation. However, the Daley length scale is not defined for non-differentiable correlation functions, which we do not want to exclude for modelling observation-error correlations. In particular, the exponential function, which corresponds to $m = 1$ in the AR family, is not differentiable at zero separation. As in Goux et al. (2024), we use the following length-scale parameter from Lindgren et al. (2011):

$$\rho = \sqrt{\kappa(2m-1)}, \tag{27}$$

which is very similar to the Daley length scale except for small values of m (see Goux et al. (2024) for a discussion). A property of this length scale is that, at a distance ρ from the AR function peak, the correlation is approximately 0.1 for all values of m. The AR correlation functions displayed in Figure 3(c) illustrate how increasing ρ increases the spatial range of the correlations. In spectral space (Figure 3(d)), ρ affects the spatial scale at which the spectrum starts decaying at small spatial scales, but, unlike m, does not change the decay rate itself.

Note that the correlation length scales specified in Balmaseda et al. (2013) for background error are based on the Daley definition, not Equation (27). In Section 4, we will use Equation (27) for both observation and background error. For $m = 10$, the length scale computed from Equation (27) is about 5% larger than the Daley length scale (Goux et al., 2024). In Section 4, we will use subscripts o and b on ρ and m to distinguish the observation- and background-error correlation parameters, respectively.

3.2.3 | Normalisation factors

The normalisation factors in $\mathbf{\Gamma}$ ensure that the specified standard deviations of the covariance matrix are not contaminated by amplitude errors in the representation of the correlation functions. The normalisation factors should be specified as the inverse of the square root of the diagonal elements of the diffusion-based covariance matrix $\mathbf{D}$. Therefore, to define the normalisation factors, we require a method to estimate the diagonal of $\mathbf{D}$. Different procedures have been proposed to estimate normalisation factors for diffusion-based formulations of $\mathbf{B}$ (see Weaver et al. (2021) for a review) and these procedures are generally applicable to diffusion-based formulations of $\mathbf{R}$ as well. For the nadir altimeter network, the normalisation factors can be computed exactly using an efficient implementation of the brute-force procedure, as described below. If the observation network changes from one assimilation cycle to the next, then the normalisation factors have to be recomputed on each cycle.

The brute-force procedure requires a method to apply $\mathbf{D}$ or a 'square-root' factor of $\mathbf{D}$ to a canonical vector (see Equations (18) and (19) in Weaver et al. (2021).) Recall that only $\mathbf{D}^{-1}$ is needed for the minimization algorithm (RBCG in our case). In order to apply $\mathbf{D}$, we need a computationally efficient method to apply $(\mathbf{M}+\mathbf{A})^{-1}$ in Equation (24). Drawing on experience from the procedure used in $\mathbf{B}$ in NEMOVAR (Weaver et al., 2016), we use the Chebyshev Iteration (CI) to define an approximate inverse. Exploiting the fact that the mass matrix is diagonal with mass lumping, which we denote by the matrix $\overline{\mathbf{M}}$, and hence can be easily factored into square-root components $\overline{\mathbf{M}}^{1/2}\overline{\mathbf{M}}^{1/2}$, we first reformulate $\mathbf{D}$ in Equation (24) as

$$\mathbf{D} = \overline{\mathbf{M}}^{-1/2}\left[\left(\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}\right)^{-1}\right]^{m}\overline{\mathbf{M}}^{-1/2}. \tag{28}$$

CI is used to solve the linear system involving the non-dimensional matrix $\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$. From Equation (28), this system must be solved m times, in sequence, where the solution of the linear system at a given iteration in the sequence is used as the right-hand side and first guess for the next linear system in the sequence. Since each system is solved only approximately, exact symmetry of the resulting (approximate) $\mathbf{D}$ is not guaranteed. To avoid possible symmetry errors, we restrict m to be an even number and factor Equation (28) as

$$\mathbf{D} = \overline{\mathbf{M}}^{-1/2}\left[\left(\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}\right)^{-1}\right]^{m/2}\left[\left(\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}\right)^{-1}\right]^{m/2}\overline{\mathbf{M}}^{-1/2}. \tag{29}$$

Then, we use CI with a fixed number of iterations to solve the $m/2$ linear systems in one of the factors and the adjoint of the CI algorithm with the same number of iterations to solve the $m/2$ linear systems in the other factor. This

procedure preserves symmetry even in the presence of substantial truncation errors (see Section 5.2 in Weaver et al. (2016) and Section 3.5 in Weaver et al. (2018) for a more in-depth discussion).

CI requires as input an estimate of the extreme eigenvalues of the non-dimensional matrix $\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$. Following Weaver et al. (2016), we estimate the extreme eigenvalues prior to the RBCG minimization using a combined CG-Lanczos algorithm. This procedure does not add substantial cost to the initialization step in NEMOVAR. After estimating the eigenvalues, the CI algorithm is used prior to the RBCG minimization to solve $\mathbf{Dz} = \mathbf{b}$ with a random right-hand side in order to estimate the number of iterations required to reach a target tolerance of 10^{-2} for the residual 2-norm divided by the 2-norm of the right-hand side. Although this tolerance is larger than the tolerance of 10^{-3} suggested by Weaver et al. (2016), it is sufficient to achieve an average relative error of 10^{-4} for the normalisation factors. This procedure determines the number of iterations that are fixed when applying $\mathbf{D}$ as explained above. In addition to ensuring symmetry, this procedure removes the need for estimating the norm of the residual on each iteration, and thus the need for global communications between different processors.

The brute-force approach requires the application of $\mathbf{D}$ to a canonical vector that targets each of the observation locations. Although this method is generally not feasible for $\mathbf{B}$ due to the large number of control variable grid points, it can be made efficient for the nadir altimeter network. Since observation points separated by a distance that is large compared to the correlation length scale do not interact significantly, we can estimate multiple normalisation factors simultaneously by applying $\mathbf{D}$ to a vector with entries equal to 0 everywhere except at regularly spaced locations separated by (at least) r times the correlation length scale where the entry is set to 1. In the case of nadir altimeters, empirical results indicate that a value of $r = 5$ is sufficient to obtain an average relative error in the normalisation factors of 10^{-4}, which is an order of magnitude smaller than the error obtained with the standard method used for $\mathbf{B}$, which involves randomization with 10000 samples. With an observation-error correlation length scale of 125 km, $r = 5$, and altimeter observations separated by 20 km on average, this method requires only 32 applications of $\mathbf{D}$ (compared to 10000 with randomization).

3.3 | Synthetic data assimilation experiments

As described in this section, we use an experimental framework involving synthetic data to evaluate the impact of different representations of $\mathbf{R}$. By combining Equations (2) and (5), we can write the innovation vector as

$$\mathbf{d} = \mathcal{G}(\mathbf{x}_\mathrm{t}) + \boldsymbol{\epsilon}_\mathrm{o} - \mathcal{G}(\mathbf{x}_\mathrm{b})\,, \tag{30}$$

where, from Equation (1),

$$\mathcal{G}(\mathbf{x}_\mathrm{t}) = \mathcal{G}(\mathbf{x}_\mathrm{b} - \boldsymbol{\epsilon}_\mathrm{b}) \simeq \mathcal{G}(\mathbf{x}_\mathrm{b}) - \mathbf{G}\boldsymbol{\epsilon}_\mathrm{b}, \tag{31}$$

$\mathbf{G}$ being the Jacobian matrix defined in Equation (4). Substituting Equation (31) in Equation (30) yields

$$\mathbf{d} \simeq \boldsymbol{\epsilon}_\mathrm{o} - \mathbf{G}\boldsymbol{\epsilon}_\mathrm{b}, \tag{32}$$

which shows that, in the linear approximation, the innovation vector can be expressed entirely in terms of the observation and background error.

In the synthetic data assimilation experiments, we use samples of background and observation error to simulate the innovation vector via Equation (32). To generate error samples with statistics consistent with $\mathbf{B}$ and $\mathbf{R}$, we make

use of 'square-root' factors $\mathbf{U}$ and $\mathbf{V}$ such that

$$\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathsf{T}} \qquad \text{and} \qquad \mathbf{R} = \mathbf{V}\mathbf{V}^{\mathsf{T}}.$$

In NEMOVAR, both $\mathbf{B}$ and $\mathbf{R}$ are constructed directly in this way, using a product of operators. For example, the factor $\mathbf{V}$ can be easily identified from Equations (22), (23) and (29). Given $\mathbf{U}$ and $\mathbf{V}$, we can generate an unlimited number of error samples, $\ell = 1, 2, \ldots$, from the equations

$$\boldsymbol{\epsilon}_{\mathrm{o}}^{(\ell)} = \mathbf{V}\boldsymbol{\eta}_{\mathrm{o}}^{(\ell)} \qquad \text{and} \qquad \boldsymbol{\epsilon}_{\mathrm{b}}^{(\ell)} = \mathbf{U}\boldsymbol{\eta}_{\mathrm{b}}^{(\ell)},$$

where $\boldsymbol{\eta}_{\mathrm{o}}^{(\ell)} \in \mathbb{R}^{p}$ and $\boldsymbol{\eta}_{\mathrm{b}}^{(\ell)} \in \mathbb{R}^{n}$ are random vectors containing uncorrelated, Gaussian fields with mean equal to zero and standard deviation equal to one.

In all of our experiments, the background-error covariance matrix used to generate the error samples is the same as the one that is specified in the data assimilation algorithm. It has been described in Section 3.1. However, for the observation-error covariance matrix, we make a distinction between the 'true' observation-error covariance matrix $\mathbf{R}$, which is used to generate the error samples, and the specified observation-error covariance matrix $\widetilde{\mathbf{R}}$, which is used in the data assimilation algorithm (where we have used notation in alignment with that of Section 2.2). Both $\mathbf{R}$ and $\widetilde{\mathbf{R}}$ are defined according to the same observation-location mesh, but are specified with potentially different covariance parameters (*i.e.*, $\rho_{\mathrm{o}} \neq \widetilde{\rho}_{\mathrm{o}}$, $m_{\mathrm{o}} \neq \widetilde{m}_{\mathrm{o}}$, $\sigma_{\mathrm{o}} \neq \widetilde{\sigma}_{\mathrm{o}}$).

As we use synthetic data, we can compute the actual analysis error when using different specifications of $\mathbf{R}$ and $\widetilde{\mathbf{R}}$. The analysis error can be expressed in terms of the analysis increment and background error as

$$\boldsymbol{\epsilon}_{\mathrm{a}} \;=\; \mathbf{x}_{\mathrm{a}} - \mathbf{x}_{\mathrm{t}} \;=\; \mathbf{x}_{\mathrm{a}} - \mathbf{x}_{\mathrm{b}} + \mathbf{x}_{\mathrm{b}} - \mathbf{x}_{\mathrm{t}} \;=\; \delta\mathbf{x}_{\mathrm{a}} + \boldsymbol{\epsilon}_{\mathrm{b}}. \tag{33}$$

As shown in Goux et al. (2024), different specifications of $\mathbf{R}$ and $\widetilde{\mathbf{R}}$ can have a strong influence on the convergence of a $\mathbf{B}$-preconditioned CG minimization algorithm, like RBCG. Therefore, as a robust performance measure, we use Equation (33) to track the evolution of the analysis error as a function of the iteration counter using the estimate of the increment on each iteration. At each iteration of RBCG, the Root Mean Square (RMS) of the analysis error is computed for each state variable to assess both the convergence rate of the minimization and the accuracy of the solution. The convergence criterion for all experiments is defined by a 10-order of magnitude reduction in the $\mathbf{B}$-norm of the gradient relative to the $\mathbf{B}$-norm of the initial gradient.

4 | NUMERICAL RESULTS

We use both the theoretical framework of Section 2.2 and the experimental framework described in Section 3 to interpret results for two general situations that involve different observation-error correlation length scales. First, in Section 4.1, we consider a situation where the observation-error correlations have length scales that are short compared to those of the background. This is the expected situation with nadir altimeter data (*e.g.*, Storto et al., 2019). Second, in Section 4.2, we consider a situation where the observation-error correlations have length scales that are large compared to those of the background. This situation is to be expected with wide-swath altimeter data such as SWOT (Surface Water Ocean Topography) due to observation errors being significantly correlated at distances up to 1000 km (Esteban-Fernandez and Peral, 2018). In Section 4.3, we show how the presence of different physical

variables, which was not accounted for in the theoretical framework, is one of the factors limiting the effectiveness of variance inflation compared to the use of a non-diagonal $\widetilde{\mathbf{R}}$. In both situations, we choose the correlation model smoothness parameters such that $m_\mathrm{o} < m_\mathrm{b}$. This choice is motivated by results from Goux et al. (2024) who showed an extreme sensitivity of the condition number and, as a consequence, the minimization convergence when $m_\mathrm{o} \geq m_\mathrm{b}$. The sensitivity to m_o is discussed in Section 4.4.

4.1 | Case 1. Short observation-error correlation length scales: $\rho_\mathrm{o} < \rho_\mathrm{b}$

4.1.1 | Expectations from theory

Figure 4 summarizes the effect of neglecting observation-error correlations in the simplified framework presented in Section 2.2. Here, the background-error correlation function is approximately Gaussian as in NEMOVAR ($m_\mathrm{b} = 10$). The background-error correlation length scale is $\rho_\mathrm{b} = 210$ km. The (true) correlation length scale of the observation error is $\rho_\mathrm{o} = 150$ km. The observation-error correlation function is a Second Order Auto Regressive (SOAR) function (*i.e.*, $m_\mathrm{o} = 2$ in the implicit diffusion model). The scale-dependent sensitivity and analysis-error variance that would be obtained with optimal parameters (as in Figure 1) are displayed as a reference. In addition, Figure 4 displays the sub-optimal (practical) scale-dependent sensitivity (bottom panels) and analysis-error variance (top panels) obtained when using a diagonal $\widetilde{\mathbf{R}}$ in the gain matrix. In Figures 4(a) and (b), the correlations are simply ignored ($\widetilde{\mathbf{R}} = \sigma_\mathrm{o}^2\mathbf{I}$). In this case, the observation-error variance is underestimated at large spatial scales and overestimated at small spatial scales (orange curve in Figure 4(a)). As a consequence, observations are overfit at large spatial scales and underfit at small spatial scales. Overfitting at large spatial scales translates into an excessively large sensitivity (left portion of Figure 4(b)) and a significant increase of the analysis-error variance at large spatial scales ($\widetilde{\sigma}_\mathrm{a}^2\widetilde{\lambda}_\mathrm{a}^{(k)} > \sigma_\mathrm{a}^2\lambda_\mathrm{a}^{(k)}$; left portion of Figure 4(a)).

At small spatial scales, $\widetilde{\sigma}_\mathrm{o}^2\widetilde{\lambda}_\mathrm{o}^{(k)}$ is overestimated. The sub-optimal sensitivity thus becomes smaller than the optimal sensitivity at smaller spatial scales, but only at the smallest scales where both are close to zero (they overlap in the figure). At the smallest spatial scales, $\sigma_\mathrm{b}^2\lambda_\mathrm{b}^{(k)}$ is much smaller than both $\widetilde{\sigma}_\mathrm{o}^2\widetilde{\lambda}_\mathrm{o}^{(k)}$ and $\sigma_\mathrm{o}^2\lambda_\mathrm{o}^{(k)}$ (right portion of Figure 4(a)), and there is thus no significant increase of $\widetilde{\sigma}_\mathrm{a}^2\widetilde{\lambda}_\mathrm{a}^{(k)}$ compared to the optimal case: in both the optimal and sub-optimal cases, the innovations (and hence the observations) are effectively discarded.

Overfitting observations at large spatial scales has a stronger impact than underfitting observations at small spatial scales on the analysis-error variance. This characteristic of the analysis justifies the common approach adopted in operational systems of inflating the observation-error variance; *i.e.*, by setting $\widetilde{\mathbf{R}} = \widetilde{\sigma}_\mathrm{o}^2\mathbf{I} = (\alpha\sigma_\mathrm{o})^2\mathbf{I}$ with $\alpha > 1$. Although the choice of α is usually tuned empirically in operational configurations, in our idealized system, we can determine an optimal value of α that results, on average, in the minimum total analysis-error variance[4]. With the optimal α (2.5 for this example), Figure 4(c) and (d) show the best possible results that can be obtained with a diagonal $\widetilde{\mathbf{R}}$. Increasing α increases $\widetilde{\sigma}_\mathrm{o}\widetilde{\lambda}_\mathrm{o}^{(k)}$ at all spatial scales and thus decreases $\widetilde{\lambda}_\mathrm{s}^{(k)}$ at all spatial scales (Figure 4(d)). This approach can thus mitigate overfitting observations at large spatial scales and thus bring the sub-optimal analysis-error variance closer to the optimal one at those scales (left portion of Figure 4(c)), but at the expense of worsening the fit at the small spatial scales. The optimal α is not large enough to increase significantly the impact of underfitting at small spatial scales on the analysis error (right portion of Figure 4(c)).

In summary, if $\widetilde{\mathbf{R}}$ is diagonal, variance inflation is advantageous for the case $\rho_\mathrm{o} < \rho_\mathrm{b}$ as it greatly reduces the analysis-error variance at large spatial scales by reducing the fit to the observations. Importantly, it does so without substantially increasing the analysis-error variance at small spatial scales. Although inflation further reduces the fit to

[4] Here, the average is computed from a sample of 100 analysis increments using innovation vectors generated randomly based on Equation (32).

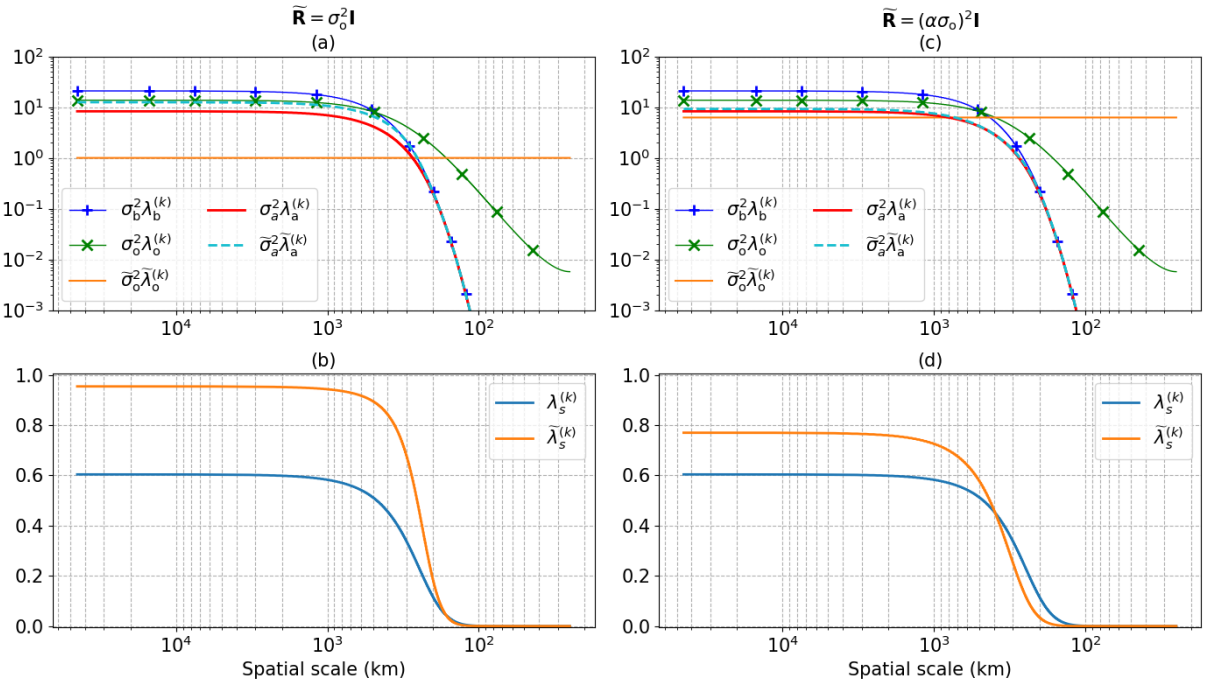


FIGURE 4 Panels (a) and (c): scale-dependent variance of the background error ($\sigma_b^2\lambda_b^{(k)}$), observation error ($\sigma_o^2\lambda_o^{(k)}$) and its misspecified counterpart ($\widetilde{\sigma}_o^2\widetilde{\lambda}_o^{(k)}$), and analysis error ($\sigma_a^2\lambda_a^{(k)}$; Equation (15)) and its sub-optimal counterpart ($\widetilde{\sigma}_a^2\widetilde{\lambda}_a^{(k)}$; Equation (20)) as a function of spatial scale. Panels (b) and (d): scale-dependent sensitivity ($\lambda_s^{(k)}$; Equation (11)) and its sub-optimal counterpart ($\widetilde{\lambda}_s^{(k)}$; Equation (17)) as a function of spatial scale. The experiment is as in Figure 1. The background-error correlation function is approximately Gaussian with a length scale $\rho_b = 210$ km; the 'true' observation-error correlation function is a SOAR function with a length scale $\rho_o = 150$ km; the misspecified observation-error covariance function has no correlation, *i.e.*, $\widetilde{\mathbf{R}} = \sigma_o^2\mathbf{I}$ in panels (a) and (b), while $\widetilde{\mathbf{R}} = \alpha^2\sigma_o^2\mathbf{I}$ in panels (c) and (d) where $\alpha = 2.5$ is an optimally derived inflation factor that minimizes the analysis-error variance.

the observations at the small spatial scales, this has little impact on the analysis error. This small impact is explained by the fact that, with or without an inflated σ_o, the sensitivity is negligible at these small spatial scales since $\sigma_o^2\lambda_o^{(k)}$ is larger than $\sigma_b^2\lambda_b^{(k)}$ by several orders of magnitude. However, variance inflation is less effective than actually accounting for observation-error correlations. With the optimal level of variance inflation, the sub-optimal sensitivity at large spatial scales remains too large compared to the optimal sensitivity. Increasing the inflation factor α further would reduce the overfitting at large spatial scales; however, in this case, the observations would be underfit at intermediate spatial scales as well as small spatial scales. Underfitting the observations at intermediate spatial scales would significantly degrade the analysis since, contrary to the small spatial scales, $\sigma_b^2\lambda_b^{(k)}$ and $\sigma_o^2\lambda_o^{(k)}$ are comparable in magnitude at these scales. Although variance inflation can mitigate the detrimental effects from neglecting correlated error, accounting for observation-error correlations is still beneficial. This is assessed experimentally in NEMOVAR in the next section.

4.1.2 | NEMOVAR experiments

We consider three experiments to determine if the conclusions deduced from the simplified system in Section 4.1.1 carry over to a realistic configuration. Correlated SSH observation errors are simulated with the implicit diffusion model with parameter settings $\rho_o = 150$ km and $m_o = 2$ to emulate the SOAR function of Section 4.1.1. For com-

parison, the smallest specified background-error correlation length scale in the meridional direction (which is approximately aligned with the satellite track) is 210 km (Balmaseda et al., 2013). The simulated innovations are then assimilated with three different specifications of $\widetilde{\mathbf{R}}$: a naive diagonal $\widetilde{\mathbf{R}}$ without variance inflation; an optimally inflated diagonal $\widetilde{\mathbf{R}}$; and an accurate non-diagonal $\widetilde{\mathbf{R}} = \mathbf{R}$. The RMS of the analysis error for each variable at each iteration (Equation (33)) is shown for the three $\widetilde{\mathbf{R}}$ specifications in panels (a), (b), (c) of Figure 5, respectively. The RMS values are computed over all model grid points. The values have been normalized by the corresponding RMS of the background error. As we observed a strong variability in the RMS of the analysis error with different random realisations of the innovation vector, we compute the average RMS from a sample of analysis errors from 100 experiments. The solid lines in Figure 5 indicate the mean of the sample. The area between the 25% and 75% quantiles of each sample is filled in to give a qualitative indication of the spread of the sample.

If observation-error correlations are not accounted for ($\widetilde{\mathbf{R}} = \sigma_o^2\mathbf{I}$), we expect a significant degradation of the analysis error due to the overfit of the observations at large spatial scales. This degradation appears very clearly in Figure 5(a), where the analysis error is larger than the background error for all variables. The analysis error is particularly large for SSH. Figure 5(b) shows that, as predicted in Section 4.1.1, variance inflation can successfully mitigate the overfitting of the observations at large spatial scales and results in an improved analysis for all variables. The minimization convergence is also accelerated by variance inflation, requiring 15 iterations compared to 31 iterations without inflation. As expected, increasing $\widetilde{\sigma}_o$ decreases the condition number of the ($\mathbf{B}$-preconditioned) Hessian matrix of the cost function (Tabeart et al., 2021; Goux et al., 2024), in this case from 9.5 to 2.4. The relatively low values of the condition number are associated with the coarseness of the grid in this configuration. The sensitivity of each variable to variance inflation is discussed in more detail in Section 4.3.

Finally, accounting for observation-error correlations with an accurate non-diagonal $\widetilde{\mathbf{R}} = \mathbf{R}$ (Figure 5(c)) leads to an even lower analysis error than with optimal inflation. Furthermore, this improvement occurs without a substantial degradation of the convergence rate, as the condition number is similar with optimal variance inflation and with a non-diagonal $\widetilde{\mathbf{R}}$ (2.4 and 2.5, respectively). This is consistent with theory where we expect the condition number of the Hessian matrix to remain similar or to be reduced when observation-error correlation length scales are short compared to background-error correlation length scales (Goux et al., 2024). The improved analysis is due to the absence of overfitting the observations at large spatial scales and underfitting the observations at small spatial scales, which occurs when using a diagonal $\widetilde{\mathbf{R}}$. The improvement at small spatial scales is highlighted by comparing the response of the different variables to the use of a non-diagonal $\widetilde{\mathbf{R}}$.

The variable most improved by the assimilation of altimeter data is SSH as it corresponds directly to the measured quantity. The other variables also see an improvement due to the balance relations. Temperature and salinity are related to SSH via a T-S relation, linearised equation of state and dynamic height relation (Weaver et al., 2005), although the effect on salinity is much smaller due to the smaller contribution of salinity to the expansion/compression of the water column. At the surface, the analysis increment for the zonal velocity component (δU) and meridional velocity component (δV) are related to the horizontal gradient of the SSH increment ($\delta\eta$) through the geostrophic balance relation (Weaver et al., 2005):

$$\begin{aligned} \delta U &= -\frac{g}{f}\frac{1}{a}\frac{\partial \delta\eta}{\partial \phi}, \\ \delta V &= \frac{g}{f}\frac{1}{a\cos\phi}\frac{\partial \delta\eta}{\partial \lambda} \end{aligned} \tag{34}$$

where g denotes standard gravity, f is the Coriolis parameter, a is the radius of the Earth, ϕ is latitude and λ is longitude. The velocity components are thus more influenced by the *gradient* of SSH than SSH itself. We can then expect them to

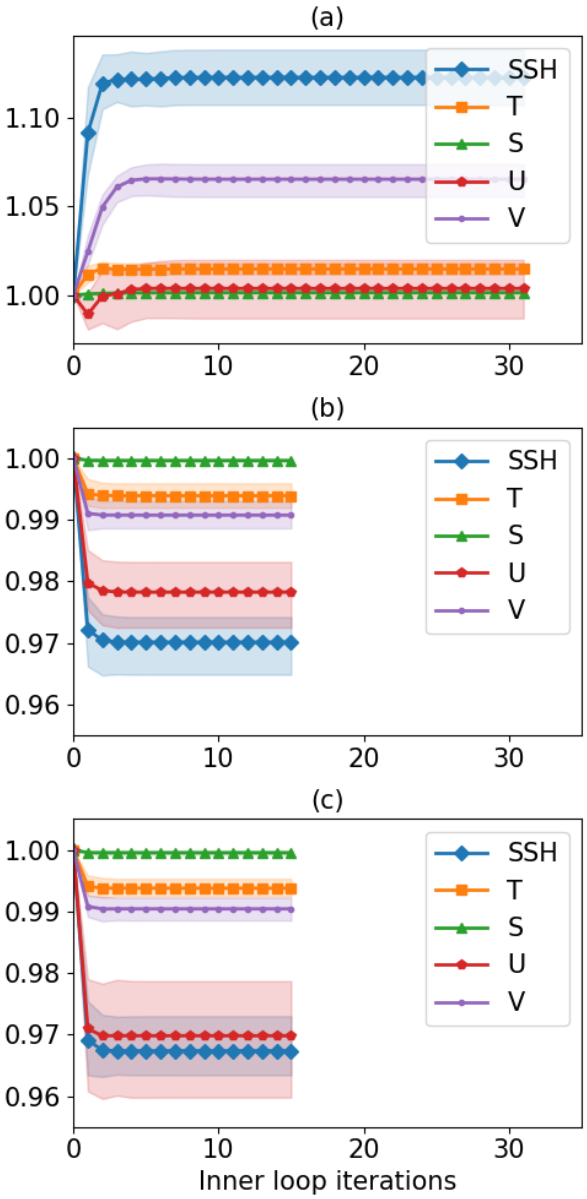


FIGURE 5 RMS of the analysis error at each iteration of the inner loop, for each variable (SSH: sea surface height, T: temperature, S: salinity, U: zonal current velocity, V: meridional current velocity). The values have been normalised by the RMS of the background error at each iteration. The background-error correlation function is approximately Gaussian ($m_b = 10$) where the length scale ρ_b varies between 210 km and 420 km. The true observation-error correlation function is a SOAR function ($m_o = 2$) with a length scale $\rho_o = 150$ km. Solid lines correspond to the mean of 100 samples of the RMS of the analysis error derived from different realisations of the random innovation vector (Equation (32)). Filled-in areas indicate the spread between the 25% and 75% quantiles of the samples. (a) $\widetilde{\mathbf{R}}$ is specified as a diagonal matrix; (b) $\widetilde{\mathbf{R}}$ is specified as a diagonal matrix with optimal inflation; and (c) $\widetilde{\mathbf{R}} = \mathbf{R}$, the true observation-error covariance matrix. For panel (b), the optimal inflation factor is $\alpha = 2.5$.

improve especially when the small spatial scales of the SSH analysis are improved, which is the case when accounting for observation-error correlations (Seaman, 1977). The improvement in the velocity components is indeed striking

when comparing panels (b) and (c) of Figure 5. Although the zonal velocity is not the observed variable, the reduction in its RMS analysis error when using a non-diagonal $\widetilde{\mathbf{R}}$ instead of optimal inflation is more significant than that of SSH. Since the nadir satellite tracks are almost aligned with the meridional axis, the SSH meridional derivative ($\partial\delta\eta/\partial\phi$) is better sampled than the zonal derivative ($\partial\delta\eta/\partial\lambda$). Furthermore, as the observation errors are correlated along track, and that these are correctly accounted for, information about the short spatial scales (gradient) is more efficiently extracted from the observations in the meridional direction. This improvement in the retrieval of the meridional derivative of SSH translates into an improvement in the retrieval of zonal velocity as their respective increments are proportional via the geostrophic relation (Equation (34)). As the zonal derivative of SSH is almost orthogonal to the satellite track, these observations provide less information on the meridional component of the geostrophic velocity.

In summary, if observation-error correlations have short length scales compared to those of background-error correlations (here, $\rho_\mathrm{o} = 150$ km, while ρ_b ranges from 210 km to 420 km in the meridional direction), not accounting for them can lead to a negative impact of the observations on the analysis due to, in particular, overfitting the observations at large spatial scales. In this situation, variance inflation can successfully reduce the overfitting at large spatial scales while simultaneously improving the convergence of the minimization. However, improvement at large spatial scales occurs at the expense of neglecting gradient information in the observations to improve the small spatial scales. On the other hand, using a non-diagonal $\widetilde{\mathbf{R}}$ allows the small-scale information to be extracted from the observations without overfitting the observations at large spatial scales. Furthermore, this improvement can be achieved without degrading the convergence rate of the minimization.

4.2 | Case 2. Large observation-error correlation length scales: $\rho_\mathrm{o} > \rho_\mathrm{b}$

In the previous section, we showed that using a non-diagonal $\widetilde{\mathbf{R}}$ has a positive impact on the analysis when $\rho_\mathrm{o} < \rho_\mathrm{b}$. In this section, we consider the opposite case when $\rho_\mathrm{o} > \rho_\mathrm{b}$. The background-error correlations are specified as in the previous section (ρ_b ranges from 210 km to 420 km in the meridional direction and $m_\mathrm{b} = 10$), but now we set the observation-error correlation length scale to $\rho_\mathrm{o} = 450$ km, while keeping the same value of the smoothness parameter ($m_\mathrm{o} = 2$).

4.2.1 | Expectations from theory

In Section 4.1, the background was less accurate than the observations at large spatial scales and more accurate at small spatial scales. With a larger ρ_o, this is no longer the case as illustrated in Figure 6. At large spatial scales, $\sigma_\mathrm{o}^2\lambda_\mathrm{o}^{(k)}$ is slightly larger than $\sigma_\mathrm{b}^2\lambda_\mathrm{b}^{(k)}$ (left portion of panels (b) and (d)) and $\lambda_\mathrm{s}^{(k)}$ is smaller than 0.4 (left portion of panels (a) and (c)), which leads to an analysis closer to the background than to the observations. However, since $\rho_\mathrm{o} > \rho_\mathrm{b}$, $\sigma_\mathrm{o}^2\lambda_\mathrm{o}^{(k)}$ begins to decay sooner than $\sigma_\mathrm{b}^2\lambda_\mathrm{b}^{(k)}$ when approaching smaller spatial scales. This leads to an increase in $\lambda_\mathrm{s}^{(k)}$ at intermediate to small spatial scales (between approximately 2000 km and 200 km in this example) where observations become predominant in the optimal analysis.

With smaller ρ_o, the shape of the optimal and suboptimal sensitivities $\lambda_\mathrm{s}^{(k)}$ and $\widetilde{\lambda}_\mathrm{s}^{(k)}$ were relatively similar as shown in Figure 4. This is not the case with larger ρ_o as can be seen in Figure 6; with a diagonal $\widetilde{\mathbf{R}}$, $\widetilde{\lambda}_\mathrm{s}^{(k)}$ always decreases monotonically when approaching smaller spatial scales, while $\lambda_\mathrm{s}^{(k)}$ (with a non-diagonal $\widetilde{\mathbf{R}}$) shows a more complex behaviour. Consequently, the optimal sensitivity cannot be approximated correctly with a diagonal $\widetilde{\mathbf{R}}$. In particular, there is no level of variance inflation that allows the analysis to extract relevant information from intermediate to small scales from the observations without significantly overfitting the observations at large spatial scales. This ineffectiveness of variance inflation in the presence of observation errors with large length scales has been observed

in practice by Reid et al. (2020) in the context of satellite-retrieved SST data assimilation.

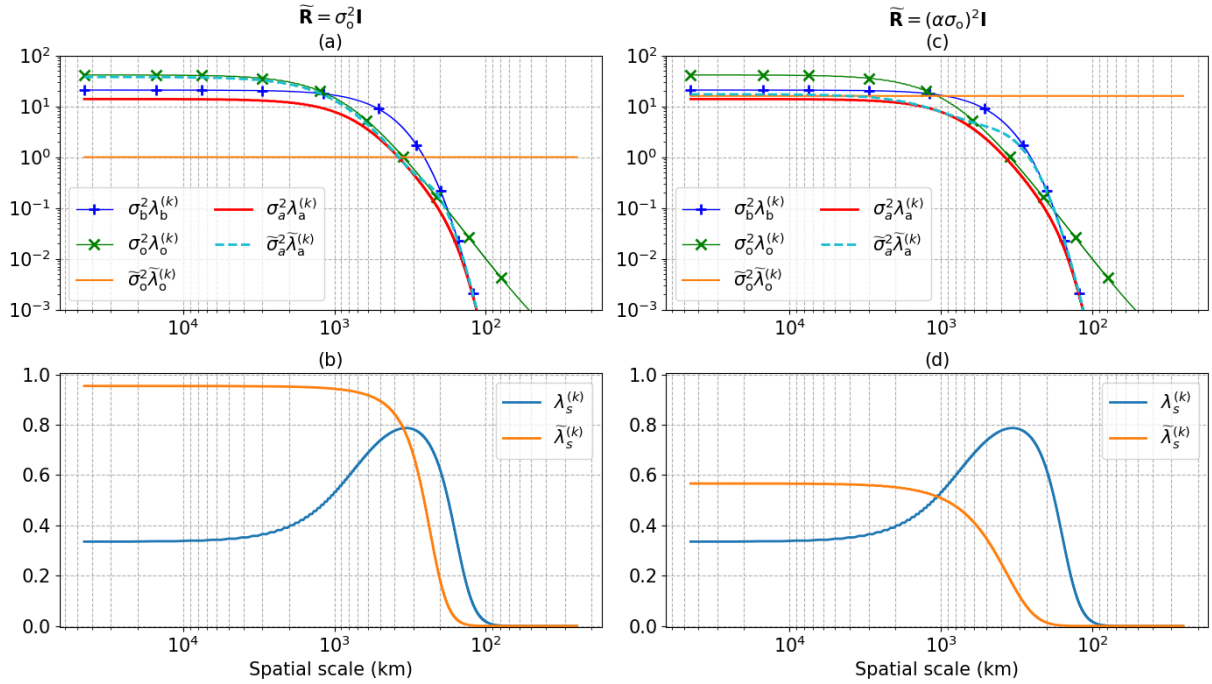


FIGURE 6 Same as for Figure 4 but with $\rho_o = 450$ km. The optimal inflation factor in this case is $\alpha = 4$.

4.2.2 | NEMOVAR experiments

As in Section 4.1.2, we consider three sets of experiments involving different specifications of $\widetilde{\mathbf{R}}$: a naive diagonal $\widetilde{\mathbf{R}}$ without variance inflation, an optimally inflated diagonal $\widetilde{\mathbf{R}}$, and an accurate non-diagonal $\widetilde{\mathbf{R}} = \mathbf{R}$. The correlated SSH observation errors are simulated with the implicit diffusion model with parameter settings $\rho_o = 450$ km and $m_o = 2$.

When the observation-error correlations are neglected, the analysis is degraded compared to the background, for all variables except zonal velocity, which achieves a 5% reduction in RMS error (Figure 7(a)). The difference between zonal velocity and the other variables can be explained by referring to panels (a) and (b) of Figure 6. The analysis overfits the observations at large spatial scales, which can explain the degradation for most variables. The scale-dependent analysis-error variance appears to be particularly sensitive to overfitting at large spatial scales. This is consistent with the theory developed in Section 2.3, which showed that overfitting the least accurate source of information (here the observations) has more impact on the analysis error than underfitting it.

However, as discussed in Section 4.1.2, the zonal velocity is influenced more by the meridional derivative of the assimilated SSH observations than the absolute value of SSH, implying that it is mainly affected by the small spatial scales in the observations. We can see in Figure 6(b) that without variance inflation, $\widetilde{\lambda}_s^{(k)}$ is relatively close to $\lambda_s^{(k)}$ at small spatial scales, and in Figure 6(a) that the increase of analysis error at these scales is small ($\sigma_a^2\lambda_a^{(k)} \simeq \widetilde{\sigma}_a\widetilde{\lambda}_a^{(k)}$). Consequently, we expect $\widetilde{\mathbf{R}} = \sigma_o^2\mathbf{I}$ to be inappropriate for most variables except for quantities such as velocity that depend on the horizontal gradient (or higher-order derivatives) of the observations. The same effect is not observed for meridional velocity, as it is related to the zonal derivative of SSH, which is almost perpendicular to the satellite track and thus less constrained by the observations.

With optimal variance inflation, overfitting at large spatial scales can be mitigated but not without preventing the extraction of small-scale information from the observations. Optimal inflation is particularly ineffective in this configuration as the small spatial scales are where the observations provide most of the information for improving the background (*i.e.*, where $\sigma_o^2 \lambda_o^{(k)} < \sigma_b^2 \lambda_b^{(k)}$). Although optimal variance inflation can prevent a negative impact of the observations on the analysis, it also negates most of their positive impact (Figure 7(b)). King and Martin (2021) reached a similar conclusion when assimilating simulated wide-swath altimeter observations with correlated error in a regional configuration. In their case, they had to apply a large amount of data thinning to compensate for correlated error when using a diagonal $\widetilde{\mathbf{R}}$. While this prevented the analysis from being degraded by the simulated observations, it also significantly reduced their impact on the analysis compared to a case where the simulated observations were taken to have uncorrelated errors.

The impact of the observations is improved by accounting for the correlated error. With an accurate non-diagonal $\widetilde{\mathbf{R}} = \mathbf{R}$, the error reduction of the SSH reaches 4% compared to 2% with optimal variance inflation (cf. Figure 7(c) and (b)). The effect on zonal velocity is even more impressive, with the relative error reduction reaching about 8%. It is interesting to note that, comparing the relative reduction of SSH and zonal velocity, the gradient of the observations provides more useful information than the observations themselves; *i.e.*, SSH observations with correlated error such that $\rho_o > \rho_b$ provide a stronger constraint on the geostrophic component of the zonal velocity than on the SSH itself. This result supports the approach to assimilate SSH observations together with their derivatives (Brankart et al., 2009; Ruggiero et al., 2016).

Using the non-diagonal $\widetilde{\mathbf{R}}$ has a slight positive impact on the convergence rate of the minimization compared to the experiments with the naive diagonal $\widetilde{\mathbf{R}}$, which uses the same observation-error variance. Although there is a risk of increasing the condition number with a non-diagonal $\mathbf{R}$ when $\rho_o > \rho_b$, it is limited by the fact $m_o < m_b$ (see Figures 7 and 8 in Goux et al. (2024)). In the case considered here, ρ_o (equal to 450 km) is not large enough compared to ρ_b (which ranges from 210 km to 420 km) for the condition number to be degraded by the non-diagonal $\widetilde{\mathbf{R}}$. The condition number is 9.2 with a non-diagonal $\widetilde{\mathbf{R}}$ compared to 9.5 with the non-inflated diagonal $\widetilde{\mathbf{R}}$. An increase in the condition number and a degraded convergence rate do occur if the same experiment is performed with $\rho_o = 800$ km (not shown). In this case the condition number is 45 and full convergence is reached in 53 iterations. On the other hand, the convergence with an optimally inflated $\widetilde{\mathbf{R}}$ is faster than with a non-diagonal $\widetilde{\mathbf{R}}$, as full convergence is reached in 11 iterations compared to 28 iterations. The improved convergence rate is explained by the larger observation-error variance in the inflated $\widetilde{\mathbf{R}}$ experiment, which decreases the condition number to 1.5. Even though reaching full convergence requires more iterations with an inflated non-diagonal $\widetilde{\mathbf{R}}$, most of the error reduction occurs in the early iterations. In fact, at each iteration, the RMS of the analysis error is lower than what can be achieved with any diagonal approximation.

4.3 | Optimal variance inflation

Optimal variance inflation allows the practical observation-error variance to achieve a better match to the true variance at large spatial scales and thus to reduce overfitting of the observations at those scales (Section 2.2). However, this can only occur at the expense of the small spatial scales where the discrepancy between practical and true observation-error variance becomes worse. In the NEMOVAR experiments, zonal velocity was shown to be more sensitive than SSH to small spatial scales. In this section, we compare the analysis errors for these two variables to evaluate their response to different levels of variance inflation.

If $\rho_o = 150$ km (*i.e.*, smaller than ρ_b everywhere in the domain), variance inflation leads to an improvement for both variables (Figure 8a). Compared to SSH, the zonal velocity seems to perform generally better with lower values

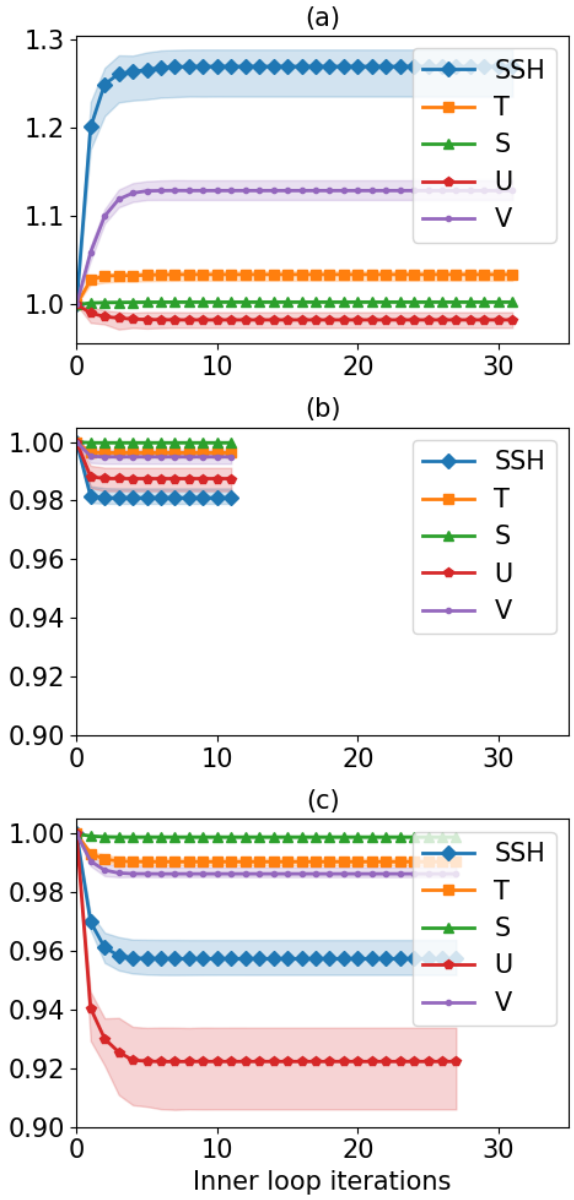


FIGURE 7 Same as for Figure 5 but with $\rho_o = 450$ km. For panel (b), the optimal inflation factor is $\alpha = 4$.

of α. This is consistent with our previous conclusion that the small spatial-scale information in the observations (*i.e.*, the meridional derivative of SSH and thus the zonal velocity through geostrophy) is best extracted with less variance inflation. For zonal velocity, there is no level of inflation with a diagonal $\widetilde{\mathbf{R}}$ for which the RMS of the analysis error is as low as that obtained with a non-diagonal $\widetilde{\mathbf{R}}$ (marked by dashed lines in Figure 8).

This is even more striking when $\rho_o = 450$ km, as seen in Figure 8(b). Even if α is set optimally for a specific variable, the analysis error of this variable would still be much larger than it would be with a non-diagonal $\widetilde{\mathbf{R}}$. Moreover, there is a large discrepancy between optimal inflation factors for SSH and zonal velocity, with $\alpha = 4$ and $\alpha = 1.5$ being the optimal values for these variables, respectively. Consequently, variance inflation forces a compromise on the accuracy

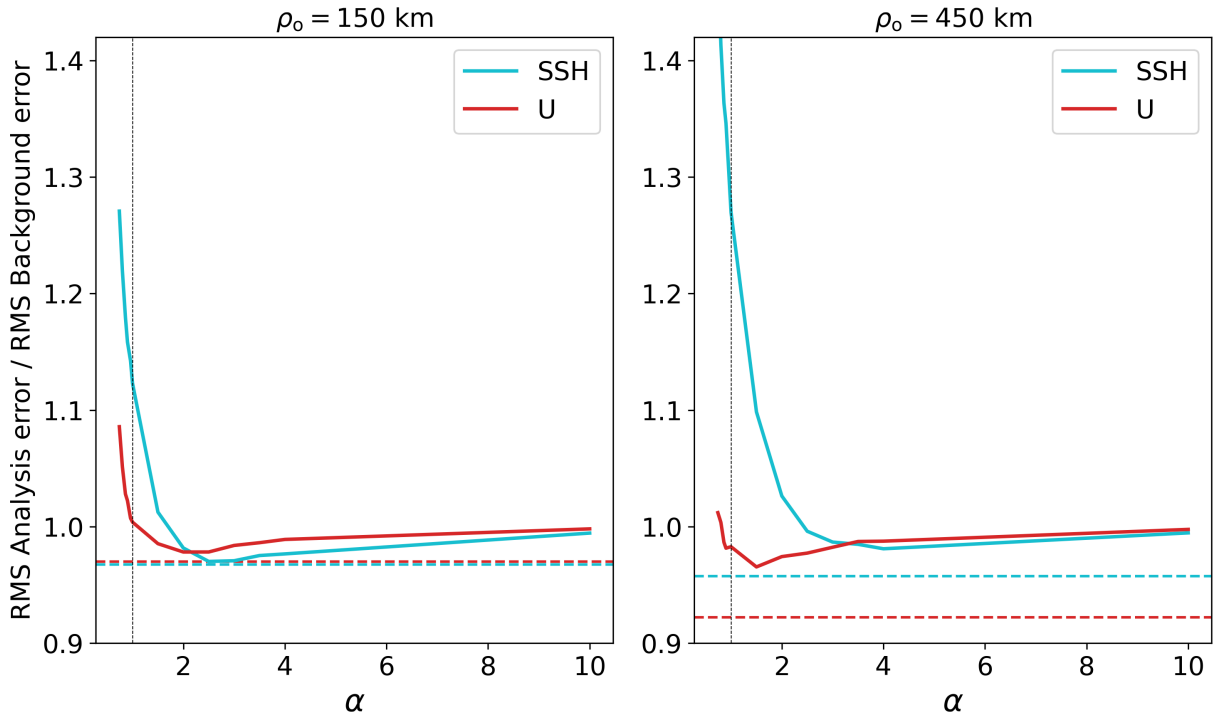


FIGURE 8 The RMS of the analysis error for SSH and zonal velocity for different values of the variance inflation factor α. The results are averaged over 100 NEMOVAR experiments and have been normalised by the RMS of the background error. The true observation-error correlation function is a SOAR function with a length scale of 150 km in panel (a) and 450 km in panel (b). The practical $\widetilde{\mathbf{R}}$ is specified as $(\alpha\sigma_o)^2\mathbf{I}$. The dotted horizontal lines indicate the values reached with a non-diagonal $\widetilde{\mathbf{R}} = \mathbf{R}$. The black vertical dashed line marks no inflation ($\alpha = 1$).

of some variables in favour of others. There is no level of variance inflation that prevents the SSH analysis from being worse than the background without negating the improvement of the zonal velocity analysis.

4.4 | Sensitivity of $\widetilde{\mathbf{R}}^{-1}$

The results from the NEMOVAR experiments illustrated the sensitivity of the analysis error and its convergence behaviour to the correlation length scale in $\widetilde{\mathbf{R}}$. Additional sensitivity experiments were conducted to evaluate the impact from adopting a diagonal $\widetilde{\mathbf{R}}$ approximation when the actual observation errors are correlated. In this section, we examine the sensitivity from a different perspective, focussing on the non-diagonal $\widetilde{\mathbf{R}}^{-1}$ and its sensitivity to noise at small spatial scales.

First, consider the linear system $\widetilde{\mathbf{R}}\mathbf{z} = \mathbf{b}$ with $\mathbf{z}, \mathbf{b} \in \mathbb{R}^p$, p being the number of observations. Let $\delta\mathbf{z}$ denote the perturbation to the solution $\mathbf{z}$ that results from a perturbation $\delta\mathbf{b}$ to the right-hand side $\mathbf{b}$; *i.e.*, $\widetilde{\mathbf{R}}(\mathbf{z} + \delta\mathbf{z}) = \mathbf{b} + \delta\mathbf{b}$ where $\delta\mathbf{z} = \widetilde{\mathbf{R}}^{-1}\delta\mathbf{b}$. The condition number of $\widetilde{\mathbf{R}}$, denoted $\kappa(\widetilde{\mathbf{R}})$, bounds the ratio of the norm of $\delta\mathbf{z}$ to the norm of $\mathbf{z}$ (Section 2.6.2 of Golub and Van Loan, 2013):

$$\frac{\|\delta\mathbf{z}\|_2}{\|\mathbf{z}\|_2} \leq \kappa(\widetilde{\mathbf{R}}) \frac{\|\delta\mathbf{b}\|_2}{\|\mathbf{b}\|_2}. \tag{35}$$

A small perturbation to $\mathbf{b}$ will not result in a large perturbation to $\mathbf{z}$ if $\kappa(\widetilde{\mathbf{R}})$ is small but could lead to a large perturbation

to $\mathbf{z}$ if $\kappa(\widetilde{\mathbf{R}})$ is large. Furthermore, in the latter case, the perturbation can be amplified by successive applications of $\widetilde{\mathbf{R}}^{-1}$ in the minimization algorithm. For example, if the experiment of Figure 7 is repeated with $\widetilde{m}_\mathrm{o} = m_\mathrm{o} = 10$ instead of $\widetilde{m}_\mathrm{o} = m_\mathrm{o} = 2$, the noise inherent in finite-precision calculations can be amplified to an extent that the analysis becomes worse than the background even though $\widetilde{\mathbf{R}} = \mathbf{R}$. The degraded performance is explained by the huge condition number of $\widetilde{\mathbf{R}}$, which is 10^9 in this case.

A large condition number does not necessarily imply that perturbations will be excessively amplified (Equation (35) is only an upper bound), but the shape of the spectrum of $\widetilde{\mathbf{R}}^{-1}$ indicates how this extreme sensitivity can occur. The spectrum of $\widetilde{\mathbf{R}}^{-1}$ is the inverse of the spectrum of $\widetilde{\mathbf{R}}$. When $\widetilde{\mathbf{R}}^{-1}$ is applied to a vector, it will amplify the small spatial scales of that vector by several orders of magnitude, with the smallest spatial scales amplified the most. In particular, referring to Figure 3(b), we see that, for $m_\mathrm{o} = 10$ and $\rho_\mathrm{o} = 450$ km, a condition number of 10^9 would result in $\widetilde{\mathbf{R}}^{-1}$ damping the largest spatial scale by a factor of 10 while amplifying the smallest spatial scale by a factor of 10^8. This property of Gaussian-like functions makes them and other correlation functions with a sharp decay rate at small spatial scales inappropriate for modelling $\mathbf{R}^{-1}$. As expected from Equation (17) in Goux et al. (2024), by decreasing m_o from 10 to 2, the largest eigenvalue of $\widetilde{\mathbf{R}}^{-1}$ decreases while its smallest eigenvalue is virtually unchanged, resulting in a lower condition number (541 in the NEMOVAR experiment). With $m_\mathrm{o} = 2$, the small spatial scales are no longer excessively amplified, as confirmed by the positive results presented earlier in Figure 7(c).

In a real-data application, the observation-error covariances are not perfectly known, which can lead to an additional risk. Let us assume that the true observation error is the sum of a correlated component $\boldsymbol{\epsilon}_\mathrm{corr}$ and an uncorrelated component $\boldsymbol{\epsilon}_\mathrm{uncorr}$:

$$\boldsymbol{\epsilon}_\mathrm{o} = \frac{1}{\sqrt{1+\beta^2}} \left(\boldsymbol{\epsilon}_\mathrm{corr} + \beta \boldsymbol{\epsilon}_\mathrm{uncorr}\right)$$

where β is a scalar set to 0.01 so that the uncorrelated error represents 1% of the total variance of the observation error. The normalisation factor $\sqrt{1+\beta^2}$ ensures that the total variance is equal to the variance of $\boldsymbol{\epsilon}_\mathrm{corr}$. The correlation model parameters are set as before: $\rho_\mathrm{o} = 450$ km and $m_\mathrm{o} = 2$. The specified $\widetilde{\mathbf{R}}$ ignores this additional uncorrelated component: *i.e.*, $\widetilde{\mathbf{R}} = \mathbb{E}\left[\boldsymbol{\epsilon}_\mathrm{corr}\boldsymbol{\epsilon}_\mathrm{corr}^\mathrm{T}\right]$ while $\mathbf{R} = \mathbb{E}\left[\boldsymbol{\epsilon}_\mathrm{o}\boldsymbol{\epsilon}_\mathrm{o}^\mathrm{T}\right]$. Even though the perturbation to the observation error is small compared to the total observation error, it leads to a significantly degraded analysis where all fields except for zonal velocity are worse than the background (Figure 9(a)). This negative impact sharply contrasts the positive impact seen in Figure 7(c) in the absence of this additional noise. The degradation comes from $\widetilde{\mathbf{R}}^{-1}$ excessively amplifying the errors at small spatial scales where the actual observation errors contain slightly more power than accounted for in $\widetilde{\mathbf{R}}$ due to the neglected uncorrelated component.

A pragmatic solution to reduce the risk of excessively amplifying the small spatial scales is to 'play it safe' by decreasing $\widetilde{\rho}_\mathrm{o}$, even if observational evidence (*e.g.*, from Desroziers et al. (2005) diagnostics) suggests otherwise, and by inflating the variances by a modest amount. For example, if $\widetilde{\rho}_\mathrm{o} = 400$ km instead of 450 km, and $\widetilde{\sigma}_\mathrm{o}$ is inflated by a factor of 2 (smaller than the factor of 4 used with the diagonal $\widetilde{\mathbf{R}}$ in Figure 7(b)), then the analysis retains a fair share of the improvements seen with an accurate non-diagonal $\widetilde{\mathbf{R}}$ while being resilient to perturbations at small spatial scales (Figure 9(b)). A clear drawback with this approach is that it requires empirical tuning and is visibly sub-optimal when comparing Figure 9(b) and Figure 7(c). A better approach would be to cap the minimum allowed value of the observation-error variances at the small spatial scales while leaving the large spatial scales unaltered. At the very small spatial scales, the background and observations are already close to the true state, so that an inaccurate specification of the observation-error variance at those scales would not have a strong impact on the analysis. For example, if $\sigma_\mathrm{o}^2\lambda_\mathrm{o}^{(k)}$ is capped at 10^{-2} in Figure 6(a) then the impact on the analysis error is minor.

Similar ideas have been explored with inter-channel observation-error correlations (Bormann et al., 2016; Geer,

2019; Weston et al., 2014), by controlling the smallest eigenvalues of an explicit eigen-decomposition of the inter-channel correlation matrix. With other modelling approaches where $\widetilde{\mathbf{R}}$ is built explicitly (*e.g.*, from a covariance function) and then inverted directly, another possibility would be to 'recondition' $\widetilde{\mathbf{R}}$ by adding a scalar multiple of the identity matrix to the matrix modelling the correlated error (Tabeart et al., 2020). Although this method also alters eigenvalues at large spatial scales, its effect is only significant at small spatial scales when the multiplicative factor is small enough. Producing the same effect with the diffusion model, while retaining a convenient form of $\widetilde{\mathbf{R}}^{-1}$, is more challenging.

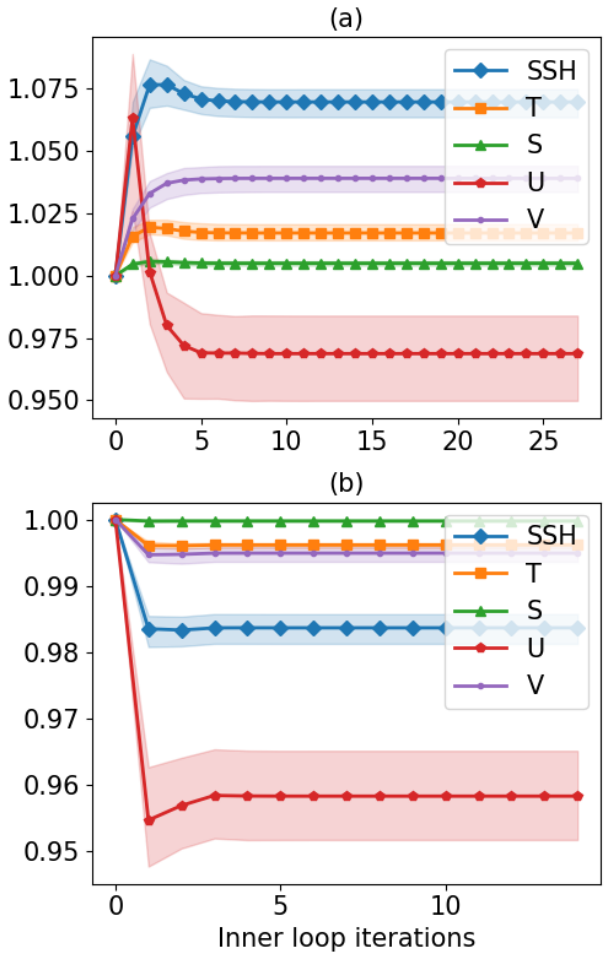


FIGURE 9 Same as Figure 7 but with the true (correlated) observation error including an additional component of uncorrelated error with variance equal to 1% of the total observation-error variance ($\sigma_o^2 = 1$). (a) $\rho_o = \widetilde{\rho}_o = 450$ km; (b) $\rho_o = 450$ km, $\widetilde{\rho}_o = 400$ km and $\widetilde{\sigma}_o = 2$ (*i.e.*, the inflation factor is $\alpha = 2$).

5 | SUMMARY AND CONCLUSIONS

The purpose of this article has been threefold. First, it provides additional insight into the impact of correlated observation error in data assimilation. Second, it addresses the specific problem of accounting for correlated error in along-track (nadir) altimeter observations in a variational data assimilation system (NEMOVAR) for the global ocean. Third, it provides the first practical demonstration of the implicit diffusion approach for modelling observation-error correlations (Guillet et al., 2019) in a realistic data assimilation system.

A key aspect of our study has been to analyse the general data assimilation problem in a simplified context using a Fourier analysis. In doing so, we have highlighted the scale-dependent characteristics of the problem and used them to interpret the results from the simulated data experiments with the ocean data assimilation system. The two key parameters in the implicit diffusion-based correlation model are the length scale ρ and the number of diffusion iterations m. These parameters control the shape of the correlation function, in particular its spatial range (ρ) and the steepness of its spectral slope at small spatial scales (m). In designing our experiments, we focussed on studying the sensitivity of the analysis errors to the specification of the basic parameters (ρ_o, m_o) in the diffusion-based correlation model of the observation-error covariance matrix (**R**). The corresponding parameters (ρ_b, m_b) used in the diffusion-based correlation model of the background-error covariance matrix (**B**) were fixed to values used in a previous operational version of NEMOVAR (Balmaseda et al., 2013). Specifically, ρ_b was spatially varying; ρ_o was constant; m_b was chosen to model an approximate Gaussian function; and, in most experiments, m_o was chosen to model a SOAR function.

The main findings from our study are as follows. We have shown that assimilating observations with correlated errors, but ignoring them in **R** in the data assimilation method, results in overfitting the observations at large spatial scales and underfitting them at small spatial scales. Variance inflation can reduce the overfitting at large spatial scales and is effective when $\rho_\mathrm{o} < \rho_\mathrm{b}$. Nevertheless, as expected, explicitly accounting for the correlations in **R** gives best results in terms of reducing the overall analysis error. The improvement is particularly noticeable in zonal velocity, which, via geostrophic balance, benefits from meridional derivative information in the along-track altimeter observations. This information about the small spatial scales is effectively extracted by the observation-error correlation model. When $\rho_\mathrm{o} > \rho_\mathrm{b}$, the derivative information in the observations becomes more important and the impact of the observation-error correlation model is even more impressive at reducing the analysis errors, particularly in zonal velocity. Variance inflation is shown to be much less effective for the case $\rho_\mathrm{o} > \rho_\mathrm{b}$ than the case $\rho_\mathrm{o} < \rho_\mathrm{b}$, as it can only prevent overfitting of observations at large spatial scales by simultaneously hindering the extraction of information at small spatial scales.

When $\rho_\mathrm{o} < \rho_\mathrm{b}$ (taken here to be roughly 70% of the minimum value of ρ_b), the minimization reaches full convergence in half the number of iterations compared to the experiment with a non-inflated diagonal **R**. In this case, the non-diagonal **R** decreases the condition number. The convergence rate and condition number are similar to those obtained in the experiment with an optimally inflated diagonal **R**. When $\rho_\mathrm{o} > \rho_\mathrm{b}$ (taken here to be roughly twice the *minimum* and slightly larger than the *maximum* value of ρ_b), the convergence rate with a non-diagonal **R** is similar to the one obtained with a non-inflated diagonal **R**, but worse than with an optimally inflated **R**. Nevertheless, our experiments show that the analysis error is always smaller with a non-diagonal **R** than with the optimally inflated diagonal **R** at any iteration of the minimisation. However, increasing ρ_o even further (to roughly twice the *maximum* value of ρ_b) significantly increases the condition number and slows down convergence.

When ρ_o is large, the condition number of **R** itself becomes large because of small eigenvalues associated with the small spatial scales. In this case, application of $\mathbf{R}^{-1}$ can potentially amplify small-scale features to extreme levels. The risk of this occurring is significantly higher with large values of m_o as demonstrated analytically by Goux et al. (2024) and confirmed by experiments in our study. For example, when $m_\mathrm{o} = 10$ (quasi-Gaussian), the condition number of **R**

becomes huge unless ρ_o is chosen to be extremely small. This extreme sensitivity to large values of m_o was the main motivation for the choice of $m_o = 2$ (SOAR) in our experiments.

In practice, the **R** that is used in the data assimilation system will always be an approximation of the true **R**. We showed that the accuracy of the analysis is particularly sensitive to errors at small spatial scales that are not fully accounted for in **R**. In this situation, we found that it was sufficient to reduce ρ_o slightly and apply a modest amount of variance inflation to retain much of the benefit to the analysis from using a non-diagonal **R** while rendering **R** more robust. A better solution would be to modify the diffusion model in a way that would allow more control over the small eigenvalues in **R**.

acknowledgements

This work has benefitted from funding from the Copernicus Climate Change Service (C3S), one of six services of the European Union's Copernicus Programme, which is implemented by ECMWF on behalf of the European Commission. It is a contribution to the contract C3S_601_INRIA: "Advancing ocean data assimilation methodology for climate applications". Additional support has been provided by the French National Programme LEFE/INSU (Les Enveloppes Fluides et l'Environnement/Institut National des Sciences de l'Univers) through the scientific themes MANU (Méthodes Mathématiques et NUmériques) and GMMC (Groupe Mission Mercator/Coriolis).

conflict of interest

This study does not have any conflicts to disclose.

Data availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

references

Balmaseda, M. A., Mogensen, K. and Weaver, A. T. (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, **139**, 1132–1161. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2063.

Bannister, R. N. (2008) A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, **134**, 1971–1996. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.339.

Bédard, J. and Buehner, M. (2019) A practical assimilation approach to extract smaller-scale information from observations with spatially correlated errors: An idealized study. *Quarterly Journal of the Royal Meteorological Society*, **146**, 468–482. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3687.

Bormann, N. and Bauer, P. (2010) Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Quarterly Journal of the Royal Meteorological Society*, **136**, 1036–1050. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.616.

Bormann, N., Bonavita, M., Dragani, R., Eresmaa, R., Matricardi, M. and McNally, A. (2016) Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Quarterly Journal of the Royal Meteorological Society*, **142**, 1767–1780. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2774.

Brankart, J.-M., Ubelmann, C., Testut, C.-E., Cosme, E., Brasseur, P. and Verron, J. (2009) Efficient parameterization of the observation error covariance matrix for square root or ensemble Kalman filters: Application to ocean altimetry. *Monthly Weather Review*, **137**, 1908–1927. URL: https://doi.org/10.1175/2008MWR2693.1.

Campbell, W. F., Satterfield, E. A., Ruston, B. and Baker, N. L. (2017) Accounting for correlated observation error in a dual-formulation 4D variational data assimilation system. *Monthly Weather Review*, **145**, 1019–1032. URL: https://doi.org/10.1175/MWR-D-16-0240.1.

Cardinali, C., Pezzulli, S. and Andersson, E. (2004) Influence-matrix diagnostic of a data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **130**, 2767–2786. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.03.205.

Chabot, V., Nodet, M., Papadakis, N. and Vidard, A. (2015) Accounting for observation errors in image data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, **67**, 23629. URL: https://a.tellusjournals.se/articles/10.3402/tellusa.v67.23629.

Chrust, M., Weaver, A. T., Browne, P., Zuo, H. and Balmaseda, M. A. (2025) Impact of ensemble-based hybrid background-error covariances in ECMWF's next generation ocean reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, **151**, e4914. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4914.

Ciarlet, P. (2002) *The Finite Element Method for Elliptic Problems*. Classics in Applied Mathematics. Philadelphia, PA: SIAM.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. (1994) A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, **120**, 1367–1387. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712051912.

Daley, R. (1991) *Atmospheric Data Analysis*. Cambridge University Press.

Derber, J. and Bouttier, F. (1999) A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A: Dynamic Meteorology and Oceanography*, **51**, 195–221. URL: https://doi.org/10.1034/j.1600-0870.1999.t01-2-00003.x.

Derber, J. and Rosati, A. (1989) A global oceanic data assimilation system. *Journal of Physical Oceanography*, **19**, 1333–1347. URL: https://doi.org/10.1175/1520-0485(1989)019%3C1333:AGODAS%3E2.0.CO;2.

Desroziers, G., Berre, L., Chapnik, B. and Poli, P. (2005) Diagnosis of observation-, background- and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, **131**, 3385–3396. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.108.

Esteban-Fernandez, D. and Peral, E. (2018) SWOT mission performance and error budget. In *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE.

Fisher, M. (2007) The sensitivity of analysis errors to the specification of background error covariances. In *Proceedings of the ECMWF Workshop on flow-dependent aspects of data assimilation, 11–13 June 2007, ECMWF.*, 27–36. URL: http://www.ecmwf.int/en/publications/.

Fletcher, S. (2022) *Data Assimilation for the Geosciences: From Theory to Application*. Elsevier Science.

Fowler, A. M., Dance, S. L. and Waller, J. A. (2018) On the interaction of observation and prior error correlations in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **144**, 48–62. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3183.

Geer, A. J. (2019) Correlated observation error models for assimilating all-sky infrared radiances. *Atmospheric Measurement Techniques*, **12**, 3629–3657. URL: https://amt.copernicus.org/articles/12/3629/2019/.

Gelb, A. (1974) *Applied Optimal Estimation*. The MIT Press: Cambridge.

Golub, G. H. and Van Loan, C. F. (2013) *Matrix Computations*. J. Hopkins Uni. Press.

Goux, O., Gürol, S., Weaver, A. T., Diouane, Y. and Guillet, O. (2024) Impact of correlated observation errors on the conditioning of variational data assimilation problems. *Numerical Linear Algebra with Applications*, **31**, e2529. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/nla.2529.

Gratton, S., Lawless, A. S. and Nichols, N. K. (2007) Approximate Gauss–Newton methods for nonlinear least squares problems. *SIAM Journal on Optimization*, **18**, 106–132. URL: https://doi.org/10.1137/050624935.

Gray, R. M. (2005) Toeplitz and circulant matrices: A review. *Foundations and Trends in Communications and Information Theory*, **2**, 155–239.

Guillet, A., Weaver, A. T., Vasseur, X., Michel, Y., Gratton, S. and Gürol, S. (2019) Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. *Quarterly Journal of the Royal Meteorological Society*, **145**, 1947–1967. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3537.

Gürol, S., Weaver, A. T., Moore, A. M., Piacentini, A., Arango, H. G. and Gratton, S. (2014) **B**-preconditioned minimization algorithms for variational data assimilation with the dual formulation. *Quarterly Journal of the Royal Meteorological Society*, **140**, 539–556. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2150.

Guttorp, P. and Gneiting, T. (2006) Studies in the history of probability and statistics XLIX on the Matérn correlation family. *Biometrika*, **93**, 989–95. URL: https://doi.org/10.1093/biomet/93.4.989.

Haben, S. A., Lawless, A. S. and Nichols, N. K. (2011) Conditioning of incremental variational data assimilation, with application to the Met Office system. *Tellus A: Dynamic Meteorology and Oceanography*, **63**, 782–792. URL: https://doi.org/10.1111/j.1600-0870.2011.00527.x.

Healy, S. B. and White, A. A. (2005) Use of discrete Fourier transforms in the 1D-Var retrieval problem. *Quarterly Journal of the Royal Meteorological Society*, **131**, 63–72. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.03.193.

Hu, G. and Dance, S. L. (2021) Efficient computation of matrix-vector products with full observation weighting matrices in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **147**, 4101–4121. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4170.

— (2024) A novel localized fast multipole method for computations with spatially correlated observation error statistics in data assimilation. *Journal of Advances in Modeling Earth Systems*, **16**, e2023MS003871. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2023MS003871.

King, R. R. and Martin, M. J. (2021) Assimilating realistically simulated wide-swath altimeter observations in a high-resolution shelf-seas forecasting system. *Ocean Science*, **17**, 1791–1813. URL: https://os.copernicus.org/articles/17/1791/2021/.

Lindgren, F., Rue, H. and Lindström, J. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**, 423–498. URL: https://doi.org/10.1111/j.1467-9868.2011.00777.x.

Madec, G. and the NEMO System Team (2023) NEMO Ocean Engine Reference Manual. *Zenodo*.

Michel, Y. (2018) Revisiting Fisher's approach to the handling of horizontal spatial correlations of observation errors in a variational framework. *Quarterly Journal of the Royal Meteorological Society*, **144**, 2011–2025. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3249.

Mignac, D., Waters, J., Lea, D. J., Martin, M. J., While, J., Weaver, A. T., Vidard, A., Guiavarc'h, C., Storkey, D., Ford, D., Blockley, E. W., Baker, J., Haines, K., Price, M. R., Bell, M. J. and Renshaw, R. (2024) Updates to the Met Office's global ocean-sea ice forecasting system including model and data assimilation changes. *EGUsphere*, **2024**, 1–31. URL: https://egusphere.copernicus.org/preprints/2024/egusphere-2024-3143/.

Mirouze, I. and Weaver, A. T. (2010) Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Quarterly Journal of the Royal Meteorological Society*, **136**, 1421–1443. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.643.

Mogensen, K., Balmaseda, M. A. and Weaver, A. T. (2012) The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. *ECMWF Technical Memorandum 668*. URL: http://www.ecmwf.int/en/publications/.

Mogensen, K., Balmaseda, M. A., Weaver, A. T., Martin, M. and Vidard, A. (2009) NEMOVAR: a variational data assimilation system for the NEMO model. *ECMWF Newsletter*, 17–20. URL: http://www.ecmwf.int/en/publications/.

Pannekoucke, O., Berre, L. and Desroziers, G. (2008) Background-error correlation length-scale estimates and their sampling statistics. *Quarterly Journal of the Royal Meteorological Society*, **134**, 497–508. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.212.

Pinnington, E. M., Casella, E., Dance, S. L., Lawless, A. S., Morison, J. I. L., Nichols, N. K., Wilkinson, M. and Quaife, T. L. (2016) Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using four-dimensional variational data assimilation. *Agricultural and Forest Meteorology*, **228-229**, 299–314. URL: https://www.sciencedirect.com/science/article/pii/S016819231630332X.

Rainwater, S., Bishop, C. H. and Campbell, W. F. (2015) The benefits of correlated observation errors for small scales. *Quarterly Journal of the Royal Meteorological Society*, **141**, 3439–3445. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2582.

Reid, R., Good, S. and Martin, M. J. (2020) Use of uncertainty inflation in OSTIA to account for correlated errors in satellite-retrieved sea surface temperature data. *Remote Sensing*, **12**, 1083. URL: https://doi.org/10.3390/rs12071083.

Ruggiero, G. A., Cosme, E., Brankart, J. M., Le Sommer, J. and Ubelmann, C. (2016) An efficient way to account for observation error correlations in the assimilation of data from the future SWOT high-resolution altimeter mission. *Journal of Atmospheric and Oceanic Technology*, **33**, 2755–2768. URL: https://doi.org/10.1175/JTECH-D-16-0048.1.

Seaman, R. S. (1977) Absolute and differential accuracy of analyses achievable with specified observational network characteristics. *Monthly Weather Review*, **105**, 1211–1222. URL: https://doi.org/10.1175/1520-0493(1977)105%3C1211:AADAOA%3E2.0.CO;2.

Stammer, D. and Cazenave, A. (2018) *Satellite Altimetry Over Oceans and Land Surfaces*. Boca Raton, FL: CRC Press.

Stewart, L. M., Dance, S. L. and Nichols, N. K. (2013) Data assimilation with correlated observation errors: experiments with a 1-D shallow water model. *Tellus A: Dynamic Meteorology and Oceanography*, **65**, 19546. URL: https://a.tellusjournals.se/articles/10.3402/tellusa.v65i0.19546.

Storto, A., Oddo, P., Cozzani, E. and Coelho, E. F. (2019) Introducing along-track error correlations for altimetry data in a regional ocean prediction system. *Journal of Atmospheric and Oceanic Technology*, **36**, 1657–1674. URL: https://doi.org/10.1175/JTECH-D-18-0213.1.

Tabeart, J. M., Dance, S. L., Haben, S. A., Lawless, A. S., Nichols, N. K. and Waller, J. A. (2018) The conditioning of least-squares problems in variational data assimilation. *Numerical Linear Algebra with Applications*, **25**, e2165. URL: https://doi.org/10.1002/nla.2165.

Tabeart, J. M., Dance, S. L., Lawless, A. S., Migliorini, S., Nichols, N. K., Smith, F. and Waller, J. A. (2020) The impact of using reconditioned correlated observation-error covariance matrices in the Met Office 1D-Var system. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1372–1390. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3741.

Tabeart, J. M., Dance, S. L., Lawless, A. S., Nichols, N. K. and Waller, J. A. (2021) New bounds on the condition number of the Hessian of the preconditioned variational data assimilation problem. *Numerical Linear Algebra with Applications*, **29**, e2405. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nla.2405.

Waller, J. A., Ballard, S., Dance, S., Kelly, G., Nichols, N. K. and Simonin, D. (2016a) Diagnosing horizontal and inter-channel observation error correlations for SEVIRI observations using observation-minus-background and observation-minus-analysis statistics. *Remote Sensing*, **8**, 581. URL: https://doi.org/10.3390/rs8070581.

Waller, J. A., Simonin, D., Dance, S. L., Nichols, N. K. and Ballard, S. P. (2016b) Diagnosing observation error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. *Monthly Weather Review*, **444**, 3534–3551. URL: https://doi.org/10.1175/MWR-D-15-0340.1.

Waters, J., Martin, M. J., Bell, M. J., King, R. R., Gaultier, L., Ubelmann, C., Donlon, C. and Van Gennip, S. (2024) Assessing the potential impact of assimilating total surface current velocities in the Met Office's global ocean forecasting system. *Frontiers in Marine Science*, **11**. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2388.

Weaver, A. T., Chrust, M., Ménétrier, B. and Piacentini, A. (2021) An evaluation of methods for normalizing diffusion-based covariance operators in variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **147**, 289–320. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3918.

Weaver, A. T. and Courtier, P. (2001) Correlation modelling on the sphere using a generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, **127**, 1815–1846. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712757518.

Weaver, A. T., Deltel, C., Machu, E., Ricci, S. and Daget, N. (2005) A multivariate balance operator for variational ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131**, 3605–3625. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.119.

Weaver, A. T., Gürol, S., Tshimanga, J., Chrust, M. and Piacentini, A. (2018) "Time"-parallel diffusion-based correlation operators. *Quarterly Journal of the Royal Meteorological Society*, **144**, 2067–2088. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3302.

Weaver, A. T. and Mirouze, I. (2013) On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, **139**, 242–260. URL: https://doi.org/10.1002/qj.1955.

Weaver, A. T., Tshimanga, J. and Piacentini, A. (2016) Correlation operators based on an implicitly formulated diffusion equation solved with the Chebyshev iteration. *Quarterly Journal of the Royal Meteorological Society*, **142**, 455–471. URL: https://doi.org/10.1002/qj.2664.

Weston, P. P., Bell, W. and Eyre, J. R. (2014) Accounting for correlated error in the assimilation of high-resolution sounder data. *Quarterly Journal of the Royal Meteorological Society*, **140**, 2420–2429. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2306.

Yaremchuk, M., D'Addezio, J. M., Panteleev, G. and Jacobs, G. (2018) On the approximation of the inverse error covariances of high-resolution satellite altimetry data. *Quarterly Journal of the Royal Meteorological Society*, **144**, 1995–2000. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3336.

Zienkiewicz, O. C. and Taylor, R. L. (2000) *The Finite Element method*. Butterworth-Heinemann.

A | IMPLEMENTATION OF THE FINITE-ELEMENT METHOD

To discretise the diffusion operator with FEM, we need a mesh that connects the observation locations as illustrated schematically in Figure 10. In 1D, the mesh is composed of:

- a set of N_n nodes, each corresponding to an observation location;
- a set of $N_n - 1$ edges, each linking two nodes together;
- a set of $N_n - 1$ cells, which partition the domain into sub-domains.

Edges and cells perfectly overlap in 1D[5]. The Voronoï cell, which is not defined explicitly on the mesh, is associated with a node and is delimited by the midpoints of the edges attached to the node (Figure 10).

Numerically, we can represent the mesh by a connectivity table that indicates which edges are related to which nodes. It can take the form of an array where each row corresponds to an edge, and contains the indices of the two nodes at both its ends. For example, the row associated with the red edge in Figure 10 would indicate the indices of the nodes highlighted in red. In 1D, building the connectivity table is straightforward. In our case, an edge of the mesh is created between nodes associated with successive observation times. Creating a mesh for 2D observations is more complex, and generally requires a meshing tool such as Atlas[6] to create connectivity tables from observation coordinates.

As defined in Equation (29), the FEM-based diffusion operator is defined in terms of the mass and stiffness matrices denoted $\mathbf{M}$ and $\mathbf{A}$, respectively. The elements on row k and column l link the k-th and l-th nodes are defined as (cf. Equations 27 and 28 in Guillet et al., 2019)

$$\begin{aligned} \mathbf{M}_{k,l} &= \int_\Omega \psi_k(z)\psi_l(z)dz, \\ \mathbf{A}_{k,l} &= \int_\Omega [\kappa(z)\nabla\psi_k(z)]^\mathsf{T}\,\nabla\psi_l(z)dz, \end{aligned} \tag{36}$$

where $(\psi_k)_{k\in[\![1,N_n]\!]}$ are the shape functions associated with $\mathcal{P}_1$ finite elements, $\kappa(z)$ is the diffusion tensor (possibly spatially varying), and Ω is the domain. In the 1D case considered in this study, Ω is the satellite track, $\nabla\psi_k$ has a single component $\partial\psi_k/\partial z$, and $\kappa(z)$ is a constant diffusion coefficient. The function ψ_k is piecewise linear on each

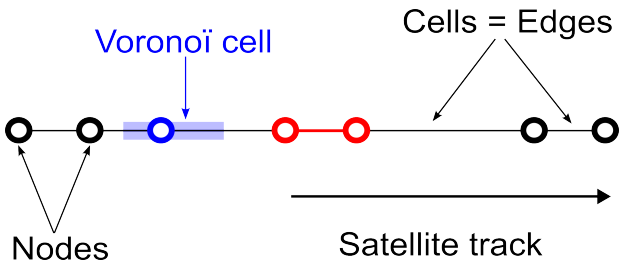


FIGURE 10 Schematic representation of the mesh used with the 1D diffusion operator for nadir altimeter data. The black circles denote the nodes associated with observations locations. The black lines denote both the edges of the mesh, which connect the observation locations, and the cells of the mesh, which divide the domain into sub-domains. The blue filled-in area is the Voronoï cell associated with the observation location highlighted in blue.

[5] In higher dimensions, cells and edges are distinct: in 2D, edges are still segments linking observation locations, while cells can be triangular surfaces.
[6] https://sites.ecmwf.int/docs/atlas/

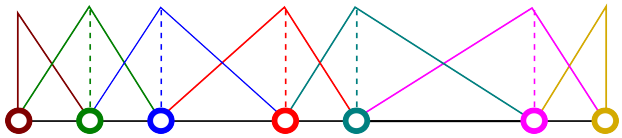


FIGURE 11 Schematic representation of the shape functions associated in 1D with $\mathcal{P}_1$ elements.

cell, equal to 1 at the k-th node, and 0 at other nodes, as illustrated in Figure 11. The local compact support of these functions is responsible for the sparse structure of $\mathbf{M}$ and $\mathbf{A}$: the integrals in Equation (36) are non-zero only if the nodes k and l are linked by an edge, or if $k = l$. If the nodes are labelled consistently with their position in the mesh then $\mathbf{M}$ and $\mathbf{A}$ are tri-diagonal matrices.

To reduce computational costs, we adopt the common approach of mass lumping (*e.g.*, Section 17.2.4 in Zienkiewicz and Taylor, 2000), which approximates $\mathbf{M}$ by a diagonal matrix $\overline{\mathbf{M}}$. Each diagonal element of $\overline{\mathbf{M}}$ is obtained by summing all the elements in the corresponding row:

$$\overline{\mathbf{M}}_{k,k} = \sum_{l=1}^{N_n} \mathbf{M}_{k,l}.$$

This assumption is equivalent to taking $\overline{\mathbf{M}}_{k,k}$ to be the surface of the Voronoï cell centred around the k-th node (Figure 10).

The most computationally efficient way to compute the value of the non-zero integrals is to cycle through the cells one by one in order to compute their contribution to all of the integrals. We thus separate the contribution of each cell to the integrals; *i.e.*, we decompose the satellite ground track into the segments that connect the observation locations:

$$\mathbf{M}_{n_k,n_l} = \sum_{\omega \subset \Omega} \mathbf{M}_{n_k,n_l}(\omega) \quad \text{with} \quad \mathbf{M}_{n_k,n_l}(\omega) = \int_\omega \psi_{n_k}(z)\psi_{n_l}(z)dz, \tag{37}$$

$$\mathbf{A}_{n_k,n_l} = \sum_{\omega \subset \Omega} \mathbf{A}_{n_k,n_l}(\omega) \quad \text{with} \quad \mathbf{A}_{n_k,n_l}(\omega) = \int_\omega \left[\kappa(z)\nabla\psi_{n_k}(z)\right]^\mathsf{T} \nabla\psi_{n_l}(z)dz, \tag{38}$$

where ω is the sub-domain represented by a cell. As the shape functions are linear within each cell, their gradient is constant on each cell. Using this property, we can simplify the integral in Equation (38) as

$$\mathbf{A}_{k,l}(\omega) = \left[\nabla\psi_k(\omega)\right]^\mathsf{T} \left(\int_\omega \kappa(z)dz\right) \nabla\psi_l(\omega). \tag{39}$$

Assuming $\kappa(z)$ is linear on each element (*i.e.*, that it can be written in terms of the basis functions $(\psi_{n_k})_{k\in[\![1,N_n]\!]}$), the integral in Equation (38) becomes

$$\int_\omega \kappa(z)dz = m(\omega)\kappa(\omega),$$

where $\kappa(\omega)$ is the mean of the values of κ at the nodes delimiting the cell, and $m(\omega)$ is the length of the cell. From Equation (39), we have

$$\mathbf{A}_{k,l}(\omega) = m(\omega)\left[\nabla\psi_k(\omega)\right]^\mathsf{T} \kappa(\omega)\ \nabla\psi_l(\omega).$$

To evaluate the contribution of a cell ω to these integrals, we use the connectivity table to identify the indices k_1 and k_2 of the nodes delimiting the cell. Let z_1 and z_2 denote the coordinates of these nodes, respectively. For nadir altimeters, observation error is correlated with respect to the observation time along the satellite track. However, we prefer to use the geographical position of the observations to define the coordinate z in order to facilitate comparisons with background-error correlations. The observation time is thus converted into a position along the satellite track (with respect to any reference as only relative distances, not absolute positions, are required) by multiplying it by the average ground-scanning velocity of the satellite, provided by the satellite operator.

The shape function ψ_{k_1} (ψ_{k_2}) is equal to 1 at z_1 (z_2) and 0 at z_2 (z_1), and varies linearly between them:

$$\psi_{k_1}(z) = \frac{z_{k_2} - z}{m(\omega)} \quad \text{and} \quad \psi_{k_2}(z) = \frac{z - z_1}{m(\omega)}, \tag{40}$$

where $m(\omega) = z_{k_2} - z_{k_1}$. The integrals in Equation (37) can be evaluated using Equation (40). Instead of building a non-diagonal $\mathbf{M}$ and then summing its elements through mass lumping, we build $\overline{\mathbf{M}}$ directly:

$$\left.\begin{array}{lcl} \overline{\mathbf{M}}_{k_1,k_1}(\omega) & = & \displaystyle\int_\omega \psi_{k_1}(z)\psi_{k_1}(z)dz + \int_\omega \psi_{k_1}(z)\psi_{k_2}(z)dz = \frac{m(\omega)}{3} + \frac{m(\omega)}{6} = \frac{m(\omega)}{2}, \\ \overline{\mathbf{M}}_{k_2,k_2}(\omega) & = & \displaystyle\int_\omega \psi_{k_2}(z)\psi_{k_2}(z)dz + \int_\omega \psi_{k_2}(z)\psi_{k_1}(z)dz = \frac{m(\omega)}{3} + \frac{m(\omega)}{6} = \frac{m(\omega)}{2}, \\ \overline{\mathbf{M}}_{k_1,k_2}(\omega) & = & 0, \\ \overline{\mathbf{M}}_{k_2,k_1}(\omega) & = & 0. \end{array}\right\} \tag{41}$$

The two components of the gradient of the shape function at nodes k_1 and k_2 is

$$\frac{\partial \psi_{k_1}}{\partial z} = -\frac{1}{m(\omega)} \quad \text{and} \quad \frac{\partial \psi_{k_2}}{\partial z} = \frac{1}{m(\omega)}.$$

The integrals in Equation (38) for the stiffness matrix $\mathbf{A}$ then become

$$\mathbf{A}_{n_1,n_1}(\omega) = \mathbf{A}_{n_2,n_2}(\omega) = \frac{\kappa(\omega)}{m(\omega)} \quad \text{and} \quad \mathbf{A}_{n_1,n_2}(\omega) = \mathbf{A}_{n_2,n_1}(\omega) = -\frac{\kappa(\omega)}{m(\omega)}. \tag{42}$$

The matrices $\mathbf{M}$ and $\mathbf{A}$ could be built explicitly by initializing their elements to 0, and then, for each cell ω, computing $m(\omega)$ and $\kappa(\omega)$, and incrementing the relevant matrix coefficients of $\overline{\mathbf{M}}$ and $\mathbf{A}$ based on Equations (41) and (42). In practice, $\overline{\mathbf{M}}$ and $\mathbf{A}$ are not stored as explicit matrices to avoid excessive memory cost. As $\overline{\mathbf{M}}$ is diagonal with mass lumping, it can be represented by a single value per node. These values form what is referred to as the *mass field*, which on the k-th node is given by

$$m(k) = \overline{\mathbf{M}}_{k,k}.$$

The similarity with the notation used for the length of each cell ($m(\omega)$) is deliberate, as the mass field can be seen as its counterpart defined on nodes rather than cells.

A typical row of a matrix-vector product $\mathbf{A}\boldsymbol{\eta}$ is

$$[\mathbf{A}\boldsymbol{\eta}]_k = \mathbf{A}_{k,k}\boldsymbol{\eta}_k + \mathbf{A}_{k,k-1}\boldsymbol{\eta}_{k-1} + \mathbf{A}_{k,k+1}\boldsymbol{\eta}_{k+1}.$$

As the rows of $\mathbf{A}$ sum to 0 by construction, $\mathbf{A}_{k,k} + \mathbf{A}_{k,k-1} + \mathbf{A}_{k,k+1} = 0$, we have

$$[\mathbf{A}\boldsymbol{\eta}]_k = \mathbf{A}_{k,k-1}(\boldsymbol{\eta}_{k-1} - \boldsymbol{\eta}_{\boldsymbol{k}}) + \mathbf{A}_{k,k+1}(\boldsymbol{\eta}_{k+1} - \boldsymbol{\eta}_{\boldsymbol{k}}),$$

and there is thus no need to store the diagonal elements $\mathbf{A}_{k,k}$ of $\mathbf{A}$. Furthermore, as $\mathbf{A}$ is symmetric, it is sufficient to store only one of its sub-diagonals. Each element on the sub-diagonal vector can be associated with an edge of the mesh. Instead of building the matrix $\mathbf{A}$, we build a *stiffness field* defined on the edge e associated with the cell ω as

$$a(e) = -\mathbf{A}_{k_1,k_2}.$$

Knowing the mass field (at nodes) and the stiffness field (on edges) is sufficient to be able to multiply a vector $\boldsymbol{\eta}$ by $\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$ as summarised in Algorithm 1. This product is the basic operation applied on each (inverse diffusion) iteration in $\mathbf{D}^{-1}$ (Equation (29)). An advantage of this formulation is that the numerical code associated with Algorithm 1 for the matrix-vector product is completely generic between 1D and higher dimensions. On the other hand, although the main steps remain the same, the procedure for constructing the mass and stiffness fields depends on the dimension of the domain.

Algorithm 1 Product of a vector $\boldsymbol{\eta}$ with $\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$

Update the value of $\boldsymbol{\eta}$ in the halo.

for each edge e_i with $i \in [\![1, N_n - 1]\!]$ **do**
  Extract from the connectivity table the indices k_1, k_2 of the nodes at both ends of the edge e_i
  $f(e_i) \leftarrow a(e_i)\left(\frac{\eta_{k_1}}{\sqrt{m_{k_1}}} - \frac{\eta_{k_2}}{\sqrt{m_{k_2}}}\right)$; ▷ Compute the flux through the edge
end for

for each edge e_i with $i \in [\![1, N_e]\!]$ **do**
  $\eta_{k_1} \leftarrow \eta_{k_1} - \frac{f(e_i)}{\sqrt{m_{k_1}}}$; ▷ Apply the flux to the first node
  $\eta_{k_2} \leftarrow \eta_{k_2} + \frac{f(e_i)}{\sqrt{m_{k_2}}}$; ▷ Apply the flux to the second node
end for

Chapter 6

Ongoing developments and outlook

6.1 Regularised diffusion operators

6.1.1 Sensitivity at small spatial scales

We have observed in Section 4.4 of Chapter 5 that the use of some observation error correlation models for $\mathbf{R}^{-1}$ can lead to an extreme sensitivity, in which unexpected noise in the small spatial scales of the innovation can lead to an abnormally large analysis error. In particular, with diffusion operators, this behaviour was observed with a large number of diffusion iterations m_{o}, but could also occur with a large observation error correlation length scale ρ_{o}. Beyond diffusion operators, this issue is common to any correlation model that would specify (explicitly or implicitly) a particularly small observation error variance at small spatial scales. For example, building $\mathbf{R}$ explicitly with a SOAR function with a large length scale or with a Gaussian correlation function could lead to the same effect. An explicit matrix $\mathbf{R}$ defined directly from a diagnostic [such as that of Desroziers et al., 2005] does not necessarily, but could potentially contain this feature as well. The reason why an observation error variance specified as particularly low at small spatial scales is problematic is that it will translate into a particularly large amplification of small spatial scale features by the operator $\mathbf{R}^{-1}$.

As an (extreme) example, let us consider a correlation operator associated with a diffusion operator with $m_{\mathrm{o}} = 12$ – or sampled from a Gaussian correlation function – with observations every 25 kilometres and a length scale $\rho_{\mathrm{o}} = 450$ km. The order of magnitude of the ratio of its largest and smallest eigenvalues would be 10^{16}. Consequently, when applying $\mathbf{R}^{-1}$ to a vector, any feature associated with the smallest spatial scale that the grid can resolve would be amplified by a factor 10^{16} times larger than the one applied to large spatial scales. This means that noise of machine-precision level could see such an amplification that it could drown actual geophysical signals. In Section 4.4 of Chapter 5, this excessive sensitivity was also observed with m_{o} equal to 2 instead of 12, which corresponds to a SOAR function instead of an approximate Gaussian function. With this change, the condition number of $\mathbf{R}$ is much lower: 541 in the NEMOVAR experiments. Although $\mathbf{R}^{-1}$ was more resilient to small perturbations in this case than with $m_{\mathrm{o}} = 12$, the excessive amplification of noise at small spatial scales could still be triggered by altering the observation error to include not only the correlated error accounted for in $\mathbf{R}$ but also a small percentage of uncorrelated error. This uncorrelated error represented only 1% of the total observation error variance, which was kept equal to that specified in $\mathbf{R}$ (see Equation 6.1). Unfortunately, such additional white noise could be difficult to detect in diagnostics, as illustrated below.

Let $\boldsymbol{\epsilon}_{\mathrm{corr}}$ denote a vector of correlated error of covariance $\mathbf{R}$ (with $\rho_{\mathrm{o}} = 450$ km, $m_{\mathrm{o}} = 2$, and $\sigma_{\mathrm{o}} = 1$). Let $\boldsymbol{\epsilon}_{\mathrm{uncorr}}$ denote a vector of uncorrelated noise with the same variance as $\boldsymbol{\epsilon}_{\mathrm{corr}}$. The total observation error is

defined as

$$\boldsymbol{\epsilon}_{\mathrm{o}} = \frac{1}{\sqrt{1+\beta^2}}\boldsymbol{\epsilon}_{\mathrm{corr}} + \frac{\beta}{\sqrt{1+\beta^2}}\boldsymbol{\epsilon}_{\mathrm{uncorr}} \tag{6.1}$$

such that β controls the ratio of uncorrelated to correlated error, and the scaling factor $\sqrt{1+\beta^2}$ keeps the variance of $\boldsymbol{\epsilon}_{\mathrm{o}}$ independent of β. Figure 6.1 shows the correlation function and eigenvalue spectrum of the correlation matrix of $\boldsymbol{\epsilon}_{\mathrm{o}}$ for different values of β. For large values of β (0.25 or 0.5), the presence of additional white noise is visible in the correlation function as the there is a sharp drop at the origin of the graph. However, the correlation functions obtained with smaller values of β (0.01 or 0.05) are almost impossible to distinguish from the correlation function obtained in the absence of white noise ($\beta = 0$).

Let us imagine a scenario where the true observation error correlations correspond to the correlation function obtained with $\beta = 0.05$ in Figure 6.1a. If we were to select our correlation model by fitting a SOAR function to an estimate of the correlation function coming from a Desroziers diagnostic [Desroziers et al., 2005], the resulting correlation model would be the correlation function associated with $\beta = 0$. Visually, this model would appear to match properly the actual error correlations, especially in the presence of sampling noise in the empirical correlation function. However, Figure 6.1b shows how important the difference between correlation models with $\beta = 0$ or $\beta = 0.05$ is. When applying $\mathbf{R}^{-1}$ to a vector using a correlation model with $\beta = 0$ instead of $\beta = 0.05$, we would apply an amplification factor at the smallest spatial scales which would be about 100 times too large. This could lead to excessive noise at small spatial scales of the analysis similar to what was observed in Section 4.2.3 of Chapter 5. Note that, for larger values of ρ_{o} or m_{o}, or denser observations, the minimum eigenvalue obtained with $\beta = 0$ would be even smaller (see Equation 17 in Goux et al. [2024]) while the minimum eigenvalue obtained with $\beta > 0$ would remain stable, thus increasing their discrepancy.

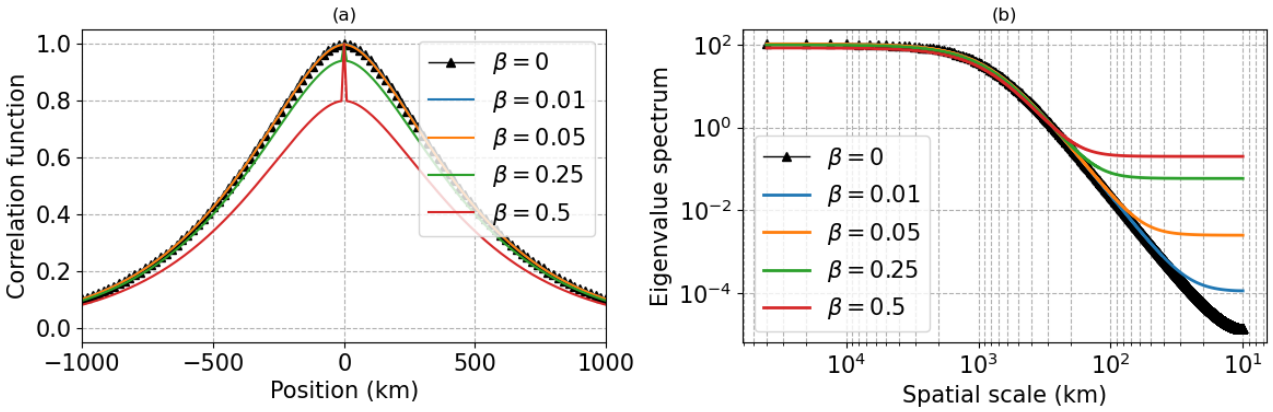


Figure 6.1: Correlation function (a) and eigenvalue spectrum of the correlation matrix (b) of the observation error as defined in Equation (6.1) for different values of β. Both are computed analytically, with $\rho_{\mathrm{o}} = 450$ km, $m_{\mathrm{o}} = 2$, and an observation every 10 km. The curves panel (a) for $\beta = 0$, 0.01 or 0.05 are almost perfectly overlapping.

If even a small, undetectable amount of white noise in the observation error is enough to generate a large analysis error, caution would suggest taking a conservative approach and assuming its presence by default. We do not have sufficient knowledge of the actual observation error correlations to know whether they are best modelled by a correlation function with $\beta = 0$ or a small non-zero value of β. When in doubt, we should use a correlation model that assumes a small amount of white noise is present as this approach would be more robust numerically. For instance, if we are not able to estimate if the variance at small spatial scales of the observation error is as large as 10^{-3} or as small as 10^{-5}, we should specify it as 10^{-3}. This choice is further supported by the

fact that, even if $\beta = 0$ was the best fit to the actual correlations, misspecifying the observation error variance at small spatial scales is not (significantly) detrimental to the quality of the analysis as explained in Section 2.3 of Chapter 5.

In practice, most methods currently developed to account for spatially correlated observation errors do not specify directly the eigenvalue spectrum of $\mathbf{R}$. For inter-channel correlations, $\mathbf{R}$ is block diagonal with blocks sufficiently small for their eigen-decompositions to be computed and altered manually. The 'eigenvalue tapering' applied to inter-channel error correlations by Weston et al. [2014], Bormann et al. [2016], Campbell et al. [2017] and Geer [2019] would create the resilience to unexpected white noise which we would like to obtain. With an explicit inversion of $\mathbf{R}$ to obtain $\mathbf{R}^{-1}$, such as with the SVD-FMM method, the 'reconditioning' approach used by Tabeart et al. [2020] could handle this issue. This approach consists of adding to $\mathbf{R}$ a scalar multiple of the identity matrix. While this method was introduced by Tabeart et al. [2020] primarily to accelerate the convergence of the CG algorithm, it should also successfully account for any unexpected noise in the small spatial scales of the innovation. However, this method is not applicable to diffusion operators, as their main advantage lies in the fact that they do not require $\mathbf{R}$ to be built and inverted. We outline below a revised method, which we are developing within the diffusion model framework to achieve the same effect as reconditioning or eigenvalue tapering without requiring access to $\mathbf{R}$.

6.1.2 From control theory to regularised diffusion operators

To do so, we draw inspiration from the methods used in control theory for the design of electrical and mechanical systems. Though more work remains to be done to formalise rigorously this link, we can already note some similarities between our description of diffusion operators and some of the classical methods from control theory. Visualising the eigenvalue spectrum of diffusion operators in a logarithmic space as we have done throughout this manuscript is relatively similar to the use of Bode diagrams [*e.g.*, Figure 4.6 of Dorf and Bishop, 2017]. These diagrams are often used to study visually the frequency response of time-invariant dynamical systems (and alter them). A common methodology is to break down a complex system into simple 'blocks' with an input, an output, and a transfer function (usually a differential equation) linking the two. The frequency response of each block is then studied on its own and approximated by a simple model. Then, the models are combined together to obtain a model of the response of the whole system. For our application, this could translate into breaking down a sophisticated linear operator into a sequence of simpler sub-operators, studying and finding a simple model for the eigenvalue spectrum of each sub-operator, and combining these models into a comprehensive model of the eigenvalue spectrum of the whole operator.

First, we need an assumption equivalent to the assumption of time-invariance used in control theory. As we are interested in the behaviour of diffusion operators spatial scale per spatial scale instead of a frequency response, this translates into an assumption of invariance in space of diffusion operators. This is exactly the assumption we made in the idealised frameworks used in Chapter 4 and 5 in order for $\mathbf{R}$ to be circulant: we assume that the correlation function is the same everywhere on the domain, and that the observations are distributed uniformly. This assumption will be critical to combine together the spectrums of the sub-operators.

Then, we can split a diffusion filter $\mathbf{D}$ into sub-operators. We use the FEM formulation from Equation (29) of Chapter 5 (recalled below) but the same process could be applied to the finite-difference formulation of Weaver and Mirouze [2013]:

$$\mathbf{D} = \overline{\mathbf{M}}^{-1/2} \left[\left(\mathbf{I} + \overline{\mathbf{M}}^{-1/2} \mathbf{A} \overline{\mathbf{M}}^{-1/2} \right)^{-1} \right]^{m} \overline{\mathbf{M}}^{-1/2}. \tag{6.2}$$

where $\overline{\mathbf{M}}$ is the diagonal, mass-lumped mass matrix. We distinguish two kinds of sub-operators in $\mathbf{D}$: a simple scaling by the diagonal matrix $\overline{\mathbf{M}}^{-1/2}$, and a diffusion step $\left(\mathbf{I} + \overline{\mathbf{M}}^{-1/2} \mathbf{A} \overline{\mathbf{M}}^{-1/2} \right)^{-1}$. The eigenvalue

spectrum of $\overline{\mathbf{M}}^{-1/2}$ is constant and equal to the inverse of the square root of the distance between observations. The eigenvalue spectrum of $\left(\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}\right)^{-1}$ can be relatively well approximated by a plateau at large spatial scales followed by a straight downward line as shown in Figure 6.2. This shape is typical of a 'first-order low pass filter', which commonly appears when modelling electrical and dynamical systems (for example, a Resistance-Capacitance filter as in Example 8.3 in Dorf and Bishop [2017]). In control theory, the point of transition between the two lines is often referred to as the 'cut-off', and the slope after the cut-off is often referred to as the 'roll-off'. The position of the cut-off depends on the correlation length scale. For a single diffusion step, the roll-off is fixed: the amplitude of the spectrum is (approximately) divided by 100 when the spatial scale is divided by 10. If we re-use the vocabulary of control theory, this corresponds to a roll-off of $\log_{10}(100) = 2$ *decibels per decade*.[1]. This asymptotic approximation of the spectrum of $\left(\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}\right)^{-1}$ can be used as a simple model to summarise the effect in spectral space of a diffusion step.

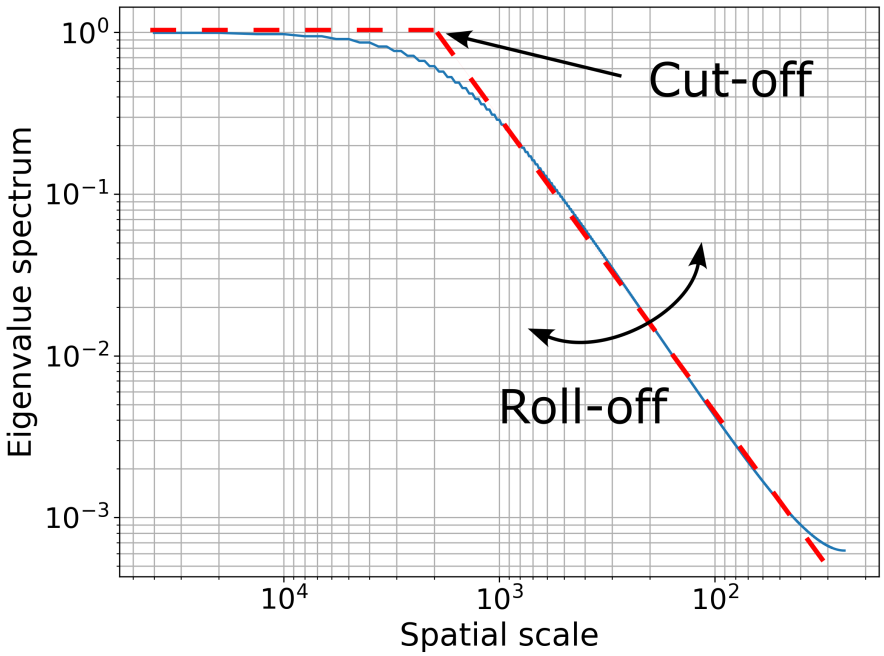


Figure 6.2: Eigenvalue spectrum (solid blue curve) of a diffusion step $\left(\mathbf{I}+\overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}\right)^{-1}$ and its asymptotic approximation (dashed red lines) with a length scale $\rho = 450$ km on a grid with a resolution of 25 km.

Based on the simplified spectrums of each sub-operator part of the sequence defining a diffusion operator, we can obtain a simplified spectrum of the whole diffusion operator. To do so, we make the assumption of homogeneity, so that these sub-operators can be represented by circulant matrices. They then all share the same eigenvectors, defined as a Fourier basis (see Section 3.1 of Goux et al. [2024]). Consequently, the spectrum of the diffusion operator is the product of the spectrums of the sub-operators. Drawing inspiration from control theory again, instead of deriving this product analytically, we use a graphical approach based on the simplified models. The simple models of the spectrums of the sub-operators are derived in a logarithmic scale. To find the shape of the simple model of the spectrum of the diffusion operator (also in a logarithmic scale), we can thus simply sum together the simple models.

The spectrum of $\overline{\mathbf{M}}^{-1/2}$ is constant , and thus only works as an offset to the spectrum of a diffusion operator.

[1] A scaling factor of 10 is often applied to the logarithm of the spectrum in control theory to define the 'gain' of the system, which thus shows a 20 decibels per decade roll-off instead.

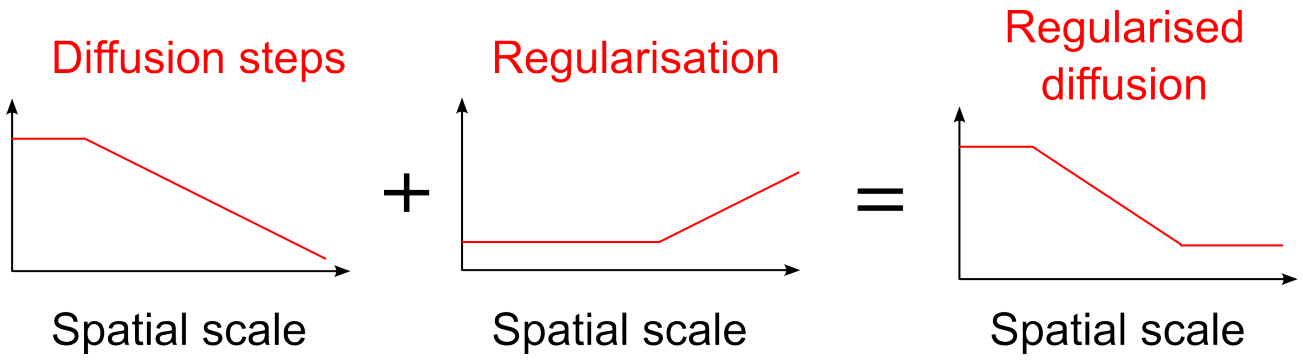


Figure 6.3: Schematic representation of the construction of a 'regularised' diffusion operator.

As diffusion operators are normalised when modelling correlations, we are not interested by the overall amplitude of the spectrum. Apart from this offset, the diffusion operator is composed of m diffusion steps. At large spatial scales, before their cut-off, the model of the spectrum of the diffusion operator is constant as it is the sum of constant models. After the cut-off, the (logarithmic) model of the spectrum of the diffusion operator is the sum of m segments with a roll-off of 2 decibels per decade each, resulting in a segment with a roll-off of $2m$ decibels per decade. In summary, the simplified model of the diffusion spectrum in logarithmic scale is a constant line at large spatial scales and, after a cut-off, a downward slope of $2m$ decibels per decade. This is consistent with what was observed when plotting the spectrums of AR functions in Figure A1 of Goux et al. [2024].

For now, we have simply shown the consistency of this approach with what we already know about the eigenvalue spectrum of diffusion operators. We will now use it as a basis to derive a 'regularised' diffusion operator, which would mimic the effect of reconditioning while remaining compatible with an operator-based approach. To do so, we need to stop the decrease in amplitude of the eigenvalue spectrum of the operator at the smallest spatial scales without perturbing the larger spatial scales. If we were to do so by adding a new 'regularisation' sub-operator in the sequence of Equation (6.2), we would need two properties from the spectrum of this operator:

- to be constant at large spatial scales in order not to perturb them;
- to increase at a rate of $2m$ decibels per decade at small spatial scales to compensate the decay from the m diffusion steps.

If the cut-off of the 'regularisation' term is located at smaller spatial scales than that of the diffusion, then we should be able to obtain an eigenvalue spectrum with the same shape as that of 'reconditioned' correlation operators, or as the spectrums with $\beta > 0$ in Figure 6.1b. A basic schematic representation of this idea is shown in Figure 6.3.

We already have at our disposal an operator whose eigenvalue spectrum fits these requirements and could act as a regularisation term: the inverse of a diffusion step. Let $\mathbf{A}_{\text{reg}}$ denote the stiffness matrix defined exactly like $\mathbf{A}$ apart from its length scale defined as ρ_{reg} instead of ρ. Then, we define our regularised diffusion operator $\mathbf{D}_{\text{reg}}$ as

$$\mathbf{D}_{\text{reg}} = \overline{\mathbf{M}}^{-1/2} \left(\mathbf{L}_{\text{reg}}\mathbf{L}^{-1}\right)^{m/2} \left(\mathbf{L}^{-1}\mathbf{L}_{\text{reg}}\right)^{m/2} \overline{\mathbf{M}}^{-1/2}. \tag{6.3}$$

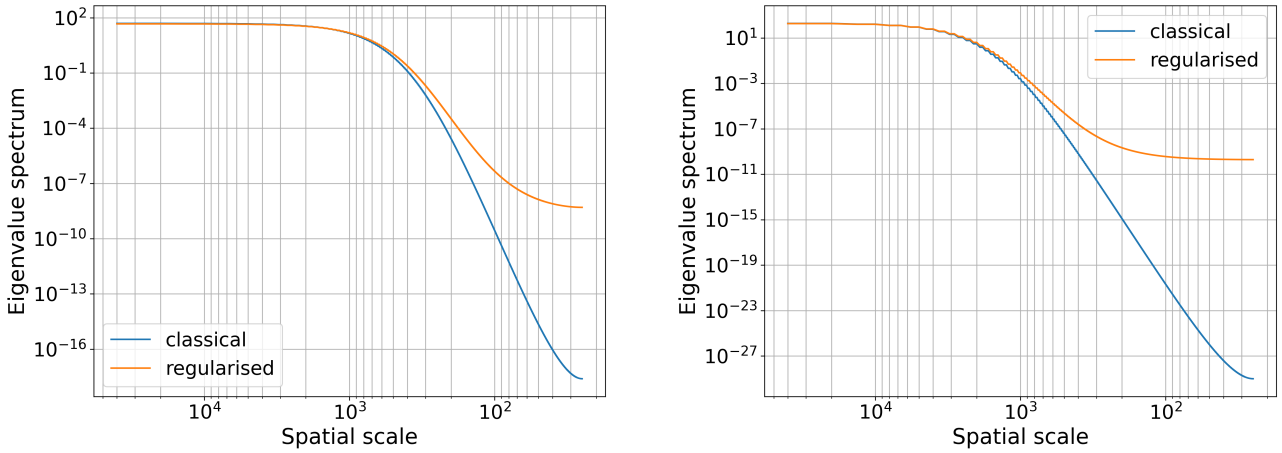
where

$$\mathbf{L} = \mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2},$$
$$\mathbf{L}_{\text{reg}} = \mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}_{\text{reg}}\overline{\mathbf{M}}^{-1/2}.$$

In this simple framework, all matrices are circulant and thus commutative, meaning that $\mathbf{D}_{\text{reg}}$ could equivalently be defined as

$$\mathbf{D}_{\text{reg}} = \overline{\mathbf{M}}^{-1/2} \mathbf{L}_{\text{reg}}^{m/2} \left(\mathbf{L}^{-1}\right)^{m} \mathbf{L}_{\text{reg}}^{m/2} \overline{\mathbf{M}}^{-1/2}.$$

However, the two formulations would not be equivalent with spatially varying correlation length scales. If ρ_{reg} is set to be smaller than ρ, then the cut-off of $\mathbf{L}_{\text{reg}}$ would occur at smaller spatial scales than that of $\mathbf{L}$, achieving the desired effect of limiting the decrease of the eigenvalue spectrum at the smallest spatial scales. This is illustrated in Figure 6.4, which shows the eigenvalue spectrums of diffusion operators (normalised to be correlation operators) with and without regularisation as defined in Equation (6.3). The variations of the actual eigenvalue spectrum are 'smoother' than that of its simplified model, but do have the behaviour we were expecting: the eigenvalue spectrum reaches a plateau at the smallest spatial scales. We can thus expect the regularised diffusion operator to be resilient to unexpected noise in the innovation. In practice, the operator we



(a) $\rho = 500$ km ; $\rho_{\text{reg}} = 150$ km

(b) $\rho = 2000$ km ; $\rho_{\text{reg}} = 500$ km

Figure 6.4: Normalised eigenvalue spectrum of classical and regularised diffusion operators, defined on a grid with a resolution of 25 km with a smoothness parameter $m = 2$.

actually apply at each iteration of the inner loop would be the inverse of $\mathbf{D}_{\text{reg}}$:

$$\mathbf{D}_{\text{reg}}^{-1} = \overline{\mathbf{M}}^{1/2} \left(\mathbf{L}_{\text{reg}}^{-1}\right)^{m/2} \mathbf{L}^{m} \left(\mathbf{L}_{\text{reg}}^{-1}\right)^{m/2} \overline{\mathbf{M}}^{1/2}.$$

A possible interpretation for this formulation is that, compared to $\mathbf{D}^{-1}$ which is purely a 'roughening' operator, $\mathbf{D}_{\text{reg}}^{-1}$ includes additional 'smoothing' operations that damp perturbations at the smallest spatial scales. In terms of computational cost, the regularised diffusion operator is more expensive than the 'classical' diffusion operator as it involves the solution of sparse linear systems defined by $\mathbf{L}_{\text{reg}}$. These systems could be solved with a truncated Chebyshev iteration as described in Section 3.2.3 of Chapter 5. As the application of a diffusion-based $\mathbf{B}$ in the inner loop would already involve the solution of similar linear systems, but of larger dimension, $\mathbf{R}$ with regularisation is unlikely to be a bottleneck in the inner loop.

6.1.3 Summary

Certain observation error correlation models attribute an extremely small variance to the small spatial scales, which can cause problems in the data assimilation algorithm. This can occur, for example, with a Gaussian correlation function ($m_{\text{o}} \geq 10$ with diffusion operators) or with very large correlation length scales. While the

first scenario can be avoided, observations with errors correlated at very large length scales (in particular larger than that of the background error) need to be considered (*e.g.*, for SWOT as discussed in Chapter 3). In this case, we have shown that low-level noise in the small spatial scales of the innovation could generate a much larger perturbation of the analysis. As this behaviour can be triggered by noise with undetectable amplitude, it is important to use observation error correlation models which are resilient to such perturbations. To do so, we must ensure that the scale-dependent observation error (*i.e.*, eigenvalue spectrum of $\mathbf{R}$) does not reach extremely low values. This can done by simply applying a 'threshold' to the explicitly accessible spectrum for inter-channel correlations. However, computing the eigen-decomposition of $\mathbf{R}$ is too expensive if spatial correlations are accounted for. If $\mathbf{R}$ is built and inverted explicitly, reconditioning techniques could be used to the same effect, but they are not applicable to diffusion operators.

We have drawn inspiration from the tools used in control theory to design the frequency response of electrical and mechanical systems as the basis to derive a 'regularised' diffusion operator. Compared to a classical diffusion operator, the regularised operator has an additional degree of freedom in the form of a regularisation length scale ρ_{reg}, which can be used to limit the presence of excessively small eigenvalues in $\mathbf{R}$. Setting $\rho_{\text{reg}} = 0$ leads to a classical diffusion operator, while setting $\rho_{\text{reg}} = \rho$ cancels out the actual diffusion. More work remains to be done to define exactly which value of $\rho_{\text{reg}} \in [0, \rho]$ should be used. To do so, methods from the field of control theory could be used once their formalism is more closely linked to that of data assimilation.

Beyond regularisation, this approach of using diffusion steps or their inverse with different length scales in the same correlation operator could also be used to add more flexibility to diffusion operators and model more complex correlation functions. For example, adding another set of diffusion steps to $\mathbf{D}_{\text{reg}}$ with an even smaller correlation length scale generates an eigenvalue spectrum reminiscent from the one which would be associated with a sum of Matérn functions with different length scales. However, we do not focus on this subject first, as creating additional degrees of freedoms is only useful if we also develop our ability to adjust them.

6.2 Extension to other observation types

In Appendix A of Chapter 5, we presented a method for modelling nadir altimeter error correlations in NEMOVAR. This method is based on the discretisation of an implicit diffusion equation with FEM. Although we only consider 1D datasets in that article, the FEM is straightforward to adapt to higher dimensional datasets. Of particular interest is the 2D case, which requires using a mesh generator to build a connectivity table for FEM. In Section 6.2.1, we present the formulation for the 2D FEM-based correlation operator, assuming that the connectivity table is available, and show how it relates to the 1D formulation. In Section 6.2.2, we discuss the challenges associated with mesh generation in 2D and our strategy to address these challenges.

6.2.1 Implementation for two-dimensional datasets

We assume the availability of a triangular 2D mesh that connects the observation locations, such as the one in the illustration of Figure 6.5. This mesh contains:

- nodes, *i.e.*, a set of points each defined by a pair of coordinates corresponding to an observation location, indexed as n_1, n_2, …, n_{N_n};
- edges, *i.e.*, a set of lines each linking two nodes together, and which do not intersect, indexed as e_1, e_2, …, e_{N_e};
- cells, *i.e.*, a set of triangles whose corners are nodes and whose sides are edges, indexed as ω_1, ω_2, …, ω_{N_ω}.

The cells can be seen as a partition of the domain Ω into sub-domains:

$$\Omega = \bigcup_{i \in [\![1, N_\omega]\!]} \omega_i.$$

Although not defined explicitly on the mesh, we will also need to consider Voronoï cells. A Voronoï cell associated with a node is the polygon whose corners are the barycentres of the surrounding cells. The blue filled-in area in Figure 6.5 shows the Voronoï cell associated with the node highlighted in blue.

Examples of triangular meshes applied to real data are given in Section 6.2.2. The mesh is represented numerically by connectivity tables, which indicate how the nodes, edges and cells relate to each other. Many representations are possible, but three of them will be of interest to us and are listed below.

- We denote by $C_{e,n}$, the connectivity table that associates edges to nodes. It is an array where each row is associated with an edge of the mesh and contains the indices of the nodes at both ends of the edge. For example, the row associated with the red edge in Figure 6.5 would contain the indices of the two nodes highlighted in red.
- The connectivity table $C_{\omega,n}$ associates cells to nodes. It is an array where each row is associated with a cell of the mesh and contains the indices of the nodes at the corners of the cell. For example, the row associated with the triangle filled in green contains the indices of the nodes highlighted in green.
- The connectivity table $C_{\omega,e}$ associates cells to edges. It is an array where each row is associated with a cell of the mesh and contains the indices of the edges surrounding it. For example, the row associated with the triangle filled in green contains the indices of the edges highlighted in green.

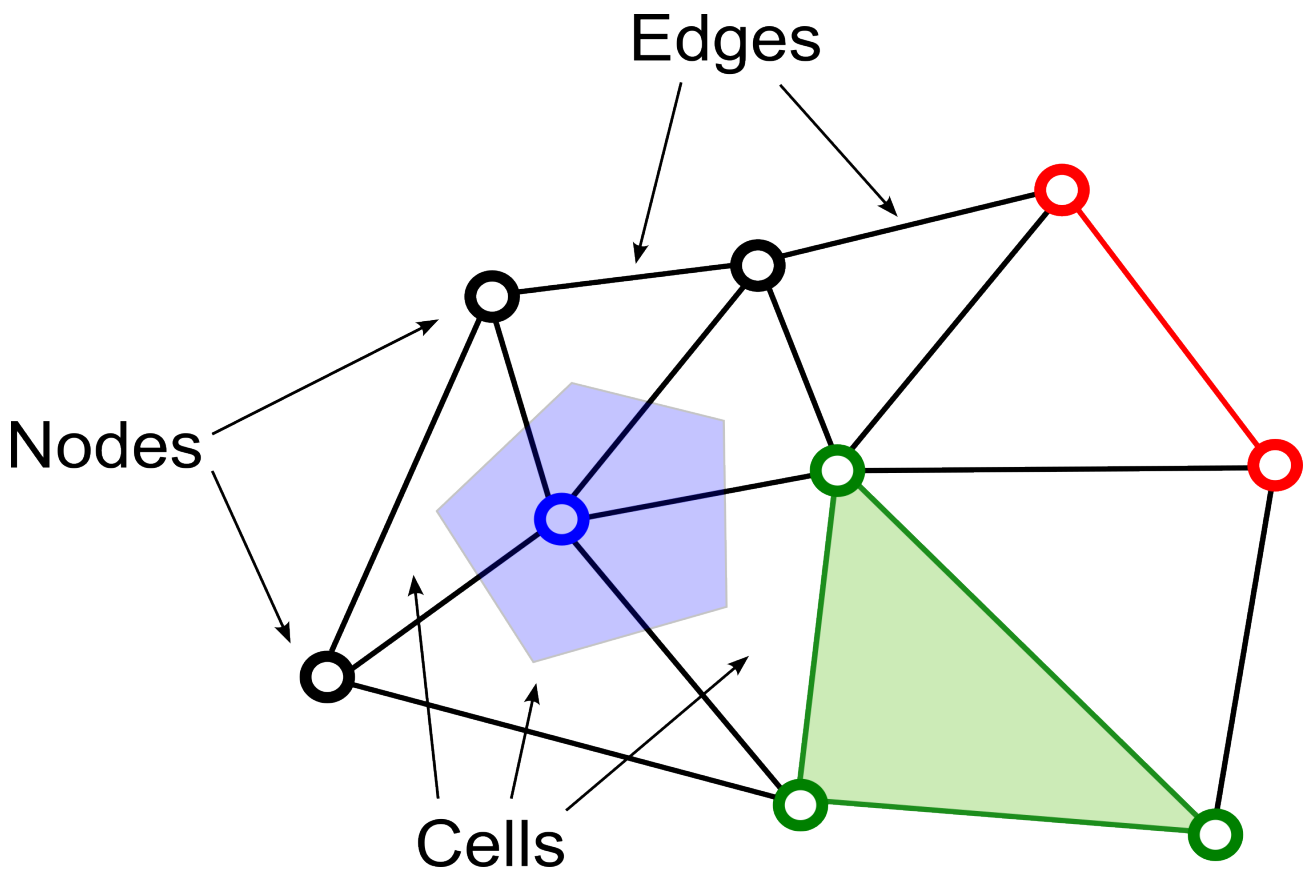


Figure 6.5: Schematic representation of a mesh used in FEM. The circles are its nodes, which correspond to the observation locations. The lines are its edges, which link nodes together. The triangles delimited by three edges are the mesh cells.

As defined in Equations (40) and (43) of Guillet et al. [2019], and as recalled below, the FEM-based diffusion operator $\mathbf{D}$ can be written in terms of the mass and stiffness matrices denoted $\mathbf{M}$ and $\mathbf{A}$[2], respectively:

$$\mathbf{D} = \left[(\mathbf{M} + \mathbf{A})^{-1} \mathbf{M} \right]^m \mathbf{M}^{-1}$$
$$\text{and} \quad \mathbf{D}^{-1} = \mathbf{M} \left[\mathbf{M}^{-1} (\mathbf{M} + \mathbf{A}) \right]^m .$$

In the set-up of the diffusion operator, we do not build $\mathbf{M}$ and $\mathbf{A}$ explicitly, but only intermediate quantities that allow the computation of matrix-vector products with them. The first steps of the analytical developments are common between 1D and 2D and are already presented in Appendix A of Chapter 5, but are recalled below for the sake of clarity.

We start from the expression of the elements of $\mathbf{M}$ and $\mathbf{A}$ which associate the node n_k to the node n_l (*i.e.*, on the row n_k and column n_l) (see Equations (27-28) in [Guillet et al., 2019]):

$$\mathbf{M}_{n_k,n_l} = \int_\Omega \psi_{n_k}(\boldsymbol{z}) \psi_{n_l}(\boldsymbol{z}) d\boldsymbol{z},$$
$$\mathbf{A}_{n_k,n_l} = \int_\Omega \left[\boldsymbol{\kappa}(\boldsymbol{z}) \boldsymbol{\nabla} \psi_{n_k}(\boldsymbol{z}) \right]^{\mathrm{T}} \boldsymbol{\nabla} \psi_{n_l}(\boldsymbol{z}) d\boldsymbol{z}, \tag{6.4}$$

[2]To denote the stiffness matrix, we use the notation $\mathbf{A}$ instead of the standard notation $\mathbf{K}$ to avoid confusion with the Kalman gain matrix.

where the functions $(\psi_{n_i})_{i\in[\![1,N_n]\!]}$ are the shape functions associated with $\mathcal{P}_1$ finite elements. They are piecewise-linear on each cell, equal to 1 at the node they are associated with and 0 at other nodes. An illustration of such a function is given in Figure 6.6.

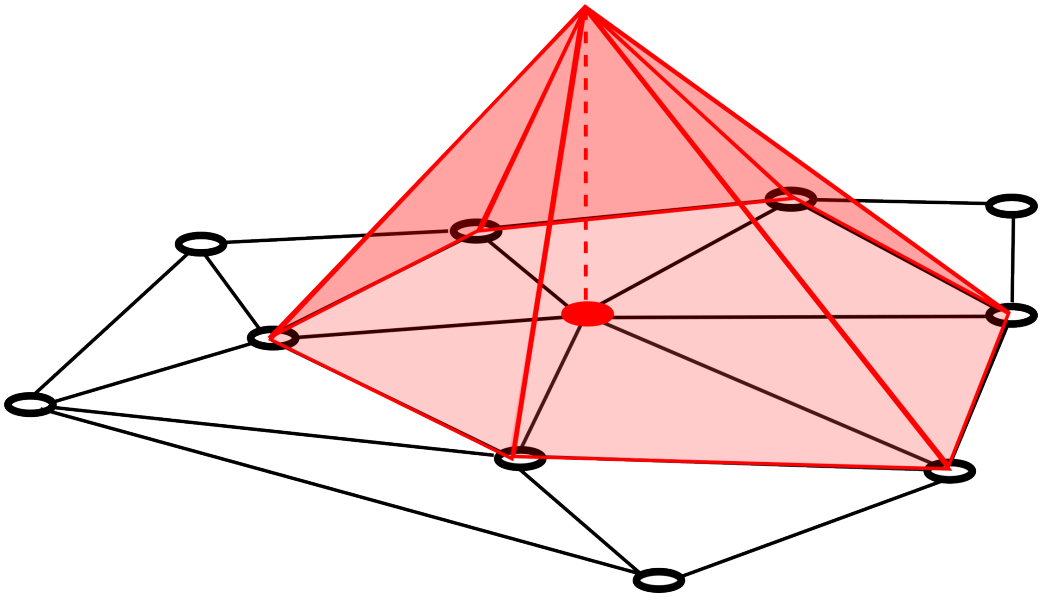


Figure 6.6: Schematic representation of the $\mathcal{P}_1$ shape function associated with the node highlighted in red. The function is piecewise-linear on each cell, equal to 1 at the node it is associated with and 0 at other nodes.

Due to the compact support of the shape functions, most of the integrals in Equation (6.4) are equal to zero: the supports of ψ_{n_k} and ψ_{n_l} only overlap if the nodes n_k and n_l are connected by an edge, or if $k = l$. It would not be efficient to loop on nodes and compute each non-zero integral from Equation (6.4) one by one. Instead, we loop on cells, and for each compute their contribution to each integral:

$$\mathbf{M}_{n_k,n_l} = \sum_{\omega\subset\Omega} \mathbf{M}_{n_k,n_l}(\omega), \quad \text{with} \quad \mathbf{M}_{n_k,n_l}(\omega) = \int_\omega \psi_{n_k}(\boldsymbol{z})\psi_{n_l}(\boldsymbol{z})d\boldsymbol{z}, \tag{6.5}$$

$$\mathbf{A}_{n_k,n_l} = \sum_{\omega\subset\Omega} \mathbf{A}_{n_k,n_l}(\omega), \quad \text{with} \quad \mathbf{A}_{n_k,n_l}(\omega) = \int_\omega \left[\boldsymbol{\kappa}(\boldsymbol{z})\boldsymbol{\nabla}\psi_{n_k}(\boldsymbol{z})\right]^{\mathrm{T}} \boldsymbol{\nabla}\psi_{n_l}(\boldsymbol{z})d\boldsymbol{z}. \tag{6.6}$$

As shape functions are linear within each cell, their gradient is constant on each cell. Using this property and the symmetry of the diffusion tensor, we can simplify the second part of Equation (6.6) as

$$\mathbf{A}_{n_k,n_l}(\omega) = \boldsymbol{\nabla}\psi_{n_k}(\omega)^{\mathrm{T}} \left(\int_\omega \boldsymbol{\kappa}(\boldsymbol{z})d\boldsymbol{z}\right) \boldsymbol{\nabla}\psi_{n_l}(\omega). \tag{6.7}$$

We assume $\boldsymbol{\kappa}$ to be linear on each cell (*i.e.*, that it can be decomposed in the basis of the functions $(\psi_{n_k})_{k\in[\![1,N_n]\!]}$), in which case the integral in Equation (6.6) becomes equal to

$$\int_\omega \boldsymbol{\kappa}(\boldsymbol{z})d\boldsymbol{z} = m(\omega)\boldsymbol{\kappa}(\omega), \tag{6.8}$$

where $\boldsymbol{\kappa}(\omega)$ is the mean of the values of $\boldsymbol{\kappa}$ defined at the nodes surrounding the cell, and $m(\omega)$ is the *mensuration* of the cell, *i.e.*, its length in 1D, surface in 2D, volume in 3D, *etc.* From Equations (6.7) and (6.8), we have

$$\mathbf{A}_{n_k,n_l}(\omega) = m(\omega)\ \boldsymbol{\nabla}\psi_{n_k}(\omega)^{\mathrm{T}}\ \boldsymbol{\kappa}(\omega)\ \boldsymbol{\nabla}\psi_{n_l}(\omega). \tag{6.9}$$

In this formulation, the matrix $\mathbf{M}$ is non-diagonal. To reduce computational costs, we can use the common assumption of mass lumping [*e.g.*, Section 17.2.4 in Zienkiewicz and Taylor [2000]], which approximates $\mathbf{M}$ by a diagonal matrix $\overline{\mathbf{M}}$. The diagonal elements of $\overline{\mathbf{M}}$ are obtained by summing all the elements of the corresponding row:

$$\overline{\mathbf{M}}_{n_k,n_k} = \sum_{l=1}^{N_n} \mathbf{M}_{n_k,n_l}.$$

This assumption is equivalent to considering that $\overline{\mathbf{M}}_{n_k,n_k}$ is the surface of the Voronoï cell centered around the node n_k. This assumption will be systematically used in the rest of this section.

With $\overline{\mathbf{M}}$ diagonal, the sequence of sub-operators representing a diffusion operator can be reformulated in terms of a product of non-dimensional operators $\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$:

$$\mathbf{D} = \overline{\mathbf{M}}^{-\frac{1}{2}} \left[\left(\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2} \right)^{-1} \right]^m \overline{\mathbf{M}}^{-\frac{1}{2}},$$
$$\mathbf{D}^{-1} = \overline{\mathbf{M}}^{\frac{1}{2}} \left(\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2} \right)^m \overline{\mathbf{M}}^{\frac{1}{2}}.$$

The process of combining the contributions of each cell to the mass and stiffness matrix is usually referred to as the *assembly* of $\overline{\mathbf{M}}$ and $\mathbf{A}$ [Ern and Guermond, 2004]. In our case, we do not build the matrices $\overline{\mathbf{M}}$ and $\mathbf{A}$ explicitly (or using a 'classical' sparse matrix format) but generate instead a 'mass field' and a 'stiffness field', which are sufficient to implement matrix-vector products with $\overline{\mathbf{M}}$ and $\mathbf{A}$, respectively. This choice makes it more efficient to combine the contributions of the cells 'on the fly' rather than computing and storing them all first. We thus refer to the 'assembly step' as the process of evaluating all the contributions from a given cell ω and adding these contributions to the mass and stiffness fields.

Assembly in 2D The assembly procedure in 2D follows the same steps as in 1D, but is more complex as there are three nodes and three edges associated with each cell. The end result in our case is a mass field defined at nodes and a stiffness field defined on edges. This allows Algorithm 2, which defines the product with $\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$, to be generic between 1D and 2D.

The row of the connectivity table $C_{\omega,n}$ associated with ω indicates which nodes define the corners of the cell ω. We assume the nodes in question to be those associated with the indices n_1, n_2 and n_3. Let $\boldsymbol{z}_1 = (x_1, y_1)$, $\boldsymbol{z}_2 = (x_2, y_2)$, and $\boldsymbol{z}_3 = (x_3, y_3)$ denote their 2D coordinates, respectively. As we are using $\mathcal{P}_1$ finite elements, the shape functions are piecewise linear over the cells of the mesh, and thus entirely determined by their values at the nodes of the mesh. The shape function ψ_{n_i} is equal to 1 at the location of the node n_i and 0 at the location of the other nodes. Under these constraints, we can express the shape functions on the cell ω as

$$\left.\begin{aligned} \psi_{n_1}(x,y) &= \frac{1}{2m(\omega)} \left[x_2y_3 - x_3y_2 + (y_2 - y_3)x + (x_3 - x_2)y \right], \\ \psi_{n_2}(x,y) &= \frac{1}{2m(\omega)} \left[x_3y_1 - x_1y_3 + (y_3 - y_1)x + (x_1 - x_3)y \right], \\ \psi_{n_3}(x,y) &= \frac{1}{2m(\omega)} \left[x_1y_2 - x_2y_1 + (y_1 - y_2)x + (x_2 - x_1)y \right], \end{aligned}\right\} \tag{6.10}$$

where

$$m(\omega) = \frac{1}{2} \left(x_2y_3 - x_3y_2 - x_1y_3 + x_3y_1 + x_1y_2 - x_2y_1 \right) \tag{6.11}$$

is the surface of ω. Rather than using Equation (6.11), it is more convenient to compute the surface $m(\omega)$ using

Heron's formula, which defines the surface area of a triangle based on the length of its sides:

$$m(\omega) = \sqrt{h(h-d_1)(h-d_2)(h-d_3)},$$

where d_1, d_2, and d_3 are the lengths of the three edges surrounding the triangle ω, and $h = (d_1 + d_2 + d_3)/2$.

Instead of building a non-diagonal $\mathbf{M}$ and then summing its elements row by row, we focus on building the diagonal $\overline{\mathbf{M}}$ directly. Substituting Equations (6.10) in Equation (6.5) yields

$$\begin{aligned}
\overline{\mathbf{M}}_{n_1,n_1}(\omega) &= \int_\omega \psi_{n_1}(\boldsymbol{z})\psi_{n_1}(\boldsymbol{z})d\boldsymbol{z} + \int_\omega \psi_{n_1}(\boldsymbol{z})\psi_{n_2}(\boldsymbol{z})d\boldsymbol{z} + \int_\omega \psi_{n_1}(\boldsymbol{z})\psi_{n_3}(\boldsymbol{z})d\boldsymbol{z} \\
&= \frac{m(\omega)}{6} + \frac{m(\omega)}{12} + \frac{m(\omega)}{12} = \frac{m(\omega)}{3}, \\
\overline{\mathbf{M}}_{n_2,n_2}(\omega) &= \int_\omega \psi_{n_2}(\boldsymbol{z})\psi_{n_1}(\boldsymbol{z})d\boldsymbol{z} + \int_\omega \psi_{n_2}(\boldsymbol{z})\psi_{n_2}(\boldsymbol{z})d\boldsymbol{z} + \int_\omega \psi_{n_2}(\boldsymbol{z})\psi_{n_3}(\boldsymbol{z})d\boldsymbol{z} \\
&= \frac{m(\omega)}{12} + \frac{m(\omega)}{6} + \frac{m(\omega)}{12} = \frac{m(\omega)}{3}, \\
\overline{\mathbf{M}}_{n_3,n_3}(\omega) &= \int_\omega \psi_{n_3}(\boldsymbol{z})\psi_{n_1}(\boldsymbol{z})d\boldsymbol{z} + \int_\omega \psi_{n_3}(\boldsymbol{z})\psi_{n_2}(\boldsymbol{z})d\boldsymbol{z} + \int_\omega \psi_{n_3}(\boldsymbol{z})\psi_{n_3}(\boldsymbol{z})d\boldsymbol{z} \\
&= \frac{m(\omega)}{12} + \frac{m(\omega)}{12} + \frac{m(\omega)}{6} = \frac{m(\omega)}{3}, \\
\overline{\mathbf{M}}_{n_1,n_2} &= \overline{\mathbf{M}}_{n_1,n_3} = \overline{\mathbf{M}}_{n_2,n_1} = \overline{\mathbf{M}}_{n_2,n_3} = 0.
\end{aligned}$$

As $\overline{\mathbf{M}}$ is diagonal, it can be represented by a single value per node. These values form what we refer to as the 'mass field', denoted $[m(n_i)]_{i\in[\![1,N_n]\!]}$:

$$m(n_i) = \overline{\mathbf{M}}_{n_i,n_i}.$$

The similarity with the notation of the mensuration of each cell $[m(\omega_i)]_{i\in[\![1,N_\omega]\!]}$ is deliberate, as the mass field can be seen as a counterpart of the mensuration defined on nodes and not cells.

We now turn our attention to the construction of a stiffness field, which can be used as a representation of the stiffness matrix. The matrix $\mathbf{A}$ is not diagonal like $\overline{\mathbf{M}}$, but it is symmetric and each of its rows sums to 0 by construction. A typical row of the matrix-vector product $\mathbf{A}\boldsymbol{\eta}$, for any vector η, is

$$[\mathbf{A}\boldsymbol{\eta}]_k = \mathbf{A}_{k,k}\boldsymbol{\eta}_k + \mathbf{A}_{k,k-1}\boldsymbol{\eta}_{k-1} + \mathbf{A}_{k,k+1}\boldsymbol{\eta}_{k+1}.$$

As rows of $\mathbf{A}$ sum to 0, $\mathbf{A}_{k,k} + \mathbf{A}_{k,k-1} + \mathbf{A}_{k,k+1} = 0$, and we can rewrite the matrix-vector product as

$$[\mathbf{A}\boldsymbol{\eta}]_k = \mathbf{A}_{k,k-1}(\boldsymbol{\eta}_{k-1} - \boldsymbol{\eta}_{\boldsymbol{k}}) + \mathbf{A}_{k,k+1}(\boldsymbol{\eta}_{k+1} - \boldsymbol{\eta}_{\boldsymbol{k}}).$$

We thus only need to store the off-diagonal elements of $\mathbf{A}$ to apply a matrix-vector product with it. As $\mathbf{A}$ is also symmetric, we only need half of these elements. The non-zero off-diagonal elements of $\mathbf{A}$ each correspond to a pair of nodes that are linked by an edge in the mesh. We can thus have a full representation of $\mathbf{A}$ by storing one value per edge, which yields the stiffness field

$$a(e) = -\mathbf{A}_{n_1,n_2},$$

where e is an edge linking the nodes n_1 and n_2. To compute the values of the stiffness fields and hence the

elements of $\mathbf{A}$, we use the fact that the gradient of the shape functions are constant on the cell:

$$\boldsymbol{\nabla}\psi_{n_1} = \frac{1}{2m(\omega)}\begin{pmatrix} y_2 - y_3 \\ x_3 - x_2 \end{pmatrix} \quad ; \quad \boldsymbol{\nabla}\psi_{n_2} = \frac{1}{2m(\omega)}\begin{pmatrix} y_3 - y_1 \\ x_1 - x_3 \end{pmatrix} \quad ; \quad \boldsymbol{\nabla}\psi_{n_3} = \frac{1}{2m(\omega)}\begin{pmatrix} y_1 - y_2 \\ x_2 - y_1 \end{pmatrix}. \tag{6.12}$$

In 2D, $\boldsymbol{\kappa}(\omega)$ is a 2×2 symmetric tensor:

$$\boldsymbol{\kappa}(\omega) = \begin{pmatrix} \kappa_{x,x}(\omega) & \kappa_{x,y}(\omega) \\ \kappa_{x,y}(\omega) & \kappa_{y,y}(\omega) \end{pmatrix}. \tag{6.13}$$

Equations (6.12) and (6.13) can be used along with Equation (6.9) to compute the contributions of the cell ω to $\mathbf{A}$.

Algorithm 1 details how the assembly code can be organised in order for the stiffness field to be associated with the appropriate edge while minimizing redundant computations. This algorithm is applicable to parallel systems if, on a given processor, the list of cells include all cells with at least one observation location attributed to this processor as a node. Algorithm 2 demonstrates how the matrix-vector product with $\mathbf{I} + \overline{\mathbf{M}}^{-\frac{1}{2}}\mathbf{A}\overline{\mathbf{M}}^{-\frac{1}{2}}$, which is the basis of the inverse diffusion operator (Equation 6.2.1), can be evaluated using the mass and stiffness fields.

Algorithm 1 Assembly procedure in 2D

Initialise the stiffness field a to 0 on each edge.
Initialise the mass field m to 0 on each node.
Create an array d of size N_e.

for each edge e_i with $i \in [\![1, N_e]\!]$ **do**
  Use the connectivity table $C_{e,n}$ to determine the indices n_1 and n_2 at both ends of the edge e_i.
  Compute the distance between n_1 and n_2 and store it as $d(e_i)$.
end for

for each cell ω_i with $i \in [\![1, N_\omega]\!]$ **do**
  Use the connectivity table $C_{\omega,e}$ to determine the indices e_1, e_2 and e_3 of the edges delimiting ω_i.
  $h \leftarrow (d(e_1) + d(e_2) + d(e_3))/2$ ▷ Compute the half-perimeter
  $S \leftarrow \sqrt{h(h - d(e_1))(h - d(e_2))(h - d(e_3))}$ ▷ Compute the surface with Heron's formula

  Use the connectivity table $C_{\omega,n}$ to determine the indices n_1, n_2 and n_3 of the nodes delimiting ω_i.
  for each node n_i in n_1, n_2, n_3 **do**
    $m(n_i) \leftarrow m(n_i) + S/2$ ▷ Add the contribution of this cell to the mass field
  end for

  $\kappa_{x,x}(\omega_i) \leftarrow (\kappa_{x,x}(n_1) + \kappa_{x,x}(n_2) + \kappa_{x,x}(n_3))/3$
  $\kappa_{x,y}(\omega_i) \leftarrow (\kappa_{x,y}(n_1) + \kappa_{x,y}(n_2) + \kappa_{x,y}(n_3))/3$
  $\kappa_{y,y}(\omega_i) \leftarrow (\kappa_{y,y}(n_1) + \kappa_{y,y}(n_2) + \kappa_{y,y}(n_3))/3$

  for each edge e in $\{e_1, e_2, e_3\}$ **do** ▷ Compute contribution of ω to the stiffness on each edge
    Using $C_{e,n}$, determine which node among n_1, n_2, n_3 is the apex node (opposed to e), denoted n_{apx}.
    Using $C_{e,n}$, determine which nodes $n_e^{(1)}, n_e^{(2)}$ among n_1, n_2, n_3 are on both ends of e.
    Fetch the coordinates $(x_{\text{apx}}, y_{\text{apx}})$, $(x_e^{(1)}, y_e^{(1)})$, and $(x_e^{(2)}, y_e^{(2)})$.
    $\nabla\Phi_{n_e^{(1)}} \leftarrow \frac{1}{2S}\begin{pmatrix} y_e^{(1)} - y_{\text{apx}} \\ x_{\text{apx}} - x_e^{(1)} \end{pmatrix}$
    $\nabla\Phi_{n_e^{(2)}} \leftarrow \frac{1}{2S}\begin{pmatrix} y_{\text{apx}} - y_e^{(2)} \\ x_e^{(2)} - x_{\text{apx}} \end{pmatrix}$
    $a(e) \leftarrow a(e) - S \times \nabla\Phi_{n_e^{(1)}}^{\mathrm{T}} \begin{pmatrix} \kappa_{x,x}(\omega) & \kappa_{x,y}(\omega) \\ \kappa_{x,y}(\omega) & \kappa_{y,y}(\omega) \end{pmatrix} \nabla\Phi_{n_e^{(2)}}$ ▷ Contribution to the stiffness field at edge e
  end for
  If working on multiple processors, update the value of the mass field m in the halo.
end for

Algorithm 2 Product of a vector $\boldsymbol{\eta}$ with $\mathbf{I} + \overline{\mathbf{M}}^{-1/2}\mathbf{A}\overline{\mathbf{M}}^{-1/2}$

If working on multiple processors, update the value of $\boldsymbol{\eta}$ in the halo.

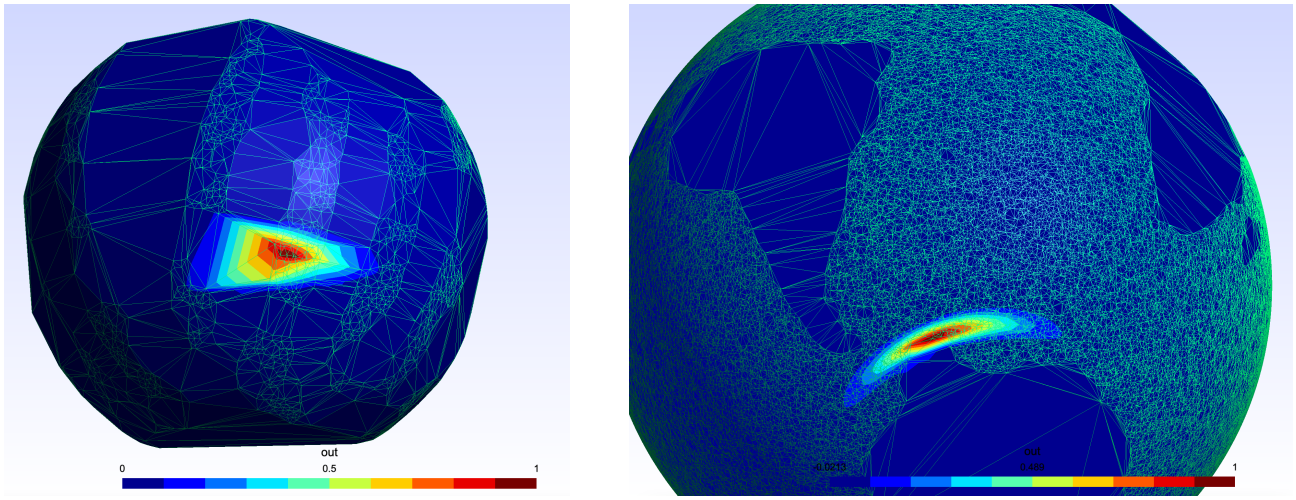
for each edge e_i with $i \in [\![1, N_e]\!]$ **do**
  Extract from the connectivity table the indices n_1, n_2 of the nodes at both ends of the edge e_i.
  $f(e_i) \leftarrow a(\omega)\left(\frac{\eta_{n_1}}{\sqrt{m_{n_1}}} - \frac{\eta_{n_2}}{\sqrt{m_{n_2}}}\right)$; ▷ Compute the flux through the edge
end for

for each edge e_i with $i \in [\![1, N_e]\!]$ **do**
  $\eta_{n_1} \leftarrow \eta_{n_1} - \frac{f(e_i)}{\sqrt{m_{n_1}}}$; ▷ Apply the flux to the first node
  $\eta_{n_2} \leftarrow \eta_{n_2} + \frac{f(e_i)}{\sqrt{m_{n_2}}}$; ▷ Apply the flux to the second node
end for

6.2.2 Meshing in 2D

One of the main challenges from moving from 1D FEM to 2D FEM is the generation of the mesh. With 1D data, the mesh is simply a 'chain' linking observation locations adjacent in time. Even though the distribution of the observations across processors can make the task more complex, building the mesh is straightforward. In 2D, however, building a triangulation linking observation locations requires more advanced tools. While many code libraries allow the construction of triangulations, few can do so based on nodes already distributed across multiple processors. The open-source library ATLAS [3], developed at ECMWF, is able to do so. Figure 6.7 shows examples of such triangulations based on two different datasets of Sea Surface Temperature (SST).

The mesh of Figure 6.7a is defined from one day of SST data from the Atmospheric Sounding Interferometer (IASI) data from the METOP B/C satellites. The denser mesh of Figure 6.7b is from the OSTIA SST data [Good et al., 2020], a global, daily, gridded SST product based on both satellite and in situ data. An anisotropic correlation function modelled with diffusion is shown on each mesh. They have been computed using a parallel Fortran prototype of a diffusion-based correlation operator implemented using the data structures and functionalities provided by ATLAS. This prototype confirms that ATLAS is an appropriate tool to handle the geometry of the observations in 2D and to facilitate exchanges between processors based on the definition of the mesh. This prototype code remains to be integrated in NEMOVAR for further experiments.



(a) Dataset: IASI SST

(b) Dataset: OSTIA SST

Figure 6.7: Anisotropic correlation function modelled with a diffusion operator on meshes defined by SST observation locations. The meridional length scale is 100 km and the zonal length scale is 1000 km. The number of diffusion iterations is set to 2. The cells are represented as by flat triangles in-between the nodes (without additional geographical information), which explains the rough aspect of the surface of the Earth in areas of sparse data.

Handling SWOT-like data with ATLAS is more difficult. The issue is that, as shown in Section 3.3, we expect the observation error correlations affecting SWOT data to be primarily confined to the path of the satellite track. For example, the error between two observations close in space should not be correlated if the observations have been taken during different satellite passes and are thus distant in time. This is linked to the nature of correlated error sources highlighted in Section 3.3: they depend on the satellite attitude or thermal environment (which evolve slowly along the satellite track) rather than on the absolute position of the observations. On the other hand, ATLAS is designed to handle grids and meshes on the sphere. Provided with a

[3] `https://sites.ecmwf.int/docs/atlas/`)

sample of 10 hours of SWOT data, the mesh naturally produced by ATLAS is the one shown in Figures 6.8a and 6.8c. Observations from different satellite passes are linked together, creating elongated triangles between the different satellite swaths. The meshes associated with different satellite passes are conjoined, leading to strong modelled correlations between these different passes.

The behaviour we would like to obtain instead is the one observed in Figures 6.8b and 6.8d. Here the connectivity of the mesh has been computed offline using the Python library Scipy[4] and loaded in ATLAS. The mesh now follows the swath of the satellite (*i.e.*, the domain can be visualised as an unrolled ribbon or a long rectangle) and correlations functions are constrained to a single satellite pass even when they overlap. As both ATLAS and Scipy are based on the C++ meshing program Qhull, it should be possible for ATLAS to generate 'swath' meshes instead of 'sphere meshes' (Willem Deconinck, personal communication). This would allow the mesh to be generated online by ATLAS, at the beginning of a assimilation cycle, to adapt to the new observation network. Whether data from instruments on other polar-orbiting satellites should be meshed on their swath or on the sphere (such as IASI in Figure 6.7a) remains to be evaluated as it depends on the nature of the correlated errors affecting them.

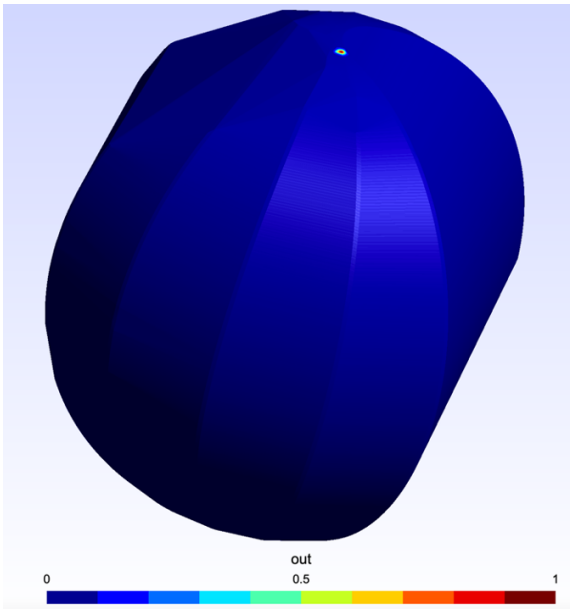
6.2.3 Summary

The developments in NEMOVAR to model observation error correlations for nadir altimeter data with a diffusion operator can be extended to other observation types. Efforts have been made to make the code developments as generic as possible for 1D and 2D applications. To this end, the code implementing the application to a vector of the observation error correlation operator, its inverse, or its square root is entirely generic to 1D and 2D. The diffusion operator is represented by a connectivity table indicating which nodes are at the ends of each edge, a stiffness field defined on each edge, and a mass field defined on each node. These quantities are produced once during the initialisation of the diffusion operator. The initialisation requires the same basic steps for each observation type but are implemented differently depending on the dimensionality of the data. The step that differs the most between the 1D and 2D data is the meshing step. Building the mesh is straightforward in 1D but requires more advanced methods in 2D. We have explored the possibility of using the ATLAS software for this purpose, and have shown with a prototype code that it can be used effectively for correlation modelling on the sphere. Work to extend the capacity of ATLAS to mesh on other domains, such as a satellite swath, is ongoing.

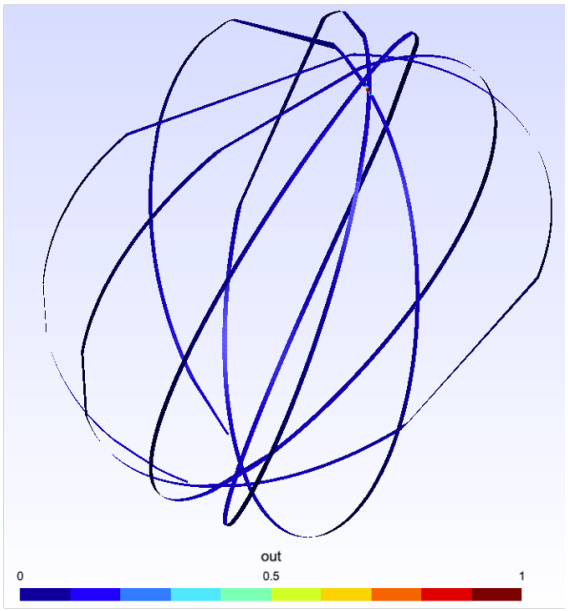
Beyond 2D correlations, the diffusion operator could be adapted to three-dimensional (3D) correlations on a 3D mesh, either to add a third spatial dimension for observations at different depths, or a time dimension to account for time correlations in surface observations. There is no theoretical barrier preventing Algorithm 1 to be adapted to 3D finite elements, and Algorithm 2 is independent of the dimensionality of the data. The main limitation would be the computational cost of the operator, which could be reduced as done for background error modelling by approximating the 3D diffusion by interleaved steps of 2D (horizontal) and 1D (vertical) diffusion. Mixing dimensions with different units (*i.e.*, space $\times$ time) is possible by using a distance measure in the assembly that accounts for both space and time.

The implementation of the diffusion operator based only on mass and stiffness fields could also allow the mesh to be modified to allow correlations that would not be represented by Matèrn-like correlation functions. For example, if two remote locations were correlated, an edge could be manually added to the mesh and attributed a stiffness value whose amplitude would control the strength of these correlations. More generally, triangulations and FEM provide an interesting way to define the connectivity, mass and stiffness fields due to their theoretical link to Matèrn functions, but there is nothing preventing us from adding edges and/or altering the stiffness

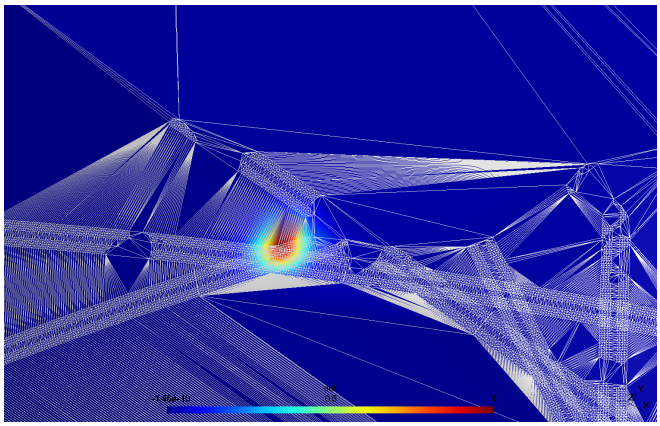
[4] Scipy.spatial.Delaunay



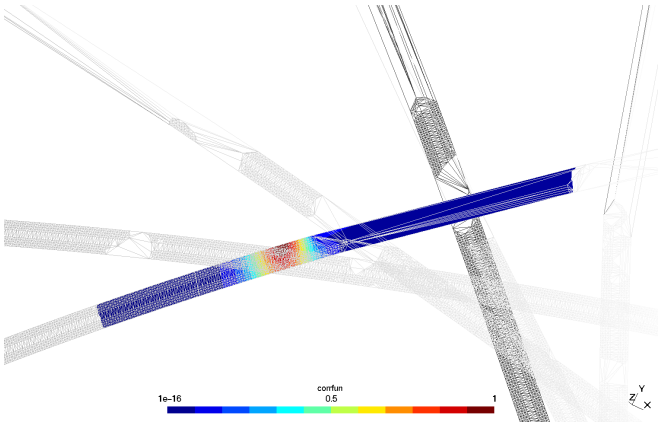
(a) Mesh on the sphere



(b) Mesh on the swath



(c) Mesh on the sphere (zoom)



(d) Mesh on the swath (zoom)

Figure 6.8: Correlation functions modelled with a diffusion operator using the same SWOT dataset but with two different meshes: one defined on the sphere in panels (a) and (c), and one defined on the swath in panels (b) and (d).

field with another method. This can be done as the formulation of the diffusion operator enforces symmetry and positive definiteness independently of the definition of the connectivity, stiffness, and mass. However, these additional degrees of freedom would only prove valuable if new methods are developed to tune them.

6.3 Convergence of the Conjugate Gradient beyond the condition number

6.3.1 Limitations of the condition number

In Chapter 4, we studied the convergence properties of the CG algorithm when applied to a linear system of the form

$$\mathbf{S}\delta\mathbf{u}^* = \mathbf{U}^{\mathrm{T}}\mathbf{b}$$

with

$$\begin{aligned}\mathbf{S} &= \mathbf{I} + \mathbf{U}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{U},\\ \mathbf{b} &= \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x}_{\mathrm{b}})).\end{aligned}$$

To do so, we focussed on the condition number of $\mathbf{S}$, denoted $\kappa(\mathbf{S})$. However, $\kappa(\mathbf{S})$ offers only a partial understanding of the convergence properties of CG. A more in-depth study of the convergence properties of CG requires understanding the effect of the distribution of the eigenvalues of $\mathbf{S}$ within the range of its extreme eigenvalues, and the effect of the right-hand side of the system, as explained in this section.

As in Chapter 4, let $(\lambda_i(\mathbf{S}))_{i\in[\![1,n]\!]}$ denote the eigenvalues of $\mathbf{S}$, and $(\mathbf{f}_i)_{i\in[\![1,n]\!]}$ the corresponding eigenvectors. The effect of the right-hand side on the system appears through the initial residual

$$\mathbf{r}_0 = \delta\mathbf{u}^{(0)} - \mathbf{U}^{\mathrm{T}}\mathbf{b}$$

where $\delta\mathbf{u}^{(0)}$ is the initial guess. Let $(\eta_i)_{i\in[\![1,n]\!]}$ represent the components of $\mathbf{r}_0$ for the basis defined by the eigenvectors of $\mathbf{S}$:

$$\eta_i = \mathbf{f}_i^{\mathrm{T}}\mathbf{r}_0, \quad i \in [\![1,n]\!].$$

The iterate $\delta\mathbf{u}^{(\ell)}$ approximating the solution $\delta\mathbf{u}^*$ at the ℓ-th iteration has the key property [Saad, 2003, Lemma 6.28]

$$\|\delta\mathbf{u}^* - \delta\mathbf{u}^{(\ell)}\|_{\mathbf{S}}^2 = \min_{p\in\mathcal{P}_\ell(0)} \|p(\mathbf{S})(\delta\mathbf{u}^* - \delta\mathbf{u}^{(0)})\|_{\mathbf{S}}^2 = \min_{p\in\mathcal{P}_\ell(0)} \sum_{i=1}^{n} p(\lambda_i(\mathbf{S}))^2 \frac{\eta_i^2}{\lambda_i(\mathbf{S})}, \tag{6.14}$$

where $\mathcal{P}_\ell(0)$ is the space of polynomials of degree ℓ that have a value of one at zero. The roots of the polynomial solution of the minimisation problem in Equation (6.14) are known as Ritz values, and can be used as an approximation of the eigenvalues of $\mathbf{S}$.

Equation (6.14) describes exactly the evolution of the Hessian norm of the error at each iteration of CG. Understanding the process leading from Equation (6.14) to the classical upper bound defined in terms of $\kappa(\mathbf{S})$ allows us to understand why and when this bound can be pessimistic. First, we use the bound

$$\|\delta\mathbf{u}^* - \delta\mathbf{u}^{(\ell)}\|_{\mathbf{S}}^2 \leq \min_{p\in\mathcal{P}_\ell(0)} \left[\max_{i\in[\![1,N]\!]} p(\lambda_i(\mathbf{S}))^2\right] \sum_{i=1}^{n} \frac{\eta_i^2}{\lambda_i(\mathbf{S})} = \min_{p\in\mathcal{P}_\ell(0)} \left[\max_{i\in[\![1,N]\!]} p(\lambda_i(\mathbf{S}))^2\right] \|\delta\mathbf{u}^* - \delta\mathbf{u}^{(0)}\|_{\mathbf{S}}^2. \tag{6.15}$$

The main approximation in this step is the removal of the effect of the initial residual on the convergence, as the selection of the polynomial p no longer depends on the factors η_i. It can lead to a strong contribution to the discrepancy between the actual error and the bound as the latter will ignore any potential advantageous properties of the initial residual.

Figure 6.9 provides a schematic illustration of the evolution of the polynomial appearing in the upper bound

(6.15). At iteration 0, the polynomial is constant and equal to 1 at all eigenvalue locations, meaning that the upper bound is exact before the minimisation starts. On subsequent iterations, the polynomial is linear, quadratic, cubic, *etc.*, reducing progressively its maximum value at eigenvalue locations.

This simple schematic clearly highlights two classical properties of CG. The first is that CG converges in, at most, N iterations. Indeed, at the N-th iteration, the polynomial is of degree N and can have a root equal to each eigenvalue, which guarantees full convergence in exact arithmetic according to Equation (6.15). In practice, this property is rarely maintained in finite precision arithmetic. The second property is that, if there are N_c 'clusters' of eigenvalues, then CG 'almost fully' converges in N_c iterations or less. The reason is that, if some eigenvalues are clustered (*i.e.*, have approximately the same value), then having a single root near them is sufficient for the polynomial to be almost 0 for all of them. This is illustrated in Figure 6.9 where the convergence is almost complete at the third iteration as two eigenvalues are clustered.

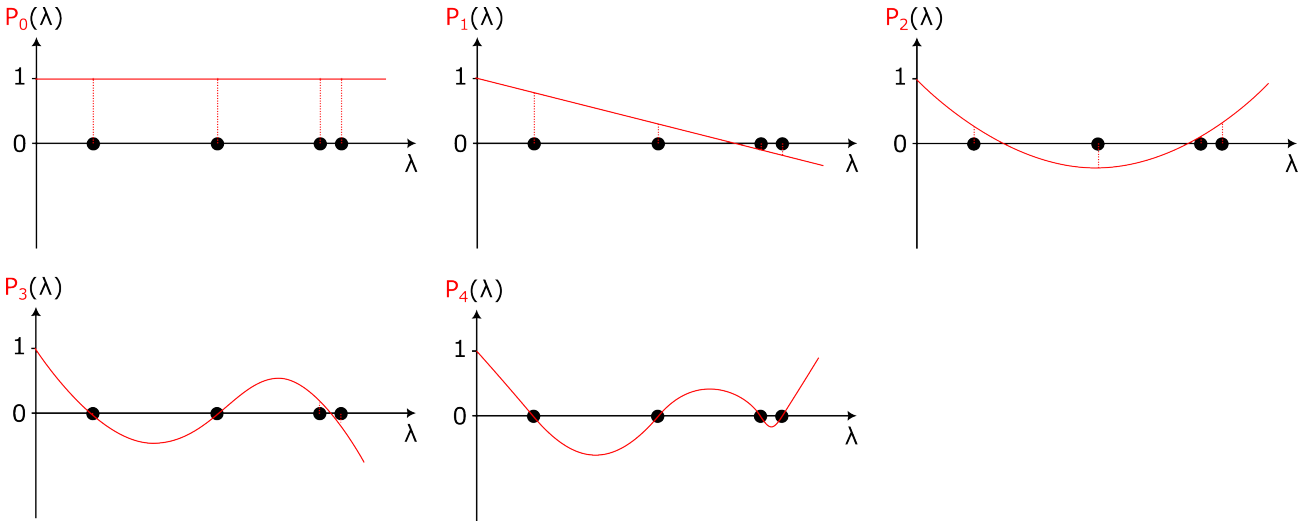


Figure 6.9: Schematic illustrating the evolution of the polynomial in the upper bound (6.15) during 4 iterations of CG applied to a 4×4 system. The abscissa indicates the amplitude of the eigenvalues, marked by black dots.

In order to obtain a bound in terms of the condition number, another approximation is required. The problem of finding the polynomial of minimal amplitude at discrete points defined by the eigenvalues is too complex to be solved analytically in the general case, but can be approximated by the simpler problem of finding the polynomial of minimal amplitude on a continuous domain bounded by the extreme eigenvalues:

$$\min_{p \in \mathcal{P}_\ell(0)} \left[\max_{i \in [\![1,N]\!]} p(\lambda_i(\mathbf{S}))^2 \right] \leq \min_{p \in \mathcal{P}_\ell(0)} \left[\max_{\lambda \in [\lambda_{min}(\mathbf{S}), \lambda_{max}(\mathbf{S})]} p(\lambda)^2 \right]. \tag{6.16}$$

The continuous problem is simple enough to be solved analytically, as the bound in (6.16) is minimised by the Chebyshev polynomials [Saad, 2003, Section 6.11.3], which leads to the well-known bound in terms of the condition number:

$$\|\delta\mathbf{u}^* - \delta\mathbf{u}^{(\ell)}\|_{\mathbf{S}}^2 \leq 2 \left(\frac{\sqrt{\kappa(\mathbf{S})} - 1}{\sqrt{\kappa(\mathbf{S})} + 1} \right)^\ell \|\delta\mathbf{u} - \delta\mathbf{u}^{(0)}\|_{\mathbf{S}}^2. \tag{6.17}$$

To obtain the bound in Equation (6.17) from the bound in Equation (6.16), the distribution of the eigenvalues in-between their extreme values is ignored. The case where this approximation would not generate a large discrepancy would be when the eigenvalues are distributed uniformly and densely between their extreme values. On the other hand, the upper bound would be pessimistic in the presence of clusters of eigenvalues, which it

does not account for.

We can thus identify two main limitations to the use of the condition number as the sole marker of speed of convergence of CG: it does not reflect the interaction between the initial residual of the system and $\mathbf{S}$ [see *e.g.*, Strakoš and Liesen, 2013], and it does not account for the distribution and potential clustering of the eigenvalues between their extreme values.

6.3.2 Convergence analysis for variational data assimilation

Equation (6.14) could be used directly to understand the convergence of CG to avoid potentially pessimistic approximations. For variational DA, we can expect many eigenvalues of $\mathbf{S}$ to be clustered around one. As mentioned in Chapter 4, if n is the dimension of the state space and p is the dimension of the observation space, then $n - p$ eigenvalues of $\mathbf{S}$ are equal to one. This property is ignored by bounds based solely on the condition number, which implicitly assumes all of these eigenvalues to be distributed uniformly between the extreme eigenvalues.

To examine this further, we consider as an example the idealized problem of Chapter 5, where the covariance matrices are assumed to be circulant. In addition, we assume that $\mathbf{H}$ is the identity matrix for the sake of simplicity. In this case, $\mathbf{S}$ is also circulant, and its eigenvectors $(\mathbf{f}_i)_{i\in[\![1,N]\!]}$ can be interpreted as Fourier basis functions (see Section 3.1 of Goux et al. [2024] in Chapter 4). Each eigenpair of $\mathbf{S}$ can then be associated with a spatial scale as discussed in Chapter 5. The scalars $(\eta_i)_{i\in[\![1,n]\!]}$ represent the Fourier transform of the initial residual $\mathbf{r}_0$. In addition to describing the convergence of the norm of the error, Equation (6.14) can also indicate for which spatial scales the error decreases first.

First, in Figure 6.10, we consider the case where the background error is correlated ($\rho_\mathrm{b} = 250$ km, $m_\mathrm{b} = 10$) and the observation error is uncorrelated (diagonal $\mathbf{R}$). The domain considered is a circle with a perimeter of 40000 km with $n = 1440$ grids points (the domain can be interpreted as a latitude circle at the Equator with a grid point every quarter degree of longitude). The covariance function of $\mathbf{B}$ is quasi-Gaussian ($m_\mathrm{b} = 10$) with a 'Stein' length scale of 250 km. The cumulative distribution of the eigenvalues of $\mathbf{S}$ – that is, the percentage of the eigenvalues smaller than λ for $\lambda \in [\lambda_{\min}(\mathbf{S}), \lambda_{\max}(\mathbf{S})]$) – is shown in the top panel. It clearly illustrates that the eigenvalues are far from being uniformly distributed: even with $\mathbf{H} = \mathbf{I}$, more than 90% of the eigenvalues are located between 1 and 1.0025, while the maximum eigenvalue is 45. As seen in the cumulative distribution built from the first Ritz value, or from P_1 in the bottom panel, the first Ritz value is close to the maximum eigenvalue. There are two main reasons for this.

The first reason is because of a generic property of CG, which tends to prioritise directions associated with large eigenvalues. This can be appreciated by visualising the polynomial of degree 1 minimizing (6.14) at the first iteration. As the polynomial is constrained to have unit amplitude at the origin, if the sole Ritz value is located near the largest eigenvalue, then the polynomial has a smaller absolute value than before the first iteration over the whole range of eigenvalues. On the other hand, if the Ritz value is located at the smallest eigenvalue which are closer to 1, then the polynomial would have a large absolute value at the location of large eigenvalues. It is thus more difficult to reduce the error in directions associated with smaller eigenvalues.

The second reason is more specific to our application. As shown in Figure 6.10c, the quantity $\eta_i^2/\lambda_i(\mathbf{S})$ is extremely close to zero around 1 (up to machine precision). This means that the scalars η_i, which represent the Fourier decomposition of the initial residual, are close to zero for all eigenvalues close to 1. To help interpret this property, Figure 6.10c also displays the spatial scale between 0 and 1, where 0 and 1 indicate the smallest and largest spatial scale resolved by the grid, respectively. The indicator is defined as i/n where n is the dimension of the system. It can be seen as a normalized, non-dimensional, counterpart of the spatial scale used throughout Chapter 5 (*e.g.*, in Figure 1 of Chapter 5). This indicator clearly shows that, in this configuration, large (small)

eigenvalues of $\mathbf{S}$ are associated with large (small) spatial scales. Consequently, η_i being close to zero for small eigenvalues means that the Fourier decomposition of $\mathbf{r}_0$ is close to zero at small spatial scales. In other words, the vector $\mathbf{r}_0$ is 'smooth'. As $\mathbf{r}_0$ is close to zero at small spatial scales, the system is 'already solved' for the associated directions, and they do not contribute to the difference between the initial guess $\delta\mathbf{u}^{(0)}$ and the solution $\delta\mathbf{u}^*$ in Equation (6.14). There is thus no strong incentive for a Ritz value to be placed at small spatial scales, as CG will first focus on large spatial scales where the error is larger.

The successive polynomials in Figure 6.10b show that the initial error is first reduced at large spatial scales by CG before it reduces the error at progressively smaller spatial scales, which have a lower contribution to the Hessian norm of the analysis error. The stopping criterion is reached when the norm of the residual normalised by the norm of the right-hand side is smaller than 10^{-6}, which is achieved in 40 iterations.

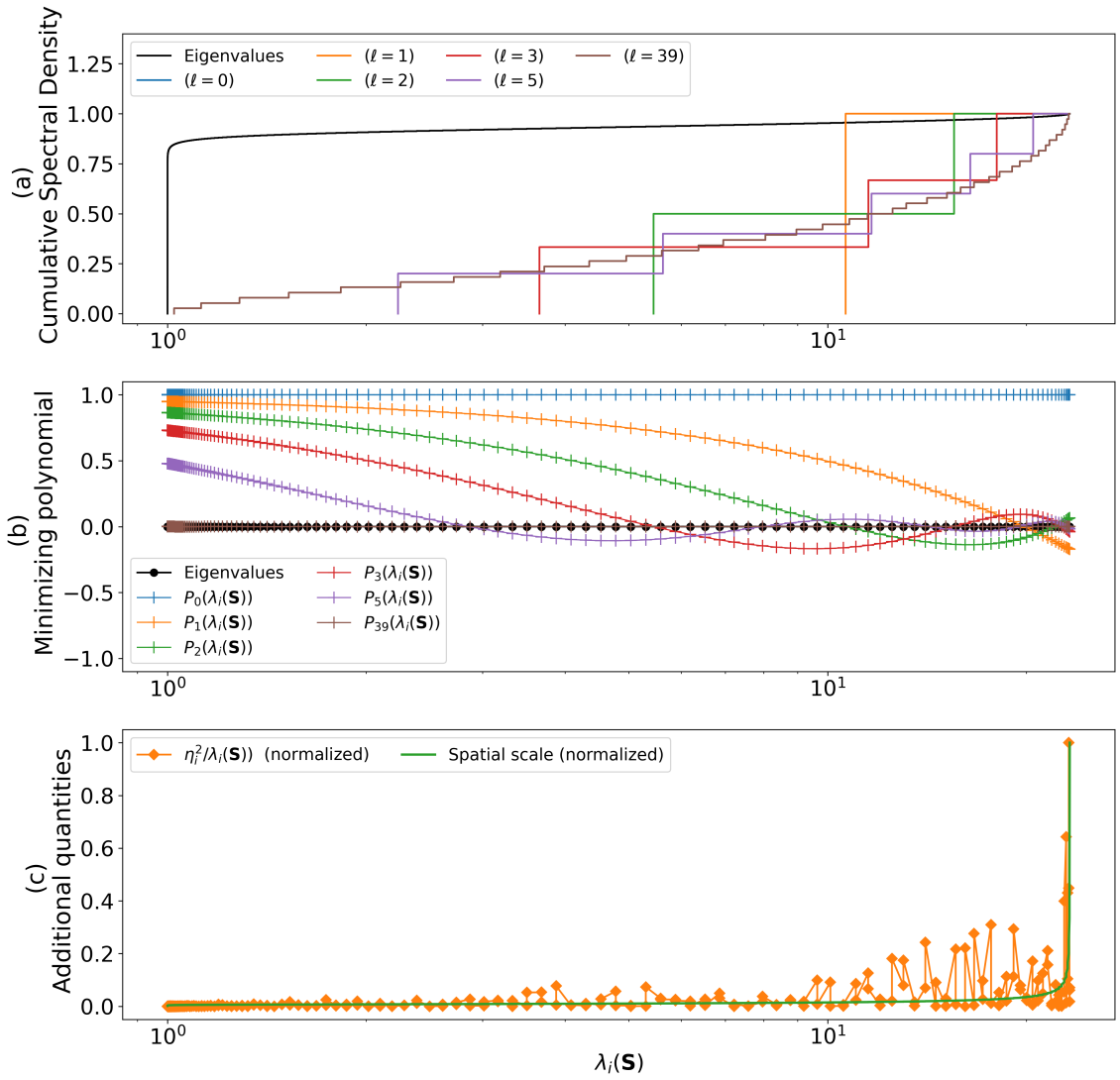


Figure 6.10: (a): Cumulative spectral density of the eigenvalues of $\mathbf{S}$, and its approximation based on Ritz values at different iterations of CG. (b): Polynomials satisfying Equation (6.14) over the range of the eigenvalues of $\mathbf{S}$, whose positions are marked by black dots on the x-axis. (c): Weights $\eta_i^2/\lambda_i(\mathbf{S})$ attributed to each eigenvalue in the minimised sum in Equation (6.14) normalised by their maximum to fit between 0 and 1 (orange triangles), and spatial scale of the associated eigenvector, also normalised between 0 and 1 (green curve). In this example, $\mathbf{R}$ is diagonal.

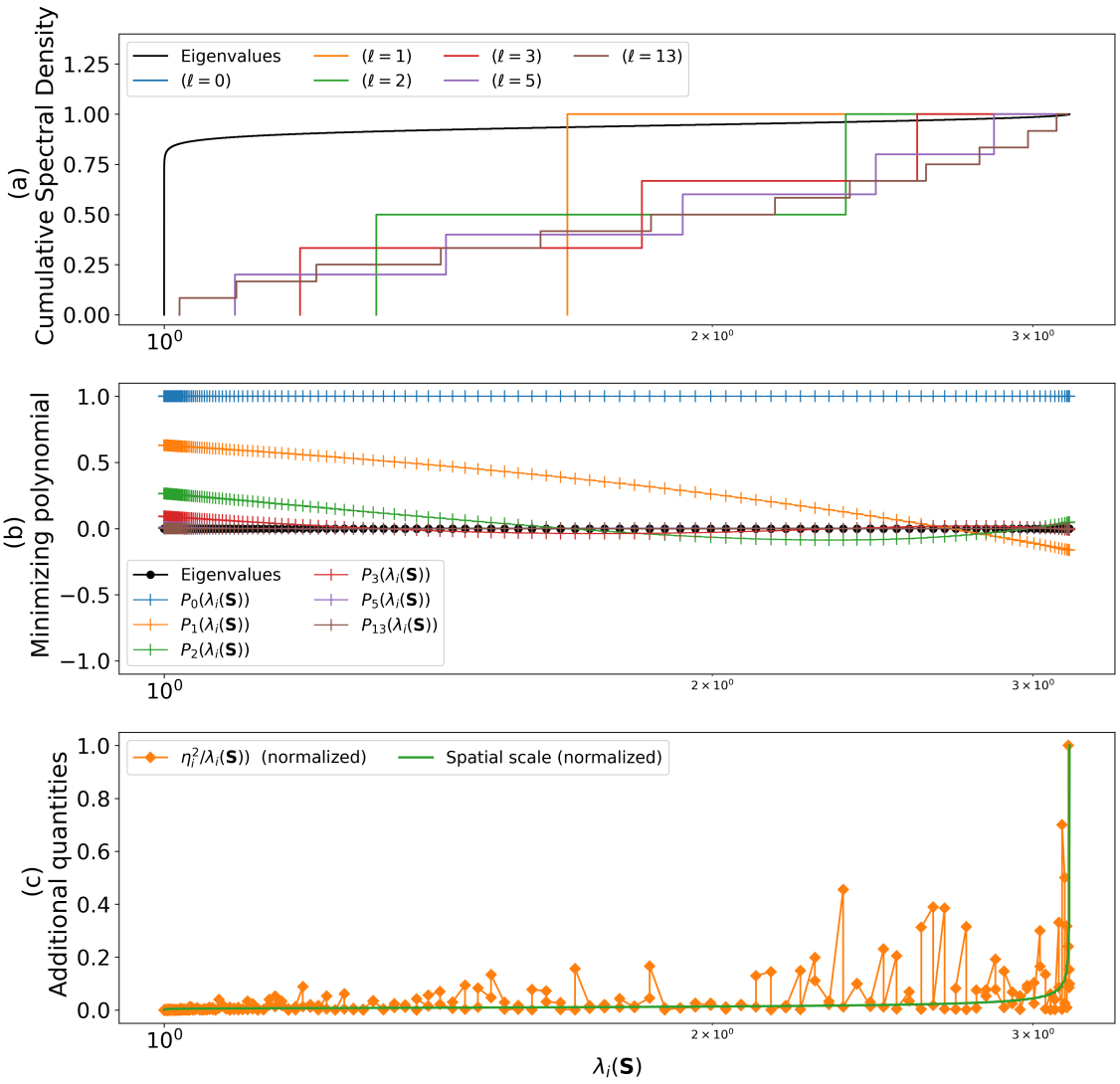


Figure 6.11: (a): Cumulative spectral density of the eigenvalues of $\mathbf{S}$, and its approximation based on Ritz values at different iterations of CG. (b): Polynomials satisfying Equation (6.14) over the range of the eigenvalues of $\mathbf{S}$, whose positions are marked by black dots on the x-axis. (c): Weights $\eta_i^2/\lambda_i(\mathbf{S})$ attributed to each eigenvalue in the minimised sum in Equation 6.14 normalised by their maximum to fit between 0 and 1 (orange triangles), and spatial scale of the associated eigenvector, also normalised between 0 and 1 green curve). In this example, $\mathbf{R}$ is defined by a SOAR function ($m_o = 2$) with a length scale half that of the background error ($\rho_o = 125$ km).

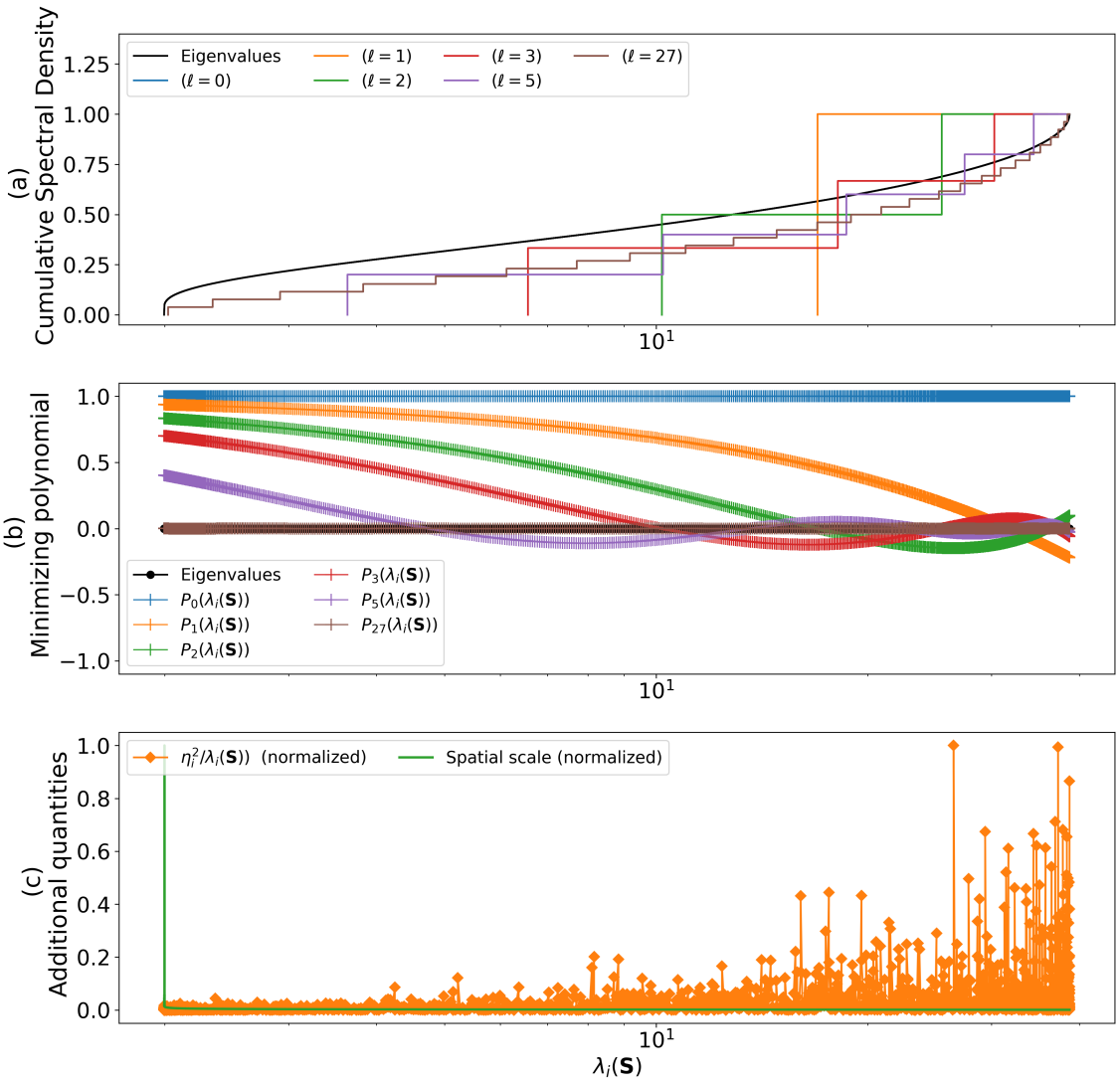


Figure 6.12: (a): Cumulative spectral density of the eigenvalues of $\mathbf{S}$, and its approximation based on Ritz values at different iterations of CG. (b): Polynomials satisfying Equation (6.14) over the range of the eigenvalues of $\mathbf{S}$, whose positions are marked by black dots on the x-axis. (c): Weights $\eta_i^2/\lambda_i(\mathbf{S})$ attributed to each eigenvalue in the minimised sum in Equation 6.14 normalised by their maximum to fit between 0 and 1 (orange triangles), and spatial scale of the associated eigenvector, also normalised between 0 and 1 (green curve). In this example, $\mathbf{R}$ is defined by a quasi-Gaussian function with $m_\mathrm{o} = 12 > m_\mathrm{b} = 10$ but with the same length scale as the background error ($\rho_\mathrm{o} = \rho_\mathrm{b} = 250$ km).

Figure 6.11 is the counterpart of Figure 6.10 but with a non-diagonal $\mathbf{R}$, where the correlation function of the observation error is a SOAR function with a length scale half of that of the background error ($\rho_\mathrm{o} = 125$ km, $m_\mathrm{o} = 2$). In agreement with results in Goux et al. [2024], the condition number obtained with this non-diagonal $\mathbf{R}$ is 3, which is lower than the value of 45 obtained with a diagonal $\mathbf{R}$. Note that such low values are rarely obtained with $\mathbf{H} \neq \mathbf{I}$ but the same decrease on the condition number is observed. Beyond this smaller range,

the density of the eigenvalues and distribution of the components η_i remain roughly similar to the experiment with a diagonal $\mathbf{R}$, leading to a faster convergence with a similar behaviour where the error decreases first at large spatial scales.

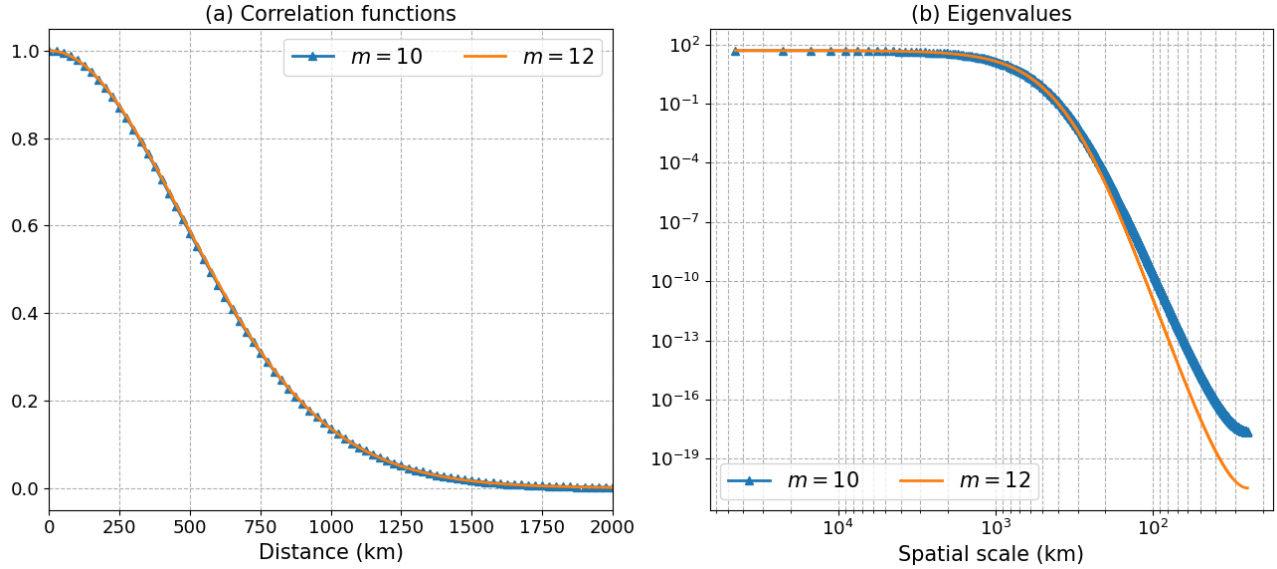


Figure 6.13: (a) AR functions and (b) their respective Fourier transforms. The length scale is 250 km, and the smoothness parameters are 10 and 12. The two functions are almost completely overlapping in physical space but their difference appears clearly in spectral space.

The poor convergence properties with values of m_o larger than m_b observed in Chapter 4 are reproduced in Figure 6.12. While this configuration is associated with a large condition number, Figure 6.12 shows that there are other reasons for the slow convergence. The observation error correlation function used in Figure 6.12 is visually close to that the background error correlation function (see Figure 6.13a) apart from a small but important difference. Although it has the same length scale ($\rho_\mathrm{o} = \rho_\mathrm{b} = 250$ km), its eigenvalue spectrum displays a slightly steeper slope ($m_\mathrm{o} = 12 > m_\mathrm{b} = 10$, see Figure 6.13). With $m_\mathrm{o} = m_\mathrm{b}$, $\mathbf{S} = 2\mathbf{I}$ and convergence would be achieved in a single iteration. However, the slightly steeper slope of the observation error eigenvalue spectrum is sufficient to obtain a radically different behaviour, as the algorithm now needs 82 iterations to meet the stopping criterion.

Panel (a) of Figure 6.12 shows that, contrary to the previous configurations, the eigenvalues of $\mathbf{S}$ cover a larger range and are also spread relatively uniformly within this range. As in the previous configurations, the Ritz values are initially located near the largest eigenvalues and slowly become closer to the smaller eigenvalues as the iterations proceed. However, contrary to the two previous configurations, large eigenvalues are now associated with small spatial scales, and small eigenvalues with large spatial scales. A possible interpretation of this property is that the 'roughening' effect of $\mathbf{R}^{-1}$ dominates the 'smoothing' effect of $\mathbf{U}$ and $\mathbf{U}^\mathrm{T}$ in the $\mathbf{B}$-preconditioned Hessian matrix $\mathbf{S}$. This means that the CG algorithm resolves the small spatial scales first in this case. This major difference with the two previous configurations likely explains why the convergence of CG is particularly slow in the early iterations of a similar configuration in Figure 8 from Goux et al. [2024] (green curve). The metric used in that example was the RMS of the difference between the analysis obtained by truncating the minimisation at the iteration ℓ and the true state $\mathbf{x}^{(\mathrm{t})}$. The small spatial scales are not those that contribute the most to this difference as the background and observations are both accurate at the small spatial scales due to their strongly correlated errors. In other words, fitting the background instead of the observations, or the reverse, at these spatial scales will not substantially bring the analysis closer to the true state.

6.3.3 Summary

The condition number of $\mathbf{S}$ is a convenient diagnostic for studying the convergence rate of the $\mathbf{B}$-preconditioned CG algorithm. However, summarising the convergence properties of CG using a single number has limitations. The first limitation is that many systems can share the same condition number while leading to very different convergence rates. This means that an upper bound on the convergence rate of CG based on the condition number can be pessimistic in many configurations. Therefore, we have to be cautious when interpreting the results from Chapter 4, or from other studies, which are based solely on the condition number. A small condition number does indicate a fast convergence, but a large condition number does not necessarily indicate a slow convergence rate.

Furthermore, with the condition number alone, we lose insight on the behaviour of CG between the first and last iteration. This can be problematic in data assimilation as the minimisation of the inner loop is usually truncated early to limit computational cost. On the other hand, analysing the convergence of the $\mathbf{B}$-preconditioned CG as done in Figures 6.10, 6.11 and 6.12 is more informative. In a very simple configuration, we were able to observe a radically different behaviour depending on how the amplification of the signal at small spatial scales by $\mathbf{R}^{-1}$ is matched by their damping by $\mathbf{B}$. For example, if the spectral decay rate of the observation error correlation function is steeper than that of the background error correlation function then the $\mathbf{B}$-preconditioned CG prioritises a correction to the small spatial scales before the large spatial scales. This would be a major concern if CG is truncated before correcting the spatial scales where the background and observation error are actually significant.

Conversely, with a diagonal $\mathbf{R}$, the $\mathbf{B}$-preconditioned CG prioritises large spatial scales over small spatial scales in the early iterations. A similar behaviour occurs with a non-diagonal $\mathbf{R}$ when the correlations have a smaller length scale and/or slower spectral decay rate than those of $\mathbf{B}$. This property is of particular interest in data assimilation, as it means that the spatial scales where the background and observation error are largest are targeted first by CG. In these configurations, not only would CG converge relatively quickly, but most of the relevant correction would be achieved in the early iterations before CG is potentially truncated.

The diagnostics presented in this section could be applied to more complex systems. For example, the assumption of $\mathbf{H} = \mathbf{I}$ is very restrictive. If there are fewer observations than grid points ($p < n$), then $n - p$ eigenvalues are exactly equal to 1, which could influence the convergence. Extending this study to selection operators as in the idealised system of Chapters 4 and 5 could be a first step. Beyond idealised systems, Ritz values can be extracted from the inner loop of NEMOVAR in a realistic configuration and visualised as in Figures 6.10, 6.11 and 6.12. A difficulty with this approach would be the lack of theoretical knowledge on the eigenvectors of $\mathbf{S}$. Building explicitly these eigenvectors would be computationally challenging, and an alternative method would be needed to estimate the weights $(\eta_i)_{i\in[\![1,n]\!]}$. Even if these eigenvectors were available, it would not be straightforward to associate them to a spatial scale, unlike in the idealised system where an analytical relation is available. A heuristic indicator could be derived either by fitting known analytical Fourier modes or based on the number of zero-crossings of the eigenvectors, but the relevancy of these indicators would need to be thoroughly evaluated.

Another component of the optimisation process which was not considered here is the potential interactions of a non-diagonal $\mathbf{R}$ with second-level preconditioners, which are used in addition to $\mathbf{B}$ (which would constitute the 'first-level'). For example, a limited memory preconditioner [LMP, see *e.g.*, Morales and Nocedal, 2000] can be used to leverage information from the solution of a similar system in the previous iteration of the outer loop to precondition the system of the current outer loop iteration. Diouane et al. [2024] also use Equation (6.14) to analyse the convergence of CG in the context of data assimilation and design a spectral scaling of their LMP to accelerate the convergence. The presence of a non-diagonal $\mathbf{R}$ could affect the choice of parameters of this

preconditioner.

Chapter 7

Conclusions

This manuscript has covered different aspects of correlated observation error in variational data assimilation. First, the need to account for correlated observation error has been illustrated in Chapter 3 with the practical example of wide-swath altimetry, which provided a strong motivation for this thesis. Based on the simulated SWOT mission error budget, we identified error correlations whose length scale could exceed that of the background error in high-resolution operational configurations. The correlations stem from instrumental errors, such as inaccuracies in the measure and control of the roll angle of the satellite. A review of the literature shows that, while assimilating simulated SWOT data without correlated errors leads to a strong positive impact on the analysis, such levels of performance cannot be reached if the observations are simulated with the full error budget. Current data assimilation systems that neglect spatial correlations in the observation error are thus unable to extract meaningful small-scale information from SWOT data without contaminating the analysis with large correlated errors. This motivates the need to develop methods to account for correlated observation error explicitly in data assimilation systems.

We investigated in Chapter 4 how a non-diagonal observation error correlation matrix ($\mathbf{R}$) could affect the convergence rate of the iterative minimisation methods used to compute the analysis in variational data assimilation, and thus how close $\mathbf{x}_k$, the solution of the truncated minimisation problem, can be from $\mathbf{x}^*$, the solution at full convergence. In particular, we focused on analysing the condition number of the Hessian matrix of the 3D-Var cost function where, following standard practice, the background error covariance matrix $\mathbf{B}$ is used as a first-level preconditioner. The condition number can be used as a first-order indicator of the convergence rate of iterative methods such as the conjugate gradient algorithm. We focused on an idealised system where both the background and observation error correlation matrices are modelled with diffusion operators. This study demonstrated that the condition number is highly sensitive to the parameters of the observation error correlation model relative to those of the background error correlation model. In particular, the condition number can be either increased or decreased compared to the condition number with a diagonal $\mathbf{R}$.

The diffusion operator used to model the observation (background) error correlations is parameterised by a length scale ρ_o (ρ_b) and a smoothness parameter m_o (m_b). If $m_\mathrm{o} \leq m_\mathrm{b}$, increasing ρ_o decreases the condition number as long as ρ_o remain smaller than ρ_b. As having $\rho_\mathrm{o} = 0$ is equivalent to using a diagonal $\mathbf{R}$, the condition number is thus systematically decreased when a non-diagonal $\mathbf{R}$ is used instead of a diagonal $\mathbf{R}$ with $\rho_\mathrm{o} \leq \rho_\mathrm{b}$ and $m_\mathrm{o} \leq m_\mathrm{b}$ (and assuming their variances are the same). In this configuration, there is a double incentive to account for observation error correlations. First, the solution of the minimisation, $\mathbf{x}^*$, becomes statistically closer to the true state $\mathbf{x}^{(\mathrm{t})}$ due to a better representation of the observation error correlations; second, the solution of the truncated minimisation $\mathbf{x}_k$ will likely be closer to $\mathbf{x}^*$ as the convergence rate of the conjugate

gradient algorithm should be improved. On the other hand, increasing ρ_o increases the condition number when ρ_o exceeds ρ_b (while maintaining $m_\mathrm{o} \leq m_\mathrm{b}$). For a large enough ρ_o, the condition number eventually becomes larger with a non-diagonal $\mathbf{R}$ than with a diagonal $\mathbf{R}$. The smaller m_o is compared to m_b, the larger ρ_o can be before that point is reached.

If $m_\mathrm{o} > m_\mathrm{b}$, the condition number with a non-diagonal $\mathbf{R}$ is almost systematically larger than with a diagonal $\mathbf{R}$ by several orders of magnitude (unless ρ_o is extremely small compared to ρ_b). In this case, our experiments show that a prohibitively large number of conjugate gradient iterations is required for $\mathbf{x}_k$ to reach $\mathbf{x}^*$. Consequently, even if actual estimates of the observation error correlations suggest using a large value of m_o (*i.e.*, a quasi-Gaussian function), it is better to use a correlation function with a smaller value of m_o (such as a SOAR function). We showed that intentionally mis-specifying m_o this way greatly improves the convergence rate of the minimisation (and thus the proximity of $\mathbf{x}_k$ to $\mathbf{x}^*$) without degrading the solution at full convergence relative to the truth (the proximity of $\mathbf{x}^*$ to $\mathbf{x}^{(\mathrm{t})}$). These results were published in *Numerical Linear Algebra with Applications* [Goux et al., 2024].

The results from Chapter 4 showed that, with a limited computational budget, the observation error correlation model that most resembles the actual observation error correlations is not necessarily the one that leads to the best solution. In particular, certain adjustments to the observation error correlation model to improve the convergence rate of the minimisation did not necessarily result in a significant degradation of the solution at convergence. In order to understand this behaviour better, we studied in Chapter 5 the sensitivity of the difference between $\mathbf{x}^*$ and a true state $\mathbf{x}^{(\mathrm{t})}$ to potential 'misspecifications' of the observation error correlations, such as an overestimation or underestimation of the error variance, or correlation length scale. We applied a Fourier analysis to an idealised system to study the effects of observation error correlations from a scale-dependent perspective; in particular, by comparing the scale-dependent variances of the observation error (*i.e.*, the eigenvalue spectrum of $\mathbf{R}$) with those of the background error mapped to observation space (*i.e.*, the eigenvalue spectrum of $\mathbf{HBH}^\mathrm{T}$). We showed that, at a given spatial scale, a misspecification of the observation error variance will only significantly affect $\mathbf{x}^*$ if the scale-dependent variances of the background and observation errors are of comparable magnitude. Moreover, overfitting the most accurate source of information (between the observations and the background) is less detrimental to $\mathbf{x}^*$ than underfitting it. Conversely, overfitting the least accurate source of information is more detrimental to $\mathbf{x}^*$ than underfitting it (though the best option is still to do neither). These results proved valuable later in the manuscript to identify in each experimental configuration at which spatial scales the scale-dependent observation error variance needs to be accurate or can afford to be inaccurate.

We used the Fourier analysis from the idealised system to interpret results from NEMOVAR involving the assimilation of simulated sea-surface height (SSH) innovations. The simulated innovations were constructed from different realisations of observation and background error, with known covariances. For NEMOVAR, we developed an error correlation model for nadir altimeter observations using an approach based on a diffusion operator, discretised using a finite-element method in order to handle unstructured observations. Its implementation in NEMOVAR was done in a rigorous and flexible way so that it serves as a basis for future operational use. In particular, it has been designed to be scalable for large datasets in parallel environments, and to be adaptable to other observation types.

With $\rho_\mathrm{o} \leq \rho_\mathrm{b}$ and $m_\mathrm{o} \leq m_\mathrm{b}$, we confirmed the positive impact of using a non-diagonal $\mathbf{R}$ where, compared to a diagonal $\mathbf{R}$, the minimising solution $\mathbf{x}^*$ becomes statistically closer to the true state $\mathbf{x}^{(\mathrm{t})}$, while similar convergence properties are maintained. With $\rho_\mathrm{o} > \rho_\mathrm{b}$ and $m_\mathrm{o} < m_\mathrm{b}$ (which could be the case with SWOT data as discussed in Chapter 3), the situation is more nuanced. In agreement with the theory developed in Chapter 4, if ρ_o is much greater than ρ_b, convergence is slow but not for values of ρ_o that are close to ρ_b. Regardless of whether convergence is slow or not, $\mathbf{x}^*$ is significantly closer to the true state $\mathbf{x}^{(\mathrm{t})}$ with a non-diagonal $\mathbf{R}$. In our

example, we showed that, compared to results with a diagonal $\mathbf{R}$, the truncated solution $\mathbf{x}_k$ was closer to the true state $\mathbf{x}^{(\mathrm{t})}$ at every iteration of the minimisation. This improvement is attributed to the use of a non-diagonal $\mathbf{R}$ wihch allows the extraction of information from the small spatial scales of the observations while rejecting more error-prone features at large spatial scales. One of the consequences of the discrepancy in accuracy between the large and small spatial scales of the observations is that observations with correlated error have more impact on the analysis fields associated with the gradient of the observations than the observations themselves. This was illustrated by the behaviour of zonal geostrophic velocity, which through geostrophic balance is associated with the meridional derivative of the SSH (aligned with the satellite track) is better corrected than the SSH itself when assimilating altimeter data with large error correlations. When $m_{\mathrm{o}} > m_{\mathrm{b}}$, the large increase of the condition number predicted in Chapter 4 is confirmed. From the theoretical analysis and experiments, we have been able to conclude that observation error variance inflation is only successful in approximating the effect of a non-diagonal $\mathbf{R}$ (though still sub-optimal) if the correlation length scale of observation error is small compared to that of the background error ($\rho_{\mathrm{o}} < \rho_{\mathrm{b}}$).

An important finding in our experiments is that a non-diagonal $\mathbf{R}$ with a large condition number can be associated with an extreme sensitivity to perturbations at small spatial scales. The observation error variance at small spatial scales can be unintentionally specified to be excessively small if, for example, a Gaussian function ($m_{\mathrm{o}} \geq 10$) is used to model observation error correlations, or if the correlation length scale of observation error is much larger than that of the background error. The consequence is that the operator associated with $\mathbf{R}^{-1}$ applies an extremely large amplification factor (its maximum eigenvalue) at the smallest spatial scales. If this factor is too large, $\mathbf{R}^{-1}$ can amplify small perturbations to the amplitude of the actual geophysical signals of interest. In our configuration in NEMOVAR, using a non-diagonal $\mathbf{R}$ with $m_{\mathrm{o}} = 10$ thus almost systematically results in large noise in the small spatial scales in the analysis. If we consider a less extreme case with lower values of m_{o} but large values of ρ_{o}, this issue does not occur in optimal conditions (*i.e.*, when $\mathbf{R}$ perfectly describes the actual errors) but can still be triggered by the addition of a small, unaccounted for, white noise in the innovation. In this configuration, the additional white noise is too small to be detected with observation error diagnostics. A conservative strategy would thus be to account for the presence of white noise when specifying observation error correlations as a safety measure. For example, manually enforcing a bound for the minimum eigenvalue is possible for inter-channel error correlations. Reconditioning could achieve a similar result but requires $\mathbf{R}$ to be built and inverted explicitly. In Section 6.1, we derive the basis for a 'regularisation' of diffusion operators aiming to bound their minimum eigenvalue without requiring access to an explicit representation of $\mathbf{R}$.

In addition to regularisation, we identified two other research areas to pursue in the future. First, the implementation of the observation error correlation operator in NEMOVAR should be extended to other observation types. In Section 6.2, we discussed how the current implementation could be adapted to 2D observations, and, in particular, how meshing the observations could be handled. Finally, in Section 6.3, we outlined a framework for evaluating the convergence of the conjugate gradient algorithm, which extends traditional approaches that focus only on the condition number. While the condition number is a first-order indicator of the convergence rate, it offers only a very incomplete description of the system. The more sophisticated approaches presented in Section 6.3, provide insight on the effect on the convergence of the distribution of the eigenvalues of the Hessian matrix between their extreme values and their interaction with the right-hand side of the system.

Bibliography

M. Asch, M. Bocquet, and M. Nodet. Data assimilation: methods, algorithms, and applications. Society for Industrial and Applied Mathematics., 2016. ISBN 9781611974546.

R. N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. Quarterly Journal of the Royal Meteorological Society, 134 (637):1971–1996, 2008. doi: https://doi.org/10.1002/qj.340.

J. Bédard and M. Buehner. A practical assimilation approach to extract smaller-scale information from observations with spatially correlated errors: An idealized study. Q. J. R. Meteorol. Soc., 146(726):468–482, 2019. doi: 10.1002/qj.3687.

M. Benkiran, M. Ruggiero, E. Greiner, P. Y. Le Traon, E. Rémy, J. M. Lellouche, R. Bourdallé-Badie, Y. Drillet, and B. Tchonang. Assessing the impact of the assimilation of SWOT observations in a global high-resolution analysis and forecasting system Part 1: Methods. Frontiers in Marine Science, 8, jul 2021. doi: 10.3389/fmars.2021.691955.

A. Bonaduce, M. Benkiran, E. Rémy, P. Y. Le Traon, and G. Garric. Contribution of future wide-swath altimetry missions to ocean analysis and forecasting. Ocean Science, 14(6):1405–1421, nov 2018. doi: 10.5194/os-14-1405-2018.

N. Bormann, M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally. Enhancing the impact of IASI observations through an updated observation-error covariance matrix. Quarterly Journal of the Royal Meteorological Society, 142(697):1767–1780, 2016. doi: https://doi.org/10.1002/qj.2774.

J.-M. Brankart, C. Ubelmann, C.-E. Testut, E. Cosme, P. Brasseur, and J. Verron. Efficient parameterization of the observation error covariance matrix for square root or ensemble Kalman filters: Application to ocean altimetry. Mon. Weather Rev., 137(6):1908–1927, 2009. doi: 10.1175/2008mwr2693.1.

W. F. Campbell, E. A. Satterfield, B. Ruston, and N. L. Baker. Accounting for correlated observation error in a dual-formulation 4D variational data assimilation system. Monthly Weather Review, 145(3):1019 – 1032, 2017. doi: 10.1175/MWR-D-16-0240.1.

M. J. Carrier, H. E. Ngodock, S. R. Smith, I. Souopgui, and B. Bartels. Examining the potential impact of SWOT observations in an ocean analysis–forecasting system. Monthly Weather Review, 144(10):3767–3782, oct 2016. doi: 10.1175/mwr-d-15-0361.1.

V. Chabot, M. Nodet, N. Papadakis, and A. Vidard. Accounting for observation errors in image data assimilation. Tellus A: Dynamic Meteorology and Oceanography, 67(1):23629, 2015. doi: 10.3402/tellusa.v67.23629.

P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. Q. J. R. Meteorol. Soc., 120(519):1367–1387, 1994. doi: 10.1002/qj.49712051912.

J. M. D'Addezio, S. Smith, G. A. Jacobs, R. W. Helber, C. Rowley, I. Souopgui, and M. J. Carrier. Quantifying wavelengths constrained by simulated SWOT observations in a submesoscale resolving ocean analysis/forecasting system. Ocean Modelling, 135:40–55, mar 2019. doi: 10.1016/j.ocemod.2019.02.001.

J. Derber and A. Rosati. A global oceanic data assimilation system. J. Phys. Oceanogr., 19(9):1333–1347, 1989. doi: 10.1175/1520-0485(1989)019⟨1333:agodas⟩2.0.co;2.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation-, background- and analysis-error statistics in observation space. Quarterly Journal of the Royal Meteorological Society, 131(613):3385–3396, 2005. doi: https://doi.org/10.1256/qj.05.108.

Y. Diouane, S. Gürol, O. Mouhtal, and D. Orban. An efficient scaled spectral preconditioner for sequences of symmetric positive definite linear systems. 2024. doi: 10.13140/RG.2.2.28678.38725.

R. C. Dorf and R. Bishop. Modern Control Systems. Pearson, 2017. ISBN 978-0-13-440762-3.

A. Ern and J.-L. Guermond. Theory and Practice of Finite Elements. Springer New York, 2004. doi: 10.1007/978-1-4757-4355-5.

D. Esteban-Fernandez and E. Peral. SWOT mission performance and error budget. In IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, jul 2018. doi: 10.1109/igarss.2018.8517385.

L. Gaultier, C. Ubelmann, and L. L. Fu. The challenge of using future SWOT data for oceanic field reconstruction. Journal of Atmospheric and Oceanic Technology, 33(1):119–126, jan 2016. doi: 10.1175/jtech-d-15-0160.1.

A. J. Geer. Correlated observation error models for assimilating all-sky infrared radiances. Atmospheric Measurement Techniques, 12(7):3629–3657, 2019. doi: 10.5194/amt-12-3629-2019.

S. Good, E. Fiedler, C. Mao, M. J. Martin, A. Maycock, R. Reid, J. Roberts-Jones, T. Searle, J. Waters, J. While, and M. Worsfold. The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. Remote Sensing, 12(4), 2020. ISSN 2072-4292. doi: 10.3390/rs12040720.

O. Goux, S. Gürol, A. T. Weaver, Y. Diouane, and O. Guillet. Impact of correlated observation errors on the conditioning of variational data assimilation problems. Numerical Linear Algebra with Applications, 31(1):e2529, 2024. doi: https://doi.org/10.1002/nla.2529.

S. Gratton, A. S. Lawless, and N. K. Nichols. Approximate Gauss–Newton methods for nonlinear least squares problems. SIAM Journal on Optimization, 18(1):106–132, 2007. doi: 10.1137/050624935.

A. Guillet, A. T. Weaver, X. Vasseur, Y. Michel, S. Gratton, and S. Gürol. Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. Q. J. R. Meteorol. Soc., 145(722):1947–1967, 2019. doi: 10.1002/qj.3537.

S. Gürol, A. T. Weaver, A. M. Moore, A. Piacentini, H. G. Arango, and S. Gratton. **B**-preconditioned minimization algorithms for variational data assimilation with the dual formulation. Q. J. R. Meteorol. Soc., 140(679):539–556, 2014. doi: 10.1002/qj.2150.

P. Guttorp and T. Gneiting. Studies in the history of probability and statistics XLIX on the Matérn correlation family. Biometrika, 93(4):989–95, 2006.

S. A. Haben, A. S. Lawless, and N. K. Nichols. Conditioning of incremental variational data assimilation, with application to the Met Office system. Tellus Ser A, 63(4):782–792, 2011. doi: 10.1111/j.1600-0870.2011.00527.x.

S. B. Healy and A. A. White. Use of discrete Fourier transforms in the 1D-Var retrieval problem. Q. J. R. Meteorol. Soc., 131(605):63–72, 2005. doi: 10.1256/qj.03.193.

M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards, 49:409–435, 1952. doi: http://dx.doi.org/10.6028/jres.049.044.

G. Hu and S. L. Dance. Efficient computation of matrix-vector products with full observation weighting matrices in data assimilation. Q. J. R. Meteorol. Soc., 147:4101–4121, 2021.

G. Hu and S. L. Dance. A novel localized fast multipole method for computations with spatially correlated observation error statistics in data assimilation. Journal of Advances in Modeling Earth Systems, 16(6): e2023MS003871, 2024. doi: https://doi.org/10.1029/2023MS003871.

R. R. King and M. J. Martin. Assimilating realistically simulated wide-swath altimeter observations in a high-resolution shelf-seas forecasting system. Ocean Science, 17(6):1791–1813, dec 2021. doi: 10.5194/os-17-1791-2021.

M. Le Hénaff, P. De Mey, B. Mourre, and P. Y. Le Traon. Contribution of a wide-swath altimeter in a shelf seas assimilation system: Impact of the satellite roll errors. Journal of Atmospheric and Oceanic Technology, 25 (11):2133 – 2144, 2008. doi: 10.1175/2008JTECHO576.1.

M. Lopez-Radcenco, A. Pascual, L. Gomez-Navarro, A. Aissa-El-Bey, B. Chapron, and R. Fablet. Analog data assimilation of along-track nadir and wide-swath SWOT altimetry observations in the western mediterranean sea. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(7):2530–2540, 2019. doi: 10.1109/JSTARS.2019.2903941.

A. C. Lorenc. Development of an operational variational assimilation scheme. J. Meteorol. Soc. Japan., 75: 339–346, 1997.

S. Metref, E. Cosme, J. Le Sommer, N. Poel, J. M. Brankart, J. Verron, and L. Gómez Navarro. Reduction of spatially structured errors in wide-swath altimetric satellite data using data assimilation. Remote Sensing, 11 (11):1336, jun 2019. doi: 10.3390/rs11111336.

S. Metref, C. Emmanuel, F. Le Guillou, J. Le Sommer, J. M. Brankart, and J. Verron. Wide-swath altimetric satellite data assimilation with correlated-error reduction. Frontiers in Marine Science, 6, jan 2020. doi: 10.3389/fmars.2019.00822.

Y. Michel. Revisiting Fisher's approach to the handling of horizontal spatial correlations of observation errors in a variational framework. Q. J. R. Meteorol. Soc., 144(716):2011–2025, 2018. doi: 10.1002/qj.3249.

K. Mogensen, M. A. Balmaseda, A. T. Weaver, M. Martin, and A. Vidard. NEMOVAR: a variational data assimilation system for the NEMO model. ECMWF Newsletter, (120):17–20, 2009.

J. L. Morales and J. Nocedal. Automatic preconditioning by limited memory quasi-newton updating. SIAM Journal on Optimization, 10(4):1079–1096, 2000. doi: 10.1137/S1052623497327854.

R. Morrow, D. Blurmstein, and G. Dibarboure. Fine-scale altimetry and the future SWOT mission. In New Frontiers in Operational Oceanography. GODAE OceanView, aug 2018. doi: 10.17125/gov2018.ch08.

E. M. Pinnington, E. Casella, S. L. Dance, A. S. Lawless, J. I. L. Morison, N. K. Nichols, M. Wilkinson, and T. L. Quaife. Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using four-dimensional variational data assimilation. Agric. For. Meteorol., 228-229: 299–314, 2016. doi: 10.1016/j.agrformet.2016.07.006.

S. Rainwater, C. H. Bishop, and W. F. Campbell. The benefits of correlated observation errors for small scales. Q. J. R. Meteorol. Soc., 141(693):3439–3445, 2015. doi: 10.1002/qj.2582.

R. Reid, S. Good, and M. J. Martin. Use of uncertainty inflation in OSTIA to account for correlated errors in satellite-retrieved sea surface temperature data. Remote Sensing, 12(7), 2020. ISSN 2072-4292. doi: 10.3390/rs12071083.

S. Ricci, A. T. Weaver, J. Vialard, and P. Rogel. Incorporating state-dependent temperature–salinity constraints in the background error covariance of variational ocean data assimilation. Monthly Weather Review, 133(1): 317 – 338, 2005. doi: 10.1175/MWR2872.1.

G. A. Ruggiero, E. Cosme, J. M. Brankart, J. Le Sommer, and C. Ubelmann. An efficient way to account for observation error correlations in the assimilation of data from the future SWOT high-resolution altimeter mission. Journal of Atmospheric and Oceanic Technology, 33(12):2755–2768, dec 2016. doi: 10.1175/jtech-d-16-0048.1.

Y. Saad. Iterative Methods for Sparse Linear Systems. Society for Industrial and Applied Mathematics, 2003. doi: 10.1137/1.9780898718003.

D. Simonin, J. A. Waller, S. P. Ballard, S. L. Dance, and N. K. Nichols. A pragmatic strategy for implementing spatially correlated observation errors in an operational system: An application to doppler radial winds. Quarterly Journal of the Royal Meteorological Society, 145(723):2772–2790, 2019. doi: https://doi.org/10.1002/qj.3592.

I. Souopgui, J. M. D'Addezio, C. D. Rowley, S. R. Smith, G. A. Jacobs, R. W. Helber, M. Yaremchuk, and J. J. Osborne. Multi-scale assimilation of simulated SWOT observations. Ocean Modelling, 154:101683, oct 2020. doi: 10.1016/j.ocemod.2020.101683.

D. Stammer and A. Cazenave. Satellite altimetry over oceans and land surfaces. CRC Press, Boca Raton, FL, 2018. ISBN 9781498743457. doi: 10.1175/JTECH-D-15-0160.1.

L. M. Stein. Interpolation of Spatial Data. Springer New York, NY, 1978.

L. M. Stewart, S. L. Dance, and N. K. Nichols. Data assimilation with correlated observation errors: experiments with a 1-D shallow water model. Tellus Ser A, 65(1):19546, 2013. doi: 10.3402/tellusa.v65i0.19546.

A. Storto, P. Oddo, E. Cozzani, and E. F. Coelho. Introducing along-track error correlations for altimetry data in a regional ocean prediction system. Journal of Atmospheric and Oceanic Technology, 36(8):1657 – 1674, 2019. doi: 10.1175/JTECH-D-18-0213.1.

Z. Strakoš and J. Liesen. Krylov subspace methods: principles and analysis. American Mathematical Society, 2013. ISBN 78-0-19-965541-0.

J. M. Tabeart, S. L. Dance, S. A. Haben, A. S. Lawless, N. K. Nichols, and J. A. Waller. The conditioning of least-squares problems in variational data assimilation. Numer. Linear Algebra Appl., 25(5):e2165, 2018. doi: 10.1002/nla.2165.

J. M. Tabeart, S. L. Dance, A. S. Lawless, S. Migliorini, N. K. Nichols, F. Smith, and J. A. Waller. The impact of using reconditioned correlated observation-error covariance matrices in the Met Office 1D-Var system. Q. J. R. Meteorol. Soc., 146(728):1372–1390, 2020. doi: 10.1002/qj.3741.

B. C. Tchonang, M. Benkiran, P. Y. Le Traon, S. J. van Gennip, J. M. Lellouche, and G. Ruggiero. Assessing the impact of the assimilation of SWOT observations in a global high-resolution analysis and forecasting system – part 2: Results. Frontiers in Marine Science, 8, 2021. ISSN 2296-7745. doi: 10.3389/fmars.2021.687414.

C. Ubelmann, L. L. Fu, S. Brown, E. Peral, and D. Esteban-Fernandez. The effect of atmospheric water vapor content on the performance of future wide-swath ocean altimetry measurement. Journal of Atmospheric and Oceanic Technology, 31(6):1446–1454, jun 2014. doi: 10.1175/jtech-d-13-00179.1.

J. A. Waller, S. Ballard, S. Dance, G. Kelly, N. K. Nichols, and D. Simonin. Diagnosing horizontal and inter-channel observation error correlations for SEVIRI observations using observation-minus-background and observation-minus-analysis statistics. Remote Sensing, 8(7):581, 2016a. doi: 10.3390/rs8070581.

J. A. Waller, D. Simonin, S. L. Dance, N. K. Nichols, and S. P. Ballard. Diagnosing observation error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. Mon. Weather Rev., 444:3534–3551, 2016b.

J. Waters, M. J. Martin, M. J. Bell, R. R. King, L. Gaultier, C. Ubelmann, C. Donlon, and S. Van Gennip. Assessing the potential impact of assimilating total surface current velocities in the Met Office's global ocean forecasting system. Frontiers in Marine Science, 11, 2024. doi: 10.3389/fmars.2024.1383522.

A. T. Weaver and P. Courtier. Correlation modelling on the sphere using a generalized diffusion equation. Q. J. R. Meteorol. Soc., 127(575):1815–1846, 2001. doi: 10.1002/qj.49712757518.

A. T. Weaver and I. Mirouze. On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. Q. J. R. Meteorol. Soc., 139(670):242–260, 2013. doi: 10.1002/qj.1955.

A. T. Weaver, C. Deltel, E. Machu, S. Ricci, and N. Daget. A multivariate balance operator for variational ocean data assimilation. Q. J. R. Meteorol. Soc., 131(613):3605–3625, 2005. doi: 10.1256/qj.05.119.

A. T. Weaver, M. Chrust, B. Ménétrier, and A. Piacentini. An evaluation of methods for normalizing diffusion-based covariance operators in variational data assimilation. Quarterly Journal of the Royal Meteorological Society, 147(734):289–320, oct 2020. doi: 10.1002/qj.3918.

P. P. Weston, W. Bell, and J. R. Eyre. Accounting for correlated error in the assimilation of high-resolution sounder data. Q. J. R. Meteorol. Soc., 140:2420–2429, 2014.

M. Yaremchuk, J. M. D'Addezio, G. Panteleev, and G. Jacobs. On the approximation of the inverse error covariances of high-resolution satellite altimetry data. Quarterly Journal of the Royal Meteorological Society, 144(715):1995–2000, 2018. doi: https://doi.org/10.1002/qj.3336.

O. C. Zienkiewicz and R. L. Taylor. The Finite Element method. Butterworth-Heinemann, 2000.