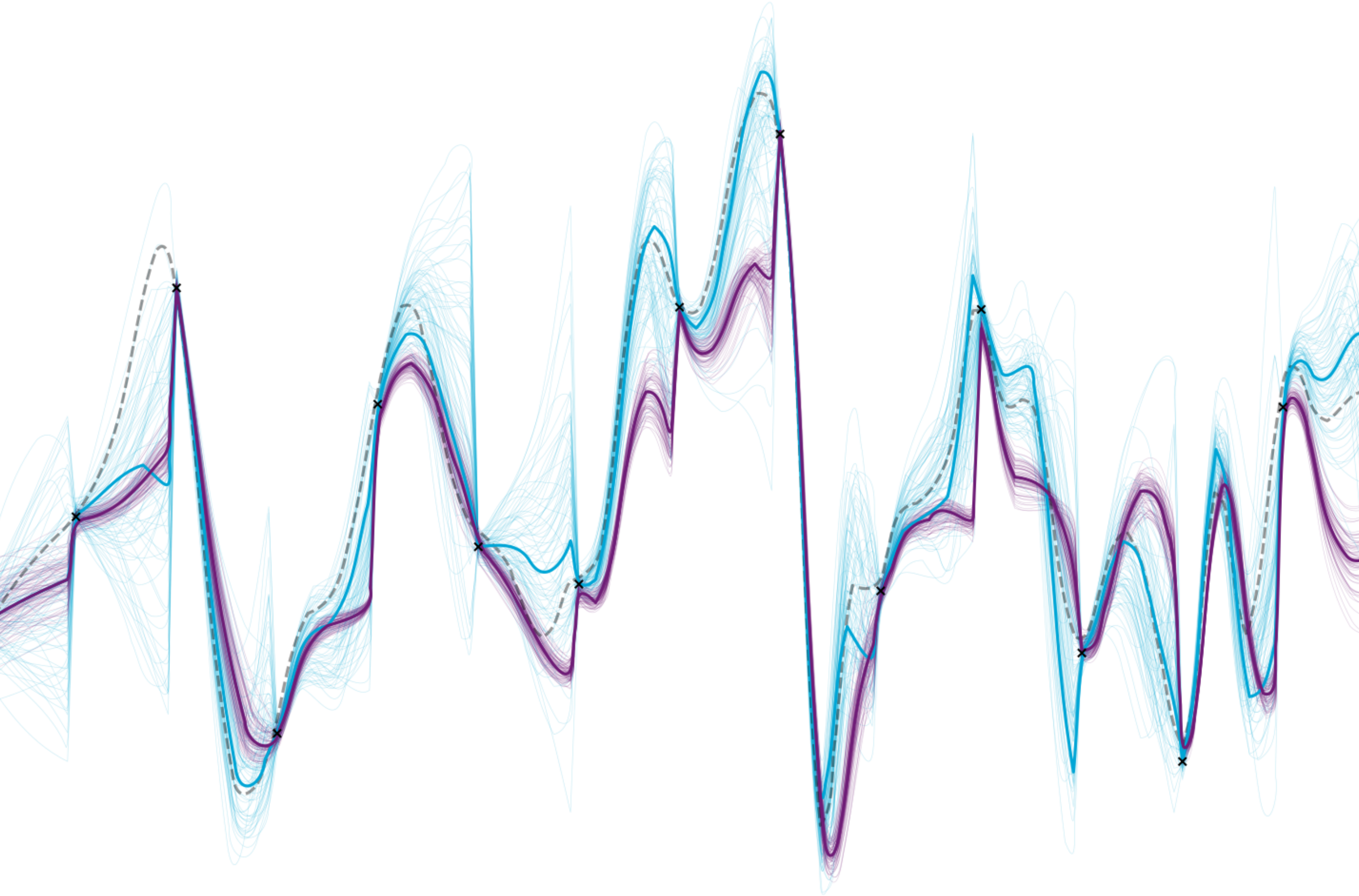


Sequential Bias-Aware Data Assimilation Methods with Parameter Uncertainty

Master's thesis in Applied Mathematics

DAVID PLAZAS

FACULTY OF ELECTRICAL ENGINEERING, MATHEMATICS AND COMPUTER SCIENCE

DELFT UNIVERSITY OF TECHNOLOGY
Delft, The Netherlands 2024

Master's thesis

A Study of Sequential Bias-Aware Data Assimilation Methods with Parameter Uncertainty

David Plazas

in partial fulfillment of the requirements for the degree of

Master of Science in Applied Mathematics
Specialization in Computational Science and Engineering

at the Delft University of Technology,
to be defended publicly on Monday, August 12th, 2024 at 10:30 AM.

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics
Mathematical Physics
Delft University of Technology
Delft, The Netherlands
2024

A Study of Sequential Bias-Aware Data Assimilation Methods with Parameter Uncertainty

David Plazas

Thesis committee

Chair:	Martin Verlaan	EEMCS-DIAM-MP	Full Professor, TU Delft
Core Member 2:	Femke Vossepoel	CITG-GSE-RSV	Full Professor, TU Delft
Core Member 3:	Max Ramgraber	CITG-GSE-RSV	Assistant Professor, TU Delft

Master's Thesis 2024

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics
Mathematical Physics

In collaboration with

Faculty of Civil Engineering and Geosciences
Department of Geoscience & Engineering
Reservoir Engineering

Delft University of Technology
2628CD Delft, The Netherlands

Cover: Author's creation

Typeset in LaTeX
Template adapted from `https://bit.ly/3ULbBV6`
Delft, The Netherlands

A Study of Sequential Bias-Aware Data Assimilation Methods with Parameter Uncertainty

David Plazas

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Summary

This thesis analyzes the effectiveness of bias-aware filtering techniques, particularly the Colored-noise Kalman Filter, in addressing parameter and bias estimation in data assimilation problems. The research explores the ability of this method to differentiate between the impacts of bias and parameter uncertainty, focusing on how the concept of feedback within the filtering process influences the estimation of both bias and parameters.

The study uses the Lorenz-96 model to conduct twin experiments, investigating various scenarios involving parameter estimation, bias estimation, and combined parameter and bias estimation. The experiments reveal that in a feedback filter configuration, where the bias directly influences the Ordinary Differential Equation system, the forcing parameter F of the Lorenz-96 model becomes indistinguishable from the bias. Conversely, a non-feedback filter configuration allows for the independent estimation of both the parameter and the bias.

In addition, the research highlights the challenges and considerations in implementing a flexible data assimilation framework, particularly in managing state augmentation, stochastic updates, and bias representation. It emphasizes the importance of carefully considering the feedback mechanism in bias-aware filtering, as it significantly impacts the estimation of parameters and bias.

The findings of this thesis offer valuable insights into the application of bias-aware filtering techniques in the presence of parameter uncertainty and provide a foundation for future research in developing robust and versatile data assimilation frameworks. The study encourages further exploration of these methods in real-world applications and with more complex bias structures to advance our understanding and ability to address uncertainties in dynamic systems effectively.

Keywords: Data Assimilation, Model Bias, Parameter Estimation, Lorenz-96, Bias-aware Filtering, Ensemble Kalman Filter, Colored-noise Kalman Filter

Acknowledgments

To my loving parents for their endless and tireless support throughout all these years of studies.

To TU Delft and the Justus & Louise van Effen Foundation for sponsoring my studies and time in the Netherlands.

To my three supervisors, Femke, Martin, and especially Max, for being so understanding and patient (especially during this last stretch of the thesis). Thank you for pushing me in the right way and dedicating time to this project.

To all my friends back home, you were my way out when things weren't working here. You kept memories alive for me. I wish I could have been there for all your accomplishments, but you'll always hold a spot in my heart.

To the IDE girlies, for the random plans, for believing in me, and always bringing a smile to my face :)

To Estepán, for being such a loyal friend. Always willing to chat, play games, or just hear me complain. Now I wish we could have spent more time together back home.

To Juanse, for all these years of friendship, since high school and through the journey of Engineering Mathematics. Thank you for teaching me how to live. I owe you my kindness and empathy.

To the IP Gals, I do not have words for you all... The sporadic plans, all the coffees, dinners, and board games nights. You truly made my time in Delft the best of my life - so far ;)

Lastly, and foremost, to my dear brother, Samuel, who inspired me to follow him on this adventure to Europe. You started this! Teaching me random math facts since we were just teenagers. You opened doors for me (and many others from Colombia) in Belgium, which became my passage to the Netherlands. I don't think I would have made it this far without someone to show me the way.

David Plazas
Delft, August 2024

Contents

List of Symbols

Below is the notation of indices, sets, parameters, and variables used throughout this thesis.

Indices

t	Index for a continuous-time instant
t_k	A discretized continuous-time instant
k	Index for a discrete-time t_k
$\mathbf{x}(t)$	Notation for a continuous-time quantity at time t
$\mathbf{x}_k$	Notation for a discrete-time quantity at time t_k

Sets

$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of non-negative real numbers
$\mathbb{R}^n$	Set of real-valued (column) vectors of size n
$\mathbb{R}^{n\times n}_{\text{spsd}}$	Set of symmetric positive semi-definite matrices of size $n \times n$
$\mathbb{Z}_+$	Set of non-negative integer numbers
Y_k	Set of observations up to time t_k, i.e, $Y_{k-1} := \{\mathbf{y}_s : 1 \leq s \leq k-1\}$

Variables

$\mathbf{x}_k$	System state vector at time t_k
$\mathbf{y}_k$	Observation at time t_k
$\mathbf{w}_k$	State error process at time t_k
$\mathbf{Q}_k$	State error covariance at time t_k

$\mathbf{v}_k$	Observation error process at time t_k
$\mathbf{R}_k$	Observation error covariance at time t_k
$\mathbf{u}_k$	Input or forcing of the system at time t_k
$\mathbf{b}_k$	Bias state at time t_k
$\tilde{\mathbf{x}}_k$	Unbiased state vector at time t_k
$\mathbf{z}_k$	Augmented state vector at time t_k
$\mathbf{q}_k$	Augmented noise process at time t_k
$\boldsymbol{\theta}_k^f$	Parameters forecast state at time t_k
$\hat{\boldsymbol{\theta}}_k$	Parameters analysis state t_k
$\mathbf{w}_k^{\boldsymbol{\theta}}$	Parameter error process at time t_k
$\mathbf{w}_k^{\mathbf{b}}$	Bias state error process at time t_k
$\mathbf{M}_k$	Linear state-transition matrix at time t_k
$\mathbf{B}_k$	Input matrix at time t_k
$\mathbf{H}_k$	Linear observation model at time t_k
$\mathbf{P}_k$	True covariance matrix at time t_k
$\mathbf{K}_k$	Kalman gain at time t_k
$\boldsymbol{\xi}_{k,i}^f$	i-th ensemble member forecast state at time t_k
$\hat{\boldsymbol{\xi}}_{k,i}$	i-th ensemble member analysis state at tie t_k
$\mathbf{w}_{k,i}$	i-th realization of the state error process $\mathbf{w}_k$ at time t_k
$\mathbf{v}_{k,i}$	i-th realization of the observation error process $\mathbf{v}_k$ at time t_k
$\mathbf{A}_k$	Coefficient matrix of the AR(1) process at time t_k

Parameters

T	The final simulation time, i.e., the system runs on $[0,\ T]$
K	The final assimilation time, i.e., predict $(K,\ T]$ without filtering
Δt	The integration time step
Δk	The assimilation time step, i.e., the time between observations
$n_{\mathbf{x}}$	Length of state vector $\mathbf{x}$
$n_{\mathbf{y}}$	Length of observation vector $\mathbf{y}$
$n_{\mathbf{u}}$	Length of the input vector $\mathbf{u}$
$n_{\boldsymbol{\theta}}$	Number of uncertain parameters to estimate
$n_{\mathbf{b}}$	Length of the bias vector $\mathbf{b}$
$\mathbf{I}_n$	Identity matrix of size $n \times n$
$\mathbf{0}$	Matrix of zeros

$\mathbf{1}_n$	Column vector of ones of size n
$G(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$
N	Number of ensemble members
$\boldsymbol{\theta}$	System parameters
$\mathbf{H}^{\mathbf{b}}$	The map from the bias state to the system state
$\mathbf{F}$	Feedback matrix for bias-aware filters

Other

$\mathbb{E}\left[\mathbf{x}\right]$	Expectation of a random variable $\mathbf{x}$
$\mathbb{V}\left[\mathbf{x}\right]$	Variance of $\mathbf{x}$, i.e. $\mathbb{V}\left[\mathbf{x}\right] \equiv \mathbb{E}\left[(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right])(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right])^T\right]$
$\mathrm{diag}(\ldots)$	Block-diagonal matrix based on the inputs
$\mathcal{M}_k(\ldots)$	Nonlinear forward model at time t_k
$\mathcal{G}_k(\ldots)$	Forward model for the bias at time t_k
$\mathcal{H}_k(\ldots)$	Nonlinear observation model at time t_k
$\mathbf{X}[\ldots]$	`NumPy`'s indexing of a tensor $\mathbf{X}$, e.g. $\mathbf{X}[\,:\,, 0\,]$
$\lfloor a \rfloor$	Floor function of a number $a \in \mathbb{R}$
$\mathbf{x}^T$, $\mathbf{A}^T$	Notation for vector $\mathbf{x}$ or matrix $\mathbf{A}$ transpose
$\mathbf{x}^f$, $\mathbf{P}^f$	Notation for a quantity forecast (*prior*)
$\hat{\mathbf{x}}$, $\hat{\mathbf{P}}$	Notation for a quantity analysis (*posterior*)
$\dot{x}(t)$	Notation for the first time derivative of a variable $x(t)$
$\ddot{x}(t)$	Notation for the second time derivative of a variable $x(t)$

List of Acronyms

4D-Var 4-Dimensional Variational data assimilation. 1, 6

AR Auto-Regressive. xii, 14, 29, 30, 37, 52–54, 56–58, II, III

BLUE Best Linear Unbiased Estimate. 20

CDA Coupled Data Assimilation. 9

ColKF Colored-noise Kalman Filter. v, ix, xx, xxi, 2, 3, 7, 10, 11, 14, 15, 17, 29, 37, 40, 41, 51–61, 63, 64, II, III

DA Data Assimilation. 1–3, 5–7, 9, 11–15, 17–19, 26, 27, 33, 35–40, 44–46, 51, 63–65

DEnKF Deterministic Ensemble Kalman Filter. 22

EAKF Ensemble Adjustment Kalman Filter. 12, 22

ECMWF European Centre for Medium-Range Weather Forecasts. 32

EKF Extended Kalman Filter. 7, 8, 22

EnKF Ensemble Kalman Filter. v, ix, xix–xxi, 1, 3, 6–9, 12, 14–17, 20–23, 26, 31, 38, 40, 43–45, 47–57, 63, 64, II, III

EnKS Ensemble Kalman Smoother. 6

EnRDA Ensemble Riemannian Data Assimilation. 13

ESN Echo State Network. 12

ETKF Ensemble Transform Kalman Filter. 13, 22

FDS Forced Dissipative Systems. 32, 33

KF Kalman Filter. ix, 1, 3, 5–8, 10, 14–17, 19–23, 27, 29–31, 40

KS Kalman Smoother. 6

LEKF Local Ensemble Kalman Filter. 11

LETKF Local Ensemble Transform Kalman Filter. 13

LHO Linear Harmonic Oscillator. ix, xix–xxi, 32, 43–48, 52–55, 63, 64

MLEKF Maximum Likelihood Ensemble Kalman Filter. 22

NWP Numerical Weather Prediction. 1, 6, 13

ODE Ordinary Differential Equation. v, 11, 18, 32, 37, 40, 41, 58, 63, 64

PDE Partial Differential Equation. 18, 35

r-EnKF regularized bias-aware Ensemble Kalman Filter. 12
RK4 4th order Runge Kutta. 36, 43
RNG Random Number Generator. 39, 43
RRSQRT Reduced Rank Square Root Kalman Filter. 8, 14

SDE Stochastic Differential Equation. 18
SEIK Singular Evolutive Interpolated Kalman. 15
SepKF Separate-bias Kalman Filter. ix, 10, 11, 15, 16, 65
SSKF Steady-State Kalman Filter. 14

UKF Unscented Kalman Filter. 8

List of Figures

List of Tables

1

Introduction

Data Assimilation (DA) plays an essential role in numerous fields where it is necessary to combine model predictions with observational data to improve forecasting accuracy. Applications of this discipline are, for instance, found in Numerical Weather Prediction (NWP), chemical transport models, ocean circulation, earthquake modeling, and more [1, 2, 3].

In general, DA methods can be classified between variational and sequential approaches. The main difference is how the observations are incorporated to correct the model predictions, even though they can be shown to be equivalent in particular (generally linear) scenarios. Perhaps the most known example of a variational approach is the 4-Dimensional Variational data assimilation (4D-Var), and of a sequential approach is the Ensemble Kalman Filter (EnKF), a popular variation of the Kalman Filter (KF).

Most of these DA techniques rely on strong assumptions about the statistical properties of errors and uncertainties. In addition, real-world applications often face challenges such as model biases, uncertain parameters, discretization errors, missing physics, or wrong assumptions when developing the model [4, 5].

Typically, the forward model incorporates an additional Gaussian white noise process to account for these sources of error in the numerical model. However, the actual model error rarely satisfies this assumption, and, in this case, the prediction model would be biased. In addition, it is well known that when the error processes in the model are not purely random, the analysis estimates of traditional DA techniques will be sub-optimal [6, 4, 7]. Yet, providing a more accurate description of the error processes is often challenging. An alternative is to estimate the biases directly in the DA process, also known as bias-aware DA.

In the context of bias-aware DA, the concept of bias can (generally) refer to either model or observation bias, the former being the main focus of attention in this work. Moreover, the effects of uncertain parameters in the model can be easily confused with model or observation bias, as their impacts on the observed data are hard to distinguish [8]. The main objective of this thesis is to analyze how different bias-aware filtering techniques perform under various forms of uncertainty.

In the literature, the biases are usually treated as extra parameters to estimate [6, 7, 9], but are sometimes assumed to follow a more complex forward model than the simple random walk model frequently used for parameter estimation [10, 8]. Other approaches also include modeling more complex noise processes in the model, e.g., colored noise, and use this approach to account for the "missing" parts of the model [7, 11, 12]. This is known as the Colored-noise Kalman Filter (ColKF), and it is the main bias-aware method used throughout this work.

This project originated from the idea of exploring how parameter uncertainty influences these bias estimation methods. Specifically, we aim to investigate whether bias-aware filters can differentiate between the effects of bias and parameter uncertainty depending on how the bias is fed back into the system. This distinction is explored through two approaches: feedback and no-feedback bias formulations (see Section 3.3.1).

Thus, this thesis aims to answer the following questions:

1. How can bias-aware filtering techniques be employed to distinguish between the effect of parameter uncertainties and systematic biases?
2. How does the concept of feedback in bias-aware filtering affect the estimation of both bias and parameters?
3. Which practical considerations must be addressed to implement a flexible (bias-aware) data assimilation framework?

To answer these questions, we created twin experiments for parameter estimation, bias estimation, and combined scenarios involving feedback and no feedback with a ColKF, using the well-known Lorenz-96 model. This model was selected for two main reasons: first, this model has been widely used by the DA community as a test case [13], and second, because its unique forcing parameter acts similarly to a bias estimation with feedback, which makes for an intriguing setting about comparing these different forms of bias, and when they can be distinguished. Nonetheless, it should be possible to identify these quantities separately in a no-feedback approach.

In summary, by the end of this thesis, we aim to examine how bias-aware filters can be effectively employed in the presence of parameter uncertainty. Moreover, a side product of this work is to implement a flexible library in Python to perform DA experiments. No Python libraries to conveniently perform bias-aware DA were found, as the augmentation of the model and system state is often done by the end user and not by the framework itself. This library is intended to support future research, particularly in the context of state, parameter, and bias estimation.

This thesis is structured as follows:

- *Chapter 2: Literature Review* reviews some of the existing research on DA, parameter estimation, bias-aware DA methods, and some of their applications.
- *Chapter 3: Theory* explains the theoretical background of the methods and models used throughout this thesis, including the KF, the EnKF, and the ColKF. Additional concepts, such as covariance localization, state augmentation for parameter estimation, and testing models, are also presented.
- *Chapter 4: Methods* dives into more details about some of the considerations needed when implementing a flexible DA library, what are the advantages of the current implementation, how to set up twin experiments, and which experimental design was followed for the results.
- *Chapter 5: Results* presents all the results obtained throughout the project, including the validation twin experiments for each model. These include state estimation, state and parameter estimation, state and bias estimation, and all combined.
- *Chapter 6: Discussion & Conclusion* discusses the results and concludes the project, including some considerations for future research.

2

Literature Review

2.1 Data Assimilation

The significance of DA methods cannot be overlooked. Since the end of the last century, they have played a central role in improving forecasts of numerical models based on real observation data. These methods are often classified as variational and sequential [14, 15].

Variational DA mainly focuses on estimating system states, parameters, or initial conditions by solving a global optimization problem that minimizes the distance between all observed data and the numerical model predictions over a time window in one go [16]. This approach is usually expensive in terms of computational costs since the global optimization solution is not trivial and often requires the calculation of an adjoint of the numerical model, which is not always available or feasible. An overview of variational methods in recent years is presented by Bannister [16], where the author summarizes and clarifies various methods, focusing on combining variational and ensemble techniques and hybrid approaches. In addition, he provides derivations of popular schemes, details common localization representations, and discusses potential future developments. Variational approaches have been used in numerous application fields, such as hydrology [17], hemodynamics [18], oceanography [19, 20], atmospheric transport models [21, 22, 23, 24], seismology [25], ocean circulation models [26, 27, 28], and more.

Sequential DA, on the other hand, corrects the state based on each new observation in chronological order [29]. In this manner, the information is propagated forward in time, which is an advantage when compared with variational methods since no calculation of the adjoint model is required, making this approach more versatile for any model [30]. Sequential DA methods have found application in many fields, for instance, oceanography [30, 20], reservoir engineering [31] and atmospheric transport [32, 33, 34]. In the case of linear dynamics (and Gaussian errors), the optimal sequential DA filtering estimator is the KF [35, 29]. The importance of this filter cannot be overlooked, especially not only in the context of DA. The next section will present a more detailed collection of references about this topic.

Furthermore, multiple reviews of the theory and applications of DA techniques are already available in the literature. Navon [36] provided a review of DA applications in the field of NWP, focusing mainly on the history and early development of the 4D-Var and how these approaches gained importance and were implemented in several NWP centers worldwide. Additionally, Sandu and Chai [37] presented an overview of both variational and sequential DA methods applied to the optimal representation of the atmosphere's chemical composition and air quality modeling. They highlight the algorithms used in operational systems and the challenges associated with DA in this field.

In the field of geosciences, Carrassi et al. [29] provided a comprehensive overview of DA techniques and their theoretical foundations, highlighting their interdisciplinary nature across statistics, dynamical systems, and numerical optimization. The authors first established the framework of DA, using probability density functions to describe model and observational errors. They then introduced three key problems in DA: prediction, filtering, and smoothing. The KF and Kalman Smoother (KS) are introduced as exact solutions for linear systems with Gaussian errors and variational methods for nonlinear cases. Ensemble methods are presented as practical alternatives for high-dimensional systems, including the EnKF and Ensemble Kalman Smoother (EnKS). In addition, the authors presented a few selected topics to showcase the challenges associated with applying DA in geosciences. These include DA for chaotic dynamics such as the atmosphere or the ocean, DA for non-Gaussian cases, DA for chemical constituents of the atmosphere, and an example of operational DA system for an ice-ocean model.

Another review on DA techniques was presented by Montzka et al. [38]. It distinguishes four types of DA approaches based on the scales used on the assimilation framework and data, namely, acknowledging that different scales and types of data can be assimilated concurrently. The authors highlight the advantages and disadvantages of assimilating various data types and scales, concluding that it is beneficial in the general case. However, a better understanding of multiscale DA methods is still needed.

Several books and compilations are available in the literature. They usually provide a more detailed and theoretical description of the general DA methods. For instance, Lahoz, Khattatov, and Menard [39] present a comprehensive overview of current practices and future prospects in DA. The book is structured into six parts: theory and methods; sources and types of observational data; applications in meteorology and atmospheric dynamics; chemical data assimilation; broader applications beyond meteorology, including oceans, land surfaces, and ionospheric models; and future directions.

Additionally, Park and Xu [40] gathered a collection of papers that mix theoretical and practical research. The methodological aspects include variational methods, ensemble methods, particle filtering, genetic algorithms, and more. Some applications include parameter estimation, radar/satellite assimilation, DA for land surface and water balance modeling, oceanic and meteorological DA, and radar rainfall estimates.

Evensen [41] introduces the formulation and solution of the data assimilation problem, focusing on methods that allow models to contain errors with evolving error statistics over time. There is a particular emphasis on the EnKF and similar techniques, highlighting their popularity due to their simple implementation, ease of interpretation, and effectiveness with nonlinear models. The book also includes a detailed list of publications that use the EnKF up to the date of publication.

The book by Chui and Chen [12] shows a detailed and theoretical formulation of the KF and some of its extensions. Among these are the colored noise filter (similar to the ColKF), square-root algorithms, Extended Kalman Filter (EKF), and more. The book also presents several examples in real-time applications, such as tracking systems, satellite navigation, trajectory estimation, and beyond.

Recently, Evensen, Vossepoel, and Leeuwen [1] provided a unified approach and formulation to various modern assimilation techniques, using the authors' combined experience in research and applications. In addition, the book is also suitable for advanced courses, providing detailed explanations and maintaining a modest mathematical level. Finally, the books by Asch, Bocquet, and Nodet [2] and Fletcher [3] are worth highlighting, as they provide a very detailed explanation of the concepts leading to modern and more advanced DA techniques.

2.2 Kalman Filtering

The KF is an algorithm by Kalman [42] and Kalman and Bucy [43] that provides estimates of the state of a discrete stochastic linear system from a series of incomplete and noisy measurements. It operates in a two-step process: forecast and analysis. In the forecast step, the filter uses a numerical model to predict the state at the next time step, along with the associated uncertainty. The analysis step incorporates new observations to correct the predicted state, reducing uncertainty.

It is widely used in various research and applied fields, such as navigation, control and tracking of vehicles [44, 45, 46], sensor fusion [47, 48], signal processing [49, 50], robotics [51], and econometrics. The review presented by [52] provided an overview of applications of the KF in five different industrial fields.

Furthermore, the KF has been extended to handle nonlinear systems through variants like the EKF [53, 54] and the Unscented Kalman Filter (UKF) [55, 56]. In addition, Verlaan and Heemink [57] proposed the Reduced Rank Square Root Kalman Filter (RRSQRT) to address some of the computational issues associated with the exact formulation by using a reduced rank approximation of the error covariances based on square root factorization and applied to a tidal flow forecasting.

Numerous books in the literature address the KF. Some of these references include the works by Jazwinski [53], Cohn [58], Bain and Crisan [59], Catlin [60], Grewal and Andrews [61], Chui and Chen [12], Govaers [62], and Schuppen [63].

However, in the context of geophysics and large-scale models, the exact formulation of the KF is usually not applied [64]. The reason is twofold. First, the size of the state vector is significantly larger (usually from 10^4 to 10^6) compared to other application areas. This makes computing and storing some of the mathematical objects in the KF recursions computationally unfeasible. Second, these models are typically nonlinear, adding an extra complication by linearizing or using more sophisticated filtering approaches, as is the case for the EKF or the UKF.

Hence, an approximation was made: the EnKF. Proposed originally by Evensen [65], this approach uses an ensemble of perturbed individual realizations of the model (often called ensemble members) to represent the uncertainty and uses a Monte Carlo approach to approximate the state mean and error covariance [66, 64]. As it will be presented later (see Section 3.2.2), the standard formulation of the EnKF assumes a linear observation model. Although extensions for nonlinear observation models have been proposed, e.g., in [67, 68, 69], for the objectives of this work, a linear observation model is assumed.

Applications of the EnKF exist in several research fields, such as chemical transport models [70, 34, 71], soil moisture [72], ocean models [73, 74, 75] and reservoir engineering [76, 77, 31, 78]. Finally, the EnKF has been applied in several operational systems agencies. Examples of these have been documented in the works by Houtekamer et al. [79], Wei et al. [80], Bonavita, Torrisi, and Marcucci [81], El Serafy and Mynett [82], Houtekamer, Mitchell, and Deng [83], McMillan et al. [84], Houtekamer et al. [85], and Rafieeinasab et al. [86].

2.3 Parameter Estimation

Sequential DA methods have been applied to estimate uncertain parameters of forward models. Evensen, Dee, and Schröter [10] presented a general introduction to the estimation of parameters for dynamical models from a theoretical standpoint. In addition, Evensen [64] discusses the EnKF and how it can also be used for parameter estimation.

Ruiz, Pulido, and Miyoshi [8] presented a thorough review of the use of ensemble-based methods in parameter estimation. The authors also highlight how DA can overcome common problems with traditional parameter estimation techniques, such as multiple local minima and sophisticated optimization methods to avoid them, and how increasingly expensive computationally this becomes. Moreover, The authors show how most DA approaches use some form of state augmentation approach for online parameter estimation, usually assuming that parameters are constant during the model forecast step. They are only updated during the assimilation of new observations. More in the particular context of this work, the authors argue that bias estimation can be regarded as a specific case of parameter estimation and that further research is still needed to evaluate how parameter estimation techniques can be mixed methods that correct other sources of model error, such as model bias, in a DA framework.

There have been multiple applications of parameter estimation using DA in the literature. Some examples are presented. Akter et al. [78] showcased a modified tunning version for an EnKF to estimate both state and parameters in the presence of model uncertainty in a reservoir context. Moreover, Sueki et al. [87] evaluated the precision and convergence speed of the EnKF for parameter estimation in an atmospheric model. Canuto et al. [88] applied an EnKF as well for estimation of patient-specific parameters in a three-dimensional cardiovascular flow simulation. Furthermore, a climate application of parameter estimation using the EnKF was shown by Annan et al. [89]. In the context of hydrological models, Moradkhani et al. [90] showed an EnKF to perform joint state and parameter estimation. Lastly, Zhang et al. [91] reviewed the applications of Coupled Data Assimilation (CDA) and parameter estimation to coupled ocean-atmosphere models.

2.4 Bias-Aware Data Assimilation

In the context of DA, “bias” can refer to different types, but they can be (generally) classified between model and observation bias. Model bias, sometimes called forecast bias, refers to errors with nonzero mean in the forecast model; they are often caused by systematic errors in the model, such as discretization errors, wrong assumptions, outdated parameters, or faulty boundary conditions [6]. Moreover, observation bias refers to nonzero mean errors in observed data due to, for instance, defective measurement equipment or differences in frames of reference when acquiring data [4]. In the early work by Dee and Da Silva [6], they showed that if either (or both) the model and observations are biased, the estimation scheme produced by the KF will result in a biased analysis state.

Early works addressing sequential data assimilation while explicitly acknowledging bias date back to Jazwinski [53] and Friedland [92]. The former is one of the early books on stochastic processes and filtering theory, and the author already suggests a state-augmentation approach to estimate system errors simultaneously with the state in the filtering procedure. Moreover, in the latter, Friedland derived the first version of what is known nowadays as the Separate-bias Kalman Filter (SepKF), corresponding to the version without feedback. The derivation assumes that the bias is constant, and it was performed for both continuous- and discrete-time filtering. This approach has been extended in subsequent publications to include nonlinear and stochastic bias [93, 94, 95, 96, 97]. However, this approach assumes that the forecast is independent of the estimated bias or *bias-blind*, as later appointed by Dee and Da Silva [6]. In that same reference, the authors take Friedland's (and the subsequent papers) theoretical framework and make the forecast no longer bias-blind by introducing a feedback loop from the biased estimator to the original analysis system.

The paper by Drécourt, Madsen, and Rosbjerg [7] applied two variations of bias-aware KFs to a groundwater model. In this work, only the concept of model bias is addressed, and the authors proposed the idea of feedback in the context of bias-aware filtering. Then, a summary of the ColKF and the SepKF, both with unified notation, is presented. These formulations also include an additional "switch" matrix **F** (see Section 3.3.1), which determines explicitly in the equations if the bias-aware filter uses feedback or not. As part of the conclusion of this work, the authors claim that the feedback versions of bias-aware filters (following [6]) are more versatile for more complicated forms of bias, as they address the drift (derivative) of the bias, rather than bias value itself (the case of a no feedback formulation). This paper provided vital concepts that were used throughout this thesis.

Similarly, Rasmussen et al. [98] evaluate the performance of both the ColKF and SepKF for correcting observation bias in integrated hydrological modeling. The experiments were conducted using synthetic and real observations, and significant improvements were found compared to a bias-blind experiment. In general, ColKF showed faster convergence to the bias values, albeit needing a larger ensemble size than the SepKF. The authors also found that both filters usually underestimated the bias, and parameters could not be estimated optimally in the real setting.

Baek et al. [99] discuss modifications to the Local Ensemble Kalman Filter (LEKF) to address forecast model bias in atmospheric state estimation. The authors propose augmenting the atmospheric state with model bias estimates, and they explore three different ways to parameterize the model bias: accounting for feedback, no feedback, or both (albeit the authors do not explicitly refer to it as feedback). Through numerical experiments using the Lorenz-96 model, the study demonstrates that effectively parameterizing the model bias can significantly improve forecast accuracy. It presents a method that allows for correlations between forecast state uncertainties and biases, enhancing the computational efficiency of the LEKF approach. In addition, the feedback version proposed in this study works directly on the Ordinary Differential Equation (ODE), in contrast to the typical bias formulation that affects the state only after discretization. This will motivate part of the aspects discussed in Section 4.1.2.

Another state augmentation approach was recently showcased by Diab Montero [9]. The author added a forcing vector directly in the ODE formulation to account for unresolved model errors and used DA to estimate the augmented state. This approach showed successful results for the Lorenz-96 model and an earthquake model.

Furthermore, Ménard [100] presents the derivation of multiple DA approaches to estimate both observation and model bias based on theoretical developments made around the time of that manuscript. The derivations are presented in variational and sequential form, as well as static and dynamic cases (offline and online). The author also discusses two different forms of detecting if bias is present in the model: through innovations or analysis increments. The paper concludes that distinguishing between model and observation biases using innovation statistics remains a significant challenge. Currently, additional information is needed to address this.

This approach was later used by Glegola, Hanea, and Kaleta [101] in application in reservoir characterization. In addition to the application-specific results, the study also concludes that the presence of bias can be determined by examining the innovations in a bias-blind experiment. However, additional information is needed to determine the source of bias (model, observation, or both). Finally, they also concluded that if bias is estimated in a parameter estimation setup, the estimated parameters have a more significant uncertainty (variance) when compared to a bias-blind experiment.

Lorente-Plazas and Hacker [102] explore the effects of simultaneous observation or model biases on a synthetic experimental setup. Their approach uses a state augmentation that includes observation and model bias vectors and uses an Ensemble Adjustment Kalman Filter (EAKF) to perform the assimilation. A critical aspect of their approach is that they feed back the estimated bias directly to the continuous model, which is an essential part of the discussion that will be carried out in this thesis, especially when considering the specific form that the forcing takes in conjunction with the model bias.

More recently, Nóvoa, Racca, and Magri [5] identified two main limitations with most sequential bias-aware techniques when the nature of the bias is unknown: the uniqueness of bias and the parameterization of the bias model. The former refers to how a bias estimation procedure is not always constrained or unique, yielding estimations that do not necessarily improve the analysis. Moreover, the latter refers to how most bias-aware methods require a parametric form of the bias to be assumed *a priori* so that the model parameters can be estimated in the assimilation process. However, to overcome these limitations, the authors proposed using a regularized bias-aware Ensemble Kalman Filter (r-EnKF) in combination with an Echo State Network (ESN) as the bias model. The former addresses the issue of uniqueness, as it introduces an extra L_2-like regularization (similar to a Tikhonov regularization) of the bias vector to the cost function of the EnKF. The latter removes the need for a specific functional form of the bias: since ESNs are generalized nonlinear autocorrelation functions [103], which are universal approximators [104], any particular form of model error (bias) can be, theoretically, represented, given that the network is trained correctly. The methodology proposed was successfully tested using two models of nonlinearly coupled oscillators with and without time delay and considering different types of bias perturbations: linear, nonlinear, and time/state-dependent.

Moreover, this paper is a continuation of two papers by the same research group, namely [105, 106], which uses a similar approach specifically for the case of thermoacoustic DA. These works showed the ESN-based method to forecast bias for the first time, but no regularized version of the EnKF was developed yet. They performed experiments to show that the proposed approach can correctly estimate state, parameters, and bias.

The paper by Nóvoa, Racca, and Magri [5] is state-of-art, so there are still some limiting factors before a method like this can be applied in another setting. The novel method opens an exciting intersection between recent machine learning methodologies and DA. However, the method still has too many “hyperparameters” that would need tuning before an approach like this can be extended into an operational setting.

More research is still missing in this aspect.

A different procedure was recently proposed by Tamang et al. [107]. They present an advanced DA technique designed for high-dimensional nonlinear dynamical systems. This technique, known as Ensemble Riemannian Data Assimilation (EnRDA), leverages the optimal mass transport theory and Wasserstein distances. Unlike traditional methods, EnRDA does not assume Gaussian distributions, allowing it to handle systematic biases more effectively. The proposed approach was tested using a chaotic Lorenz-96 model and a two-layer quasi-geostrophic atmospheric circulation model. The paper also claims that despite its computational complexity, this method is promising to enhance the predictability of large-scale dynamical systems.

Another approach worth mentioning is the low-frequency filtering approach initially discussed by Asher et al. [108]. The primary purpose of this bias filter is to correct gradually long frequency trends on the model, the so-called "unresolved drivers", that induce an error between the simulated water level from a hydrodynamics model and the actual observations. The basic idea is that the filtering is performed on a moving average space, i.e., both state forecast and real observations are averaged over time windows, and the state is corrected using this averaged analysis, accounting for a low-frequency offset of the model. However, these corrections are made to the pressure field above the water level to ensure longer-term corrections on the state. The methodology was successfully tested using a large-scale hydrodynamic model of the Caribbean Sea and assimilating storm-surge observations during hurricane Matthew.

Several sources can also be found in the field of bias-aware DA, particularly in the case of observation bias. Although this topic is beyond the scope of this thesis, some sources are worth mentioning. Derber and Wu [109] present an observation bias correction scheme based on linear models for satellite radiance observations aiding a NWP model. The coefficients of said linear models are included in the analysis scheme, similar to a parameter estimation approach. Moreover, Fertig et al. [110] present a Local Ensemble Transform Kalman Filter (LETKF) to correct the same type of observations, with a state augmentation approach and assuming a persistence model for the bias evolution. In addition, Eyre [111] explores how biases in the observations can lead to bias in both the forecast and analysis state estimations in NWP models. Furthermore, Ridler et al. [112] propose a bias-aware Ensemble Transform Kalman Filter (ETKF) to estimate and correct bias in the observations in a complex groundwater model. Finally, Jin et al. [22] proposed a machine learning approach to correct dust storm observations when assimilated into a chemical transport model. This study concludes that assimilation in a bias-blind setting can lead to worse results than standard forecast and that, as expected, the best results are obtained when using machine learning-based bias-aware assimilation.

2.4.1 Colored-noise Kalman Filter (ColKF)

The ColKF (and similar state augmentation methods) is probably one of the most common approaches to bias estimation. Often, DA literature does not explicitly refer to it as a colored noise approach. For instance, Evensen [66] explained how to implement an Auto-Regressive (AR) error process to account for model error in an EnKF.

Erdal, Neuweiler, and Wollschläger [11] investigates the use of a feedback ColKF to account for unresolved structure in unsaturated zone modeling. The authors find that introducing bias corrections in the EnKF can significantly improve the predictive capability of a simplified model. The study highlights the importance of considering model inadequacy in data assimilation for unsaturated zone modeling, particularly due to unresolved structure. Moreover, the authors emphasize the importance of imposing spatial correlation when generating the noise process for the bias forecast. Otherwise, the filter cannot properly correct the bias terms at locations where observations are not available.

Additionally, Sørensen and Madsen [113] use an approach similar to the ColKF but with different underlying methods to perform the assimilation. This paper investigates the sensitivity of three filter schemes: the EnKF, the RRSQRT, and the Steady-State Kalman Filter (SSKF), in assimilating water levels in a three-dimensional hydrodynamic model of shelf seas. The paper incorporates a stochastic propagation operator that includes the numerical and an AR(1) model to model system errors. This process is used to incorporate the errors in boundary conditions and wind velocity. The study explores the sensitivity of error covariance matrices within the Kalman filter schemes and their impact on assimilation performance, aiming to ensure these schemes perform well under parameter uncertainty.

Da Silva and Colonius [114] proposes a bias-aware EnKF estimator to handle biases in both forecast and observation models for aerodynamic flows. The authors suggest a method that decomposes the bias into slow and fast components, modeling the former as colored noise processes and the latter as white-noise processes. Addressing both discretization and model-form errors improved the accuracy of state estimations in fluid systems.

Furthermore, Chumchean, Seed, and Sharma [115] explores a KF method to adjust radar rainfall estimates by correcting mean field bias in real time. They model the logarithmic mean bias as an AR(1). This model assumes that the error in the radar estimates is not only spatially invariant but also exhibits Markovian dependence, meaning the current error is correlated with the past errors in a predictable manner. This statistical technique allows for dynamic adjustment of the bias, improving the accuracy of the radar rainfall estimates over time.

Moreover, Raboudi et al. [116] focus on EnKF techniques in systems with colored observation noise. The authors introduce two algorithms, extensions of existing KF methods, specifically designed to accommodate colored noise scenarios. They argue that incorporating temporal autocorrelation leads to a more accurate representation of the underlying processes. Although observation noise is not treated in this work, the concept of ColKF can also be extended to accommodate observation bias.

Finally, the book by Chui and Chen [12, Ch. 5] on Kalman filtering presents a slightly different approach to including a colored-noise process into a KF. The authors assume that the system and observation models are affected only by colored-noise processes. While the same state augmentation approach as the ColKF is followed, the derivation shows that modified versions of the initialization of the filter, state and covariance recursions, and Kalman gain are obtained. While this approach is interesting from a theoretical standpoint, the simpler formulation of the ColKF proposed by Drécourt, Madsen, and Rosbjerg [7] was used in this thesis.

2.4.2 Separate-bias Kalman Filter (SepKF)

Although the SepKF is not used in this work, it is still one of the most used methodologies to perform sequential bias-aware DA, and it has been employed in various research disciplines. Thus, some of the references in the literature are worth mentioning.

Cao et al. [117] implemented a SepKF to correct system bias of a soil moisture model. The experiments were tested using constant- and sinusoidal-type bias in a synthetic experiment and finally in a real setting using on-site observations from Northwest China. Moreover, Nerger and Gregg [118] implemented the same two-stage filter for state and bias estimation in an ocean-biochemical model (specifically for estimating surface chlorophyll concentrations). Their approach uses a local Singular Evolutive Interpolated Kalman (SEIK) filter for data assimilation, in contrast with the commonly used KF (or EnKF). Furthermore, Chepurin, Carton, and Dee [119] used a SepKF for a tropical Pacific Ocean circulation model, and Deng, Tang, and Wang [120] applied this filter to assimilate temperature and salinity in a primitive equation model of the Pacific Ocean.

In the works by De Lannoy et al. [121] and De Lannoy et al. [122], the authors explore the estimation of state and bias in a soil moisture model. In particular,

[121] focuses on the estimation performance based on the assimilation experiment setup (e.g. assimilation frequency and density), while [122] focuses on the bias estimation pipeline. Given its systematic approach to testing variations of the SepKF, the latter work's importance must be highlighted. In particular, the authors propose five variations of the two-stage filter approach, depending on how the forecast and analysis states are computed and how the forward model is re-initialized after analysis. These approaches also explicitly describe whether state and bias update is performed and if the filter coincides with a feedback or no feedback formulation.

Another variation of the SepKF was presented by Pauwels et al. [123], where they simultaneously estimate model and observation bias, as well as system state, in a scheme similar to the one already proposed by Dee and Da Silva [6], and using an EnKF. The central assumption for the derived filtering scheme is that the observation and forecast errors are independent of each other and of the error of the unbiased model state variables. This is not the case for other state augmentation techniques. Although the derivations are presented for a linear case, a hybrid approach (by [122]) is followed in the test cases: the state uses an EnKF whereas the bias filters use an exact KF to perform the analysis. Finally, despite the setup assumptions, reasonable results can still be obtained in a rainfall-runoff model in both a synthetic and real data assimilation experiment.

Draper et al. [124] introduce a two-stage filter that dynamically estimates and removes systematic observation-minus-forecast (O-F) biases from observations, enabling the correction of model errors at sub-seasonal scales. This approach parallels the SepKF developed by Dee and Da Silva [6]. Still, it differs from the original by Friedland [92] in optimizing an analysis equation that includes the estimated bias, obtaining a modified Kalman gain. This method is demonstrated by assimilating geostationary skin temperature observations into the Catchment land surface model.

Although not directly for bias estimation, another similar idea to the SepKF was presented by Yang and Delsole [125] for parameter estimation. In their work, the authors show that in two particular cases of parameter structure (additive and multiplicative), the analysis can be performed with two separate filters, one for the state and one for the parameters.

3

Theory

This chapter contains the theoretical background, methods, and testing models used throughout this study. Firstly, a brief introduction to discrete-time dynamical systems is provided, followed by a general description of the filtering problem in the context of sequential data assimilation. Then, sequential data assimilation (KF) and ensemble methods (EnKF) are presented. Afterward, the concept of bias in data assimilation systems is discussed, and a detailed description of the selected bias-aware filtering technique is given (ColKF). The chapter concludes by introducing the main testing models used for experimentation.

3.1 The Dynamic Model

Processes are often nonlinear in real-world applications, making these models a central topic of interest when utilizing DA techniques. While the original derivations of these methods are typically presented for linear cases, various extensions have been developed to handle nonlinear dynamics [54, 56, 51]. Among these extensions, ensemble methods have emerged as a superior alternative since they can effectively propagate nonlinear dynamics while still performing "optimal" linear corrections [65, 35, 66]. This work provides a comprehensive theoretical background, including linear filtering, to ensure the study is self-contained.

In general, a discrete-time stochastic system can be described as

$$\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}) + \mathbf{w}_k, \tag{3.1a}$$

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \mathbf{v}_k, \tag{3.1b}$$

$$\mathbf{x}_0 \sim G(\boldsymbol{\mu}_0, \mathbf{P}_0) \tag{3.1c}$$

$$\mathbf{w}_k \sim G(\mathbf{0}, \mathbf{Q}_k),\ \forall k \in \mathbb{Z}_+, \tag{3.1d}$$

$$\mathbf{v}_k \sim G(\mathbf{0}, \mathbf{R}_k),\ \forall k \in \mathbb{Z}_+, \tag{3.1e}$$

where $\mathcal{M}_k : \mathbb{R}^{n_\mathbf{x}} \times \mathbb{R}^{n_\mathbf{u}} \to \mathbb{R}^{n_\mathbf{x}}$ is a nonlinear forward operator (with parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{n_{\boldsymbol{\theta}}}$) that takes the state $\mathbf{x}_k$ with input $\mathbf{u}_k$ from the time step k to

$k+1$, and $\mathcal{H}_k : \mathbb{R}^{n_\mathbf{x}} \to \mathbb{R}^{n_\mathbf{y}}$ is a general observation operator that extracts the observable output $\mathbf{y}$ from the state at time k. The vectors $\mathbf{w}_k$ and $\mathbf{v}_k$ represent the process and observation noise, respectively, which, by the specification above, are both time-uncorrelated Gaussian processes with covariance matrices $\mathbf{Q}_k \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}_{\text{spsd}}$ and $\mathbf{R}_k \in \mathbb{R}^{n_\mathbf{y} \times n_\mathbf{y}}_{\text{spsd}}$, $\forall k \in \mathbb{Z}_+$, respectively. Finally, $\boldsymbol{\mu}_0 \in \mathbb{R}^{n_\mathbf{x}}$, and $\mathbf{P}_0 \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}_{\text{spsd}}$ represent the mean and covariance of the initial state of the dynamic system.

In the context of DA in geophysics, the model (3.1) is often referred to as the *forward*, *prediction*, or *forecast* model since it is the numerical approximation of a real-world phenomenon that is used to predict the state until the following observation is available for assimilation. Moreover, it is essential to mention that a system like (3.1) is often the result of a time-discretization of an ODE, generally in the state space form of

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t); \boldsymbol{\theta}), \tag{3.2a}$$

$$\mathbf{y}(t) = h(\mathbf{x}(t)), \tag{3.2b}$$

which, at the same time, in the context of geophysics, is often the result of a spatial discretization of a Partial Differential Equation (PDE). The noise processes $\mathbf{w}_k$ and $\mathbf{v}_k$ in (3.1) are a way to add the uncertainty of the system, resulting from modeling errors, uncertain parameters, discretization errors, etc. The discrete-time formulation in (3.1) allows for easier manipulation and estimation of said stochasticity, which would not be the case if uncertainty is directly considered in (3.2). The latter case would take us to a Stochastic Differential Equation (SDE), which is known to be more challenging to treat.

Finally, the subscript k used throughout this thesis, e.g., in eq. (3.1), refers to a discrete-time index. However, in the context of DA, it usually refers to the so-called assimilation time, which is the instant at which the model prediction will be corrected by the DA framework. Formally, suppose the continuous model runs on a time interval $[0, T]$, then the subscripts k refer to finite sub-sequence $\{t_k\}_{k=0}^{K} \subseteq [0, T]$ of assimilation times. This will be better explained in Section 4.1.1. This is done mainly to keep the notation simple since the results in sequential DA are (generally) the same, independently of the time between observations.

3.2 Sequential Data Assimilation

3.2.1 Kalman Filter (KF)

This section gives an overview of the filtering problem for a linear Gaussian system for which the solution results in the well-known KF. This section is based on ideas presented by Lewis, Lakshmivarahan, and Dhall [126] and Schuppen [63]. The traditional KF is realistically not applicable in many DA applications, including the object of this work. However, this result is used by many other sequential DA methods, and its importance cannot be overstated.

Hence, consider the following discrete-time linear Gaussian system

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{M}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \end{aligned} \tag{3.3}$$

where the initial condition $\mathbf{x}_0$, and processes $\mathbf{w}_k$ and $\mathbf{v}_k$, follow the same specification as eqs. (3.1c) to (3.1e) respectively. Additionally, the matrix $\mathbf{M}_k \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}$ is the state-transition matrix, $\mathbf{B}_k \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{u}}$ maps the inputs to the states, and $\mathbf{H}_k \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{y}}$ is the observation model, and in general, all of these matrices may depend on time, thus the subscript. Note that system (3.3) is a particular case of (3.1) by setting $\mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}) \equiv \mathbf{M}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k$ and $\mathcal{H}_k(\mathbf{x}_k) \equiv \mathbf{H}_k \mathbf{x}_k$. Note that the parameter vector $\boldsymbol{\theta}$ is left out from (3.1) for two reasons: first, this simplifies the notation for this section. Second, if the parameter vector $\boldsymbol{\theta}$ were uncertain and its estimation were needed, the formulation (3.3) would no longer be linear, and the KF would no longer be applicable (see Section 3.2.4).

Nevertheless, the linear filtering problem can then be formulated as follows: given that the state $\mathbf{x}_k$ evolves according to (3.3) and given that set of observations $Y_{k-1} := \{\mathbf{y}_s : 1 \leq s \leq k-1\}$ is available, the objective is to find an estimate $\hat{\mathbf{x}}_k$ of the actual state of the system $\mathbf{x}_k$ that minimizes the conditional mean squared error $\mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T(\mathbf{x}_k - \hat{\mathbf{x}}_k)|Y_{k-1}\right]$ [126].

The derivation of the filter is omitted, but the reader can refer to [127] for a comprehensive derivation from a least-squares approach and to [63] for a conditional distribution approach. The KF is then given by the forecast (prediction) recursions

$$\mathbf{x}^f_{k+1} = \mathbf{M}_k \hat{\mathbf{x}}_k + \mathbf{B}_k \mathbf{u}_k \tag{3.4}$$

$$\mathbf{P}^f_{k+1} = \mathbf{M}_k \hat{\mathbf{P}}_k \mathbf{M}_k^T + \mathbf{Q}_k \tag{3.5}$$

$$\hat{\mathbf{x}}_0 = \boldsymbol{\mu}_0, \; \hat{\mathbf{P}}_0 = \mathbf{P}_0,$$

and the assimilation (analysis) recursions

$$\hat{\mathbf{x}}_k = \mathbf{x}_k^f + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f\right) \tag{3.6}$$

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \tag{3.7}$$

$$\hat{\mathbf{P}}_k = \left(\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}_k \mathbf{H}_k\right) \mathbf{P}_k^f \tag{3.8}$$

where $\mathbf{I}_n$ represent the identity matrix of size $n \times n$, and the estimated state and covariance respectively satisfy

$$\begin{aligned} \hat{\mathbf{x}}_k &= \mathbb{E}\left[\mathbf{x}_k | Y_{k-1}\right], \\ \hat{\mathbf{P}}_k &= \mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T | Y_{k-1}\right] \\ &= \mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T\right]. \end{aligned}$$

Additionally, note how the analysis equation (3.6) can be rewritten as

$$\hat{\mathbf{x}}_k = \left(\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}_k \mathbf{H}_k\right) \hat{\mathbf{x}}_k + \mathbf{K}_k \mathbf{y}_k,$$

which resembles a linear interpolation formula between two points. The analysis equation can be interpreted as a weighted sum between the observations and the state forecast to produce a corrected prediction, on which the weight is given by the Kalman gain $\mathbf{K}_k$. The resulting analysis $\hat{\mathbf{x}}$ is an unbiased minimum variance estimate of the state. Moreover, the KF recursions result in what is known as a Best Linear Unbiased Estimate (BLUE), even without the Gaussian assumptions [2, 29].

Finally, the concept of innovation is usually mentioned when talking about the KF. In this work, the innovation process is assumed to refer to the difference between the analysis and the real measurement, i.e., $\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k$. Moreover, it is well-known that the Kalman update guarantees that this error is a white noise process [63] specified by

$$\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k \sim G\left(\mathbf{0}, \mathbf{H}_k \hat{\mathbf{P}}_k \mathbf{H}_k^T + \mathbf{R}_k\right). \tag{3.9}$$

This result will be used later to assess the filter's performance. An innovation process whose expected value is significantly different than zero strongly indicates systematic bias in the model.

3.2.2 Ensemble Kalman Filter (EnKF)

Although the KF provides a closed optimal form for the filtering problem, the size of the state vectors makes the KF recursions computationally unfeasible for use in an operational setting, particularly the storage and propagation of the forecast covariance matrix $\mathbf{P}_k^f$ in (3.5). Moreover, the KF is formulated as the solution to the filtering problem in linear Gaussian systems, yet most forecast models are nonlinear, as (3.1).

The main idea of the EnKF is to use an ensemble of states that are propagated independently to create a sample containing the stochastic component of the filter. From this sample, the forecast state $\mathbf{x}_k^f$ and covariance matrix $\mathbf{P}_k^f$ can be estimated during the assimilation step. This makes the implementation more computationally efficient, as the explicit update for $\mathbf{P}_k^f$ is replaced by a simple sample covariance estimator using the ensemble states [65, 67].

To allow for a comparison with the traditional KF, the linear case of the EnKF will first be introduced. That is, for a system in the form of (3.3). This section is mainly inspired by the formulation presented by Lewis, Lakshmivarahan, and Dhall [126], although the EnKF was presented first by Evensen [65].

First, each member $\hat{\boldsymbol{\xi}}_{0,i}$, $i = 1, \dots, N$ of the ensemble is initialized by drawing a sample from the initial state Gaussian distribution $G(\boldsymbol{\mu}_0, \mathbf{P}_0)$. Now, let $\boldsymbol{\xi}_{k,i}^f$ and $\hat{\boldsymbol{\xi}}_{k,i}$ denote the forecast and analysis state vectors of the i-th ensemble member, respectively, at time k. The time updates are now given by

$$\boldsymbol{\xi}_{k+1,i}^f = \mathbf{M}_k \hat{\boldsymbol{\xi}}_{k,i} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_{k,i}, \ \forall i = 1, 2, \dots, N \tag{3.10}$$

$$\mathbf{x}_k^f = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\xi}_{k,i}^f \tag{3.11}$$

$$\mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^{N} \left(\boldsymbol{\xi}_{k,i}^f - \mathbf{x}_k^f \right) \left(\boldsymbol{\xi}_{k,i}^f - \mathbf{x}_k^f \right)^T, \tag{3.12}$$

where each $\mathbf{w}_{k,i}$ is a sample of $G(\mathbf{0}, \mathbf{Q}_k)$, and the analysis (assimilation) updates are given by

$$\hat{\boldsymbol{\xi}}_{k,i} = \boldsymbol{\xi}_{k,i}^f + \mathbf{K}_k \left(\mathbf{y}_k + \mathbf{v}_{k,i} - \mathbf{H}_k \boldsymbol{\xi}_{k,i}^f \right), \ \forall i = 1, 2, \dots, N \tag{3.13}$$

$$\hat{\mathbf{x}}_k = \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\xi}}_{k,i} \tag{3.14}$$

$$\hat{\mathbf{P}}_k = \frac{1}{N-1} \sum_{i=1}^{N} \left(\hat{\boldsymbol{\xi}}_{k,i} - \hat{\mathbf{x}}_k \right) \left(\hat{\boldsymbol{\xi}}_{k,i} - \hat{\mathbf{x}}_k \right)^T \tag{3.15}$$

where each $\mathbf{v}_{k,i}$ is a sample of $G(\mathbf{0}, \mathbf{R}_k)$, and where $\mathbf{K}_k$ is the given by (3.7), same as for the exact KF. Note that in steps (3.10) and (3.13) lay a major difference with respect to the exact KF. These steps perform what is known as *stochastic updates*: adding realizations of the system and observation error processes ($\mathbf{w}_k$ and

$\mathbf{v}_k$ respectively).

In particular, (3.13) uses the so-called *virtual observations*, which refer to the perturbation of $\mathbf{y}_k$ with the associated observation error process $\mathbf{v}_k$ inside the parenthesis. These perturbed observations are required to maintain consistency of this formulation with the traditional KF [128, 126]. Meaning that, as the ensemble size $N \to \infty$, the EnKF converges to the KF in the case of linear dynamics. However, the so-called *square-root* schemes proposed in the literature avoid perturbing the measurements by preserving the exact covariance in the update [29]. Examples of these methods include the ETKF [129, 130], EAKF [131], the Maximum Likelihood Ensemble Kalman Filter (MLEKF) [132], and the Deterministic Ensemble Kalman Filter (DEnKF) [133].

Finally, another noteworthy remark is that the estimators provided in (3.12) and (3.15) are typical unbiased estimators of a sample covariance matrix, in this case of $\mathbf{P}_k^f$ and $\hat{\mathbf{P}}_k$ respectively. However, in theory, this step can be replaced by any sample covariance estimator, like covariance shrinkage estimation [134, 135, 136] or other methodologies.

One significant advantage of the EnKF is that it only uses the state transition matrix $\mathbf{M}_k$ when forecasting the ensemble. This allows the framework to be extended naturally to a model similar to (3.1). In the context of nonlinear models, the traditional KF can still be applied by computing the Jacobian of the nonlinear model $\mathbf{M}_k \approx \frac{dM_k}{d\mathbf{x}}$. This is known as the EKF.

However, in the EnKF, (3.1) can be used directly to forecast the next state. Therefore, assuming the observation model is still linear, the EnKF procedure remains unchanged, with the exception of (3.10), which is replaced with

$$\boldsymbol{\xi}_{k+1,i}^{f} = \mathcal{M}_k(\hat{\boldsymbol{\xi}}_{k,,}\mathbf{u}_k;\boldsymbol{\theta}) + \mathbf{w}_{k,i} \tag{3.16}$$

which corresponds to a nonlinear stochastic prediction step for the i-th ensemble member. Consequently, through the remainder of this work, the observation operator in (3.1b) is assumed to be linear, that is, $\mathcal{H}_k(\cdot) \equiv \mathbf{H}_k \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{y}}$, $\forall k$. Since the analysis step explicitly requires the observation matrix, the complication of computing the Jacobian of $\mathcal{H}_k$ is avoided by making this assumption. Moreover, since the matrix $\mathbf{M}_k$ is no longer required for computing $\mathbf{P}_k^f$, the state can still be propagated using the nonlinear model. In summary, EnKFs work with a system in the form

$$\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}) + \mathbf{w}_k, \tag{3.17a}$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \tag{3.17b}$$

where the initial condition $\mathbf{x}_0$, and processes $\mathbf{w}_k$ and $\mathbf{v}_k$, follow the same specification as (3.1c), (3.1d), and (3.1e), respectively. Unless otherwise specified, this is the *forecast* model assumed for the remainder of this work.

3.2.3 Covariance Localization

This section gives a short introduction to the concept of covariance localization. Since the EnKF is a Monte Carlo approximation of the exact KF, the size of the ensemble determines how accurate the estimation of the state mean and covariance is. However, computational resources limit the number of ensemble members, which is often low compared to the number of states. This results in problems such as undersampling, inbreeding, filter divergence, and spurious correlations [137]. One of the strategies to overcome these problems is covariance localization.

Covariance localization [138, 139, 130] is a process by which the covariance of "distant" states is cut off in the forecast covariance matrix $\mathbf{P}_k^f$. It is typically done by applying a Hadamard (Schur) product between the forecast error covariance matrix $\mathbf{P}_k^f$ and a localization matrix $\Psi \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}$ (sometimes called the *mask*). Firstly, the Hadamard product, also known as element-wise product, between two matrices $A, B \in \mathbb{R}^{n \times n}$, denoted $A \odot B$, is given by

$$(A \odot B)_{ij} = A_{ij} B_{ij}, \ \forall i, j = 1, \ldots, n.$$

Secondly, the elements of the localization matrix Ψ satisfy

$$\Psi_{ij} = \psi(z_{ij}), \ \forall i, j = 1, \ldots, n, \tag{3.18}$$

where z_{ij} is the distance between the grid points associated with the states x_i and x_j in the physical space, and where ψ is a local correlation function. Hence, the Kalman gain (3.7) is replaced by

$$\mathbf{K}_k = \left(\Psi \odot \mathbf{P}_k^f\right) \mathbf{H}_k^T \left(\mathbf{H}_k \left(\Psi \odot \mathbf{P}_k^f\right) \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1}, \tag{3.19}$$

during the analysis step. It can be proved that Ψ is also a covariance matrix, and thus, so is $\Psi \odot \mathbf{P}_k^f$. The reader is referred to the dissertation by Petrie [137], which provides a detailed description of localization for the EnKF.

One typical local correlation function ψ is the 5^{th}-order piecewise rational function with compact support proposed by Gaspari and Cohn [140], given by

$$\psi(z) = \begin{cases} -\frac{1}{4}\left(\frac{|z|}{c}\right)^5 + \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 - \frac{5}{3}\left(\frac{|z|}{c}\right)^2 + 1, & 0 \leq |z| \leq c, \\ \frac{1}{12}\left(\frac{|z|}{c}\right)^5 - \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 + \frac{5}{3}\left(\frac{|z|}{c}\right)^2 - 5\left(\frac{|z|}{c}\right) + 4 - \frac{2}{3}\left(\frac{c}{|z|}\right) & c \leq |z| \leq 2c, \\ 0, & |z| \geq 2c, \end{cases}$$

and where c is a length scale defined such that the correlation reduces to zero after more than twice this distance and following a Gaussian-like bell curve within this scale. A plot of the local correlation function is shown in Fig. 3.1 for different values of c.

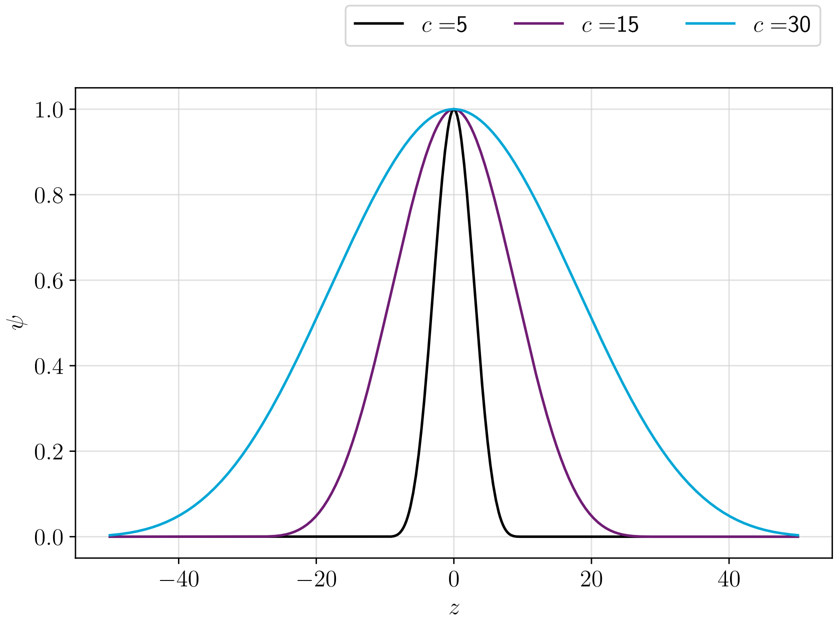


Figure 3.1: Local correlation function ψ for different length scales c.

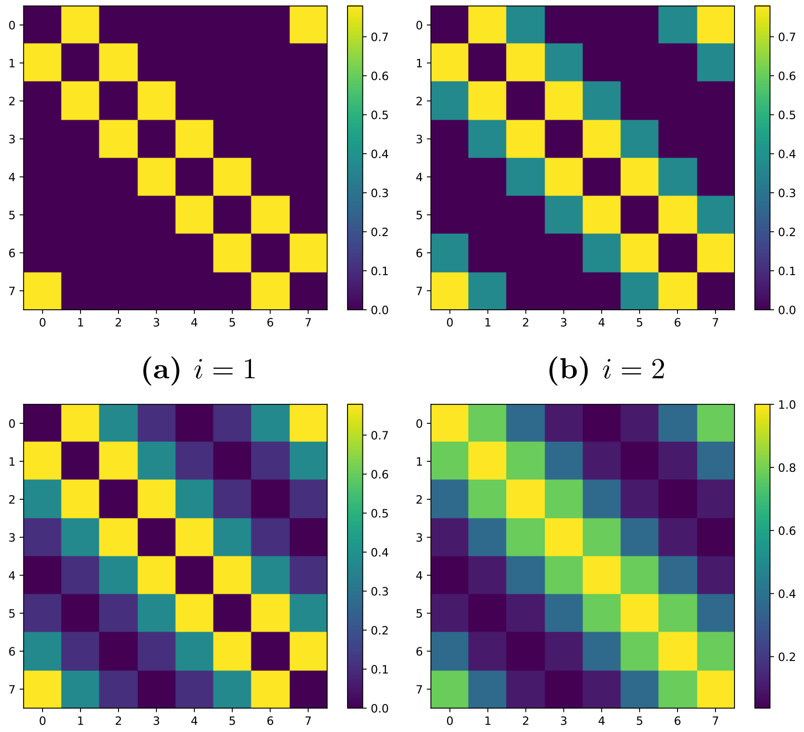
However, this function is more common in geophysical applications, where an Euclidean distance exists between the grid cells in the state [137]. Since this work only focuses on the synthetic case scenario using a Lorenz-96 model, we use a Gaussian decay approach. This approach was recently used by Diab Montero [9] to create a localization mask for the model based on the state indices rather than physical distance. His implementation was modified to include the cyclic nature of the boundary of the Lorenz-96 model.

The pseudocode of this procedure is presented in Algorithm 1. In this procedure, the matrix $\mathbf{D}_i \in \mathbb{R}^{n_\mathsf{x} \times n_\mathsf{x}}$ represents a matrix with ones on the i-th off-diagonal entries. In this manner, $i > 0$ represents diagonals above the main diagonal, and $i < 0$ represents diagonals below the main diagonal. As a reference, $\mathbf{D}_0 \equiv \mathbf{I}_{n_\mathsf{x}}$.

Algorithm 1 Localization mask generation.

```
Input: $r \in \mathbb{Z}_+$                              ▷ The covariance influence radius
Input: $n_\mathbf{x} \in \mathbb{Z}_+$                   ▷ The size of the state vector
$\Psi \Leftarrow \mathbf{0}_{n_\mathbf{x} \times n_\mathbf{x}}$                  ▷ Initialize the mask
for $i = 1$ to $3r$ do
    $\alpha \Leftarrow \exp\left(-\left({}^{i}\!/_{r}\right)^2\right)$
    $\Psi \Leftarrow \Psi + \alpha\left(\mathbf{D}_i + \mathbf{D}_{n_\mathbf{x}-i}\right)$     ▷ Upper diagonals
    $\Psi \Leftarrow \Psi + \alpha\left(\mathbf{D}_{-i} + \mathbf{D}_{i-n_\mathbf{x}}\right)$    ▷ Lower diagonals
$\Psi \Leftarrow \Psi + \mathbf{I}_{n_\mathbf{x}}$
Output: $\Psi \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}$                 ▷ The localization mask
```

This procedure is depicted in Fig. 3.2 for a test case with $n_\mathbf{x} = 8$ cells and a covariance influence radius of $r = 2$. Note how the off-diagonal entries are progressively filled with the Gaussian decay coefficient in the nearest and furthest points, correctly accounting for the cyclic boundary of the Lorenz-96 model. The last panel shows the resulting covariance mask.



(a) $i = 1$

(b) $i = 2$

(c) $i = 3$

(d) Final mask

Figure 3.2: Sequential construction of the covariance mask for localization using Gaussian decay. Experiment with $n_\mathbf{x} = 8$ cells and a covariance influence radius of $r = 2$.

3.2.4 Parameter and State Estimation

The forecast model is an abstraction of natural processes. It includes parameters representing real-world factors that may not be directly measurable or rely on assumptions that may not be entirely correct. Therefore, parameter estimation is commonly part of the DA workflow so that models can closely align with actual conditions. This, in turn, leads to better simulations and more reliable forecasts, improving our ability to produce valid estimates.

For parameter estimation in a filtering setting, the basic idea is that the parameters are deemed "variable" in time $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}_k$, and they are considered part of an augmented state vector. The parameters are assumed to follow a random walk model [10], given by

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{w}_k^{\boldsymbol{\theta}}, \tag{3.20}$$

where $\mathbf{w}_k^{\boldsymbol{\theta}} \sim G\left(\mathbf{0}, \mathbf{Q}_k^{\boldsymbol{\theta}}\right)$, and where $\mathbf{Q}_k^{\boldsymbol{\theta}} \in \mathbb{R}_{\text{spsd}}^{n_{\boldsymbol{\theta}} \times n_{\boldsymbol{\theta}}}$ such that $\mathbf{Q}_k^{\boldsymbol{\theta}} = \operatorname{diag}\left(\sigma_{k,1}^2, \ldots, \sigma_{k,n_{\boldsymbol{\theta}}}^2\right)$. Usually, in an applied setting, all parameter forecast errors are taken as zero ($\mathbf{Q}_k^{\boldsymbol{\theta}} \equiv \mathbf{0}$, $\forall k \in \mathbb{Z}_+$) to represent a persistent model (constant parameter) at the forecast step [8]. This means that the parameters are only updated in the analysis step as a consequence of the learned covariance between them and the state through the observation assimilated. However, we assume that $\mathbf{Q}_k^{\boldsymbol{\theta}}$ can be non-zero to keep generality in the formulation.

This specification leads to an augmented formulation of (3.17) as

$$\mathbf{x}_{k+1} = \mathcal{M}_k\left(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}_k\right) + \mathbf{w}_k, \tag{3.21a}$$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{w}_k^{\boldsymbol{\theta}}, \tag{3.21b}$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \tag{3.21c}$$

which can be summarized into a new stochastic system formulation, using an augmented state vector $\mathbf{z}_k := \begin{bmatrix}\mathbf{x}_k^T & \boldsymbol{\theta}_k^T\end{bmatrix}^T$ as

$$\mathbf{z}_{k+1} = \widetilde{\mathcal{M}}_k(\mathbf{z}_k, \mathbf{u}_k) + \tilde{\mathbf{w}}_k, \tag{3.22a}$$

$$\mathbf{y}_k = \tilde{\mathbf{H}}_k \mathbf{z}_k + \mathbf{v}_k, \tag{3.22b}$$

where

$$\widetilde{\mathcal{M}}_k(\mathbf{z}_k, \mathbf{u}_k) = \begin{bmatrix}\mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}_k) \\ \boldsymbol{\theta}_k\end{bmatrix}, \quad \tilde{\mathbf{w}}_k = \begin{bmatrix}\mathbf{w}_k \\ \mathbf{w}_k^{\boldsymbol{\theta}}\end{bmatrix}, \quad \text{and } \tilde{\mathbf{H}}_k = \begin{bmatrix}\mathbf{H}_k & \mathbf{0}\end{bmatrix},$$

and this is just a new form of (3.17) that can be used with an EnKF scheme to perform DA.

3.3 Bias-Aware Data Assimilation

3.3.1 Bias & Feedback

Several sources in the literature discuss how the KF remains optimal as long as, among other assumptions, the conditions (3.1d) and (3.1e) are satisfied [6, 4, 7]. Therefore, in the presence of model bias, the error of the state $\mathbf{x}_k$ would not meet this requirement. In this manner, if an estimate of the model bias $\mathbf{b}_k \in \mathbb{R}^{n_\mathbf{b}}$ is available, one could compute an unbiased state vector

$$\tilde{\mathbf{x}}_k = \mathbf{x}_k - \mathbf{H}^\mathbf{b}\mathbf{b}_k, \tag{3.23}$$

and perform DA using $\tilde{\mathbf{x}}_k$, where $\mathbf{H}^\mathbf{b} \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{b}}$ is a map from the bias state to the system state in a case where they are not of the same length.

In this sense, if we assume that the dynamics driving $\mathbf{b}_k$ also satisfy the Gaussian white noise requirements, one could theoretically achieve optimality using the KF. However, when and how the state should be corrected within the DA procedure is still an open question. To address this, Drécourt, Madsen, and Rosbjerg [7] presented the concept of having feedback and no feedback in the bias-aware filter. Although they did not introduce this concept, they provided a unified formulation that included both cases. The initial formulation without feedback was introduced by Friedland [92], while the feedback version was introduced later by Dee and Da Silva [6]. These concepts are hereby discussed.

In concise terms, in a *no-feedback* filter, the estimated bias does not directly affect the model dynamics. It only influences the state during the filtering update, making it well-suited for settings where the bias is relatively constant. Whereas a *feedback* filter interacts directly with the model dynamics, which can result in more complex behavior.

In more technical terms, in a bias-aware filter where there is *no-feedback*, the forward model uses the **biased** state vector $\mathbf{x}$ as input to produce a forecast state. This method is ideal for scenarios where the bias's nature is partially understood or remains relatively stable over the assimilation run time. Moreover, when there is *feedback*, the forward model uses the **bias-corrected** estimate of the state vector $\tilde{\mathbf{x}}_k$ as input to produce a forecast state. As the state is corrected at every integration time step, the estimated process $\mathbf{b}_k$ represents the bias' time derivative, allowing for more complex forms of bias that can be estimated [7]. However, two versions of the feedback filter can be distinguished, depending on how the model is corrected using the bias estimate, as it can be done at either the continuous or discrete formulation of the forward model.

Additionally, Drécourt, Madsen, and Rosbjerg [7] also introduced the feedback matrix $\mathbf{F} \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}$ to the mathematical formulation of the bias-aware filters. In particular $\mathbf{F} \equiv \mathbf{0}$ means that there is no feedback, whereas $\mathbf{F} \equiv \mathbf{I}_{n_\mathbf{x}}$ if there is feedback. The same notation is maintained in this work. A graphical representation of the feedback switch in the filtering process is presented in Fig. 3.3, and its effects on the dynamics are presented in Fig. 3.4a in the case of feedback and in Fig. 3.4b for the no-feedback case.

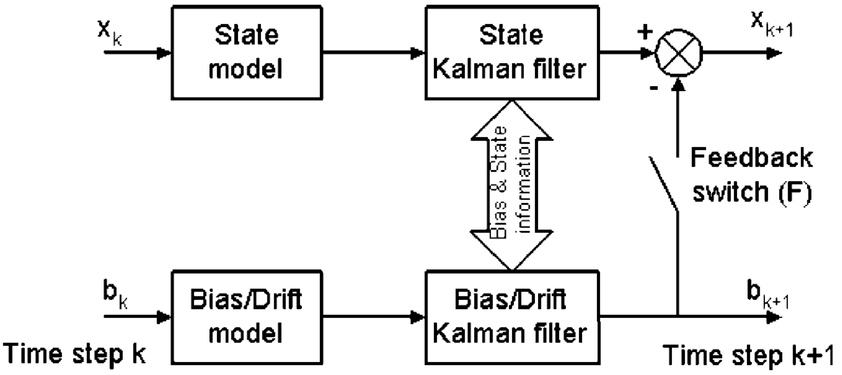
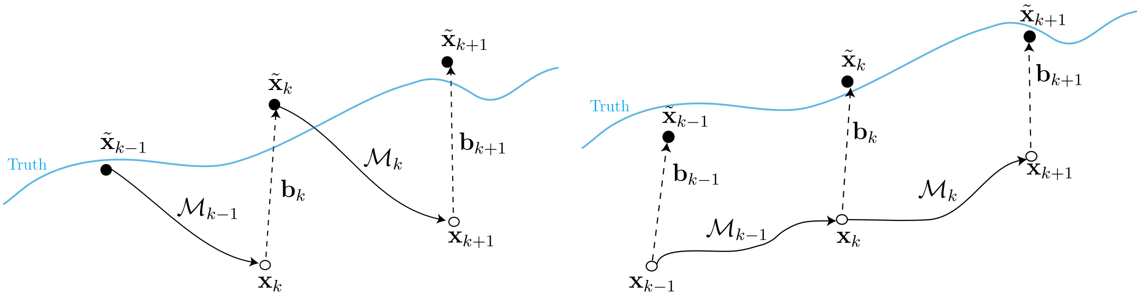


Figure 3.3: Inclusion of the feedback switch on the state and bias estimation process. Figure by Drécourt, Madsen, and Rosbjerg [7].



(a) State dynamics in a feedback setting. The unbiased state $\tilde{\mathbf{x}}$ is used as an initial condition to forecast the next state.

(b) State dynamics in a no-feedback setting. The biased state $\mathbf{x}$ is used as an initial condition to forecast the next state.

Figure 3.4: Illustrations of the concept of feedback for bias-aware filters, based on the ones presented by Baek et al. [99].

In light of these concepts, and inspired by the formulation of Drécourt, Madsen, and Rosbjerg [7], a general formulation of a bias-aware discrete stochastic system with uncertain parameters is proposed in the form

$$\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}_k) - \mathbf{F}\mathbf{H}^{\mathbf{b}}\mathbf{b}_k + \mathbf{w}_k, \tag{3.24a}$$

$$\mathbf{b}_{k+1} = \mathcal{G}_k(\mathbf{x}_k, \mathbf{b}_k) + \mathbf{w}_k^{\mathbf{b}}, \tag{3.24b}$$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{w}_k^{\boldsymbol{\theta}}, \tag{3.24c}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + (\mathbf{F} - \mathbf{I}_{n_\mathbf{x}})\,\mathbf{H}_k\mathbf{H}^{\mathbf{b}}\mathbf{b}_k + \mathbf{v}_k, \tag{3.24d}$$

where $\mathbf{b}_k \in \mathbb{R}^{n_\mathbf{b}}$ represents the bias state at time k, which is propagated through the parameterization $\mathcal{G}_k : \mathbb{R}^{n_\mathbf{x}} \times \mathbb{R}^{n_\mathbf{b}} \to \mathbb{R}^{n_\mathbf{b}}$. Moreover, the process $\mathbf{w}_k^{\mathbf{b}} \sim G\left(\mathbf{0}, \mathbf{Q}_k^{\mathbf{b}}\right)$, and where $\mathbf{Q}_k^{\mathbf{b}} \in \mathbb{R}_{\text{spsd}}^{n_\mathbf{b} \times n_\mathbf{b}}$. The matrix $\mathbf{F} \in \mathbb{R}^{n_\mathbf{x} \times n_\mathbf{x}}$ is the *feedback* matrix, i.e. $\mathbf{F} \in \{\mathbf{0}, \mathbf{I}_{n_\mathbf{x}}\}$. In practice, the parameterization for the bias propagation $\mathcal{G}_k(\cdot)$ is typically taken as a persistence model, similar to how parameters are modeled. In the case of the ColKF, $\mathcal{G}_k(\cdot)$ represents the AR model.

Nonetheless, a point must be made about how the bias is fed back to the model in a feedback implementation. Based on (3.24), in a feedback formulation, that is $\mathbf{F} = \mathbf{I}_{n_\mathbf{x}}$, the bias $\mathbf{b}_k$ corrects the output of the forward model $\mathcal{M}_k$ directly. However, this bias could be used *inside* the forward model. This issue is raised again in Section 4.1.2, specifying how it can interact with the model before time discretization.

3.3.2 Colored-noise Kalman Filter (ColKF)

One of the main assumptions of the KF is that the model and observation error processes are uncorrelated Gaussian noise with mean zero, as specified in (3.1d) and (3.1e). However, in the presence of systematic errors (biases) not considered explicitly in the model, these processes may no longer satisfy these assumptions. Thus, some approaches have considered more complex forms of uncertainty in the errors.

The ColKF extends the KF, allowing for a correlated error process on both model and observations. Following the methodology presented by Drécourt, Madsen, and Rosbjerg [7], an assumption is made: the colored noise process affects only the forecasting model. This assumption aligns with this work since we focused on addressing model bias (systematic errors) and parameter uncertainty in the forecasting model.

Again, the methodology in the case of a linear system is first described. However, their formulation of the linear ColKF is extended to include the $\mathbf{H}^{\mathbf{b}}$ mapping for situations where $n_{\mathbf{b}} \neq n_{\mathbf{x}}$. Consider the following linear stochastic model

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{M}_k \mathbf{x}_k - \mathbf{F}\mathbf{H}^{\mathbf{b}}\mathbf{b}_k + \mathbf{w}_k \\ \mathbf{b}_{k+1} &= \mathbf{A}_k \mathbf{b}_k + \mathbf{w}_k^{\mathbf{b}} \\ \mathbf{y}_k &= \mathbf{H}_k \mathbf{x}_k + \left(\mathbf{F} - \mathbf{I}_{n_{\mathbf{x}}}\right) \mathbf{H}_k \mathbf{H}^{\mathbf{b}} \mathbf{b}_k + \mathbf{v}_k, \end{aligned} \tag{3.25}$$

where the processes $\mathbf{w}_k$ and $\mathbf{v}_k$ also satisfy (3.1d) and (3.1e) respectively, and similarly $\mathbf{w}_k^{\mathbf{b}} \sim G(\mathbf{0}, \mathbf{Q}_k^{\mathbf{b}})$, $\forall k \in \mathbb{Z}_+$. Processes $\mathbf{w}_k$ and $\mathbf{w}_k^{\mathbf{b}}$ are assumed to be uncorrelated, that is $\mathbb{E}\left[\mathbf{w}_k^{\mathbf{b}} \mathbf{w}_k^T\right] = \mathbf{0}$. The process $\mathbf{b}_k$ represents a first-order multivariate AR model, with coefficients given by the diagonal matrix $\mathbf{A}_k$, and the matrix $\mathbf{F} \in \mathbb{R}^{n_{\mathbf{x}}}$ is constant.

Typically, the coefficients in $\mathbf{A}_k$ of the AR model are constrained to be less or equal to 1. In the context of time series and linear stochastic systems, this ensures that the process is weakly stationary (see, e.g., [141]). However, the update on the KF will prevent the process's divergence [7]; thus, this assumption is no longer necessary.

Moreover, in theory, the matrix $\mathbf{F} \in \mathbb{R}^{n_{\mathbf{x}}}$ can take any value, but, following the description of feedback presented in Section 3.3.1, only the cases $\mathbf{F} \equiv 0$ (no-feedback) and $\mathbf{F} \equiv \mathbf{I}_{n_{\mathbf{x}}}$ (feedback) are considered. The initial condition of the AR(1) process $\mathbf{b}_{-1}$ can be initialized using an educated initial guess or just as zero. However, if this is the case, the mean of the AR(1) will remain zero until the first filter update. The mean of the AR (bias estimate) can only be changed in an assimilation step [7]. In addition, the term $-\mathbf{F}\mathbf{H}^{\mathbf{b}}\mathbf{b}_k$ is taken as negative in the propagation of the state $\mathbf{x}$ to more intuitively resemble (3.23).

This case's derivation of the KF is straightforward. The process starts by making the system uncorrelated by introducing an augmented state vector $\mathbf{z}_k := \begin{bmatrix} \mathbf{x}_k^T & \mathbf{b}_k^T \end{bmatrix}^T \in \mathbb{R}^{n_{\mathbf{x}}+n_{\mathbf{b}}}$ and defining the system as

$$\begin{aligned} \mathbf{z}_{k+1} &= \tilde{\mathbf{M}}_k \mathbf{z}_k + \mathbf{q}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{z}_k + \mathbf{v}_k, \end{aligned} \tag{3.26}$$

where

$$\tilde{\mathbf{M}}_k = \begin{bmatrix} \mathbf{M}_k & -\mathbf{F}\mathbf{H}^{\mathbf{b}} \\ \mathbf{0} & \mathbf{A}_k \end{bmatrix}, \quad \mathbf{q}_k = \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_k^{\mathbf{b}} \end{bmatrix}, \quad \text{and } \mathbf{C}_k = \begin{bmatrix} \mathbf{H}_k & \left(\mathbf{F} - \mathbf{I}_{n_{\mathbf{x}}}\right) \mathbf{H}_k \mathbf{H}^{\mathbf{b}} \end{bmatrix},$$

and applying the KF recursions (3.4)-(3.8), as follows

$$\mathbf{z}^f_{k+1} = \tilde{\mathbf{M}}_k \hat{\mathbf{z}}_k \tag{3.27}$$

$$\mathbf{P}^f_{k+1} = \tilde{\mathbf{M}}_k \hat{\mathbf{P}}_k \tilde{\mathbf{M}}_k^T + \mathbf{U}_k \tag{3.28}$$

$$\hat{\mathbf{z}}_0 = \begin{bmatrix} \boldsymbol{\mu}_0 \\ \mathbb{E}\left[\mathbf{b}_0\right] \end{bmatrix}, \; \hat{\mathbf{P}}_0 = \begin{bmatrix} \mathbf{P}_0 & \mathbf{0} \\ \mathbf{0} & \mathbb{V}\left[\mathbf{b}_0\right] \end{bmatrix},$$

where

$$\mathbf{U}_k = \mathbb{E}\left[\mathbf{q}_k \mathbf{q}_k^T\right] = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{\mathbf{b}}_k \end{bmatrix},$$

and the analysis recursions

$$\hat{\mathbf{z}}_k = \mathbf{z}^f_k + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{C}_k \mathbf{z}^f_k\right), \tag{3.29}$$

$$\mathbf{K}_k = \mathbf{P}^f_k \mathbf{C}_k^T \left(\mathbf{C}_k \mathbf{P}^f_k \mathbf{C}_k^T + \mathbf{R}_k\right)^{-1}, \tag{3.30}$$

$$\hat{\mathbf{P}}_k = \left(\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}_k \mathbf{C}_k\right) \mathbf{P}^f_k. \tag{3.31}$$

Alternatively, in the notation of (3.25), we have

$$\begin{bmatrix} \mathbf{x}^f_{k+1} \\ \mathbf{b}^f_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_k & -\mathbf{F}\mathbf{H}^{\mathbf{b}} \\ \mathbf{0} & \mathbf{A}_k \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_k \\ \mathbf{b}_k \end{bmatrix},$$

$$\mathbf{P}^f_{k+1} = \begin{bmatrix} \mathbf{M}_k & -\mathbf{F}\mathbf{H}^{\mathbf{b}} \\ \mathbf{0} & \mathbf{A}_k \end{bmatrix} \hat{\mathbf{P}}_k \begin{bmatrix} \mathbf{M}_k^T & \mathbf{0} \\ -(\mathbf{F}\mathbf{H}^{\mathbf{b}})^T & \mathbf{A}_k^T \end{bmatrix} + \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}'_k \end{bmatrix},$$

$$\hat{\mathbf{z}}_0 = \begin{bmatrix} \boldsymbol{\mu}_0 \\ \mathbb{E}\left[\mathbf{b}_0\right] \end{bmatrix}, \; \hat{\mathbf{P}}_0 = \begin{bmatrix} \mathbf{P}_0 & \mathbf{0} \\ \mathbf{0} & \mathbb{V}\left[\mathbf{b}_0\right] \end{bmatrix},$$

$$\begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\mathbf{b}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{x}^f_k \\ \mathbf{b}^f_k \end{bmatrix} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}^f_k - \left(\mathbf{F} - \mathbf{I}_{n_\mathbf{x}}\right) \mathbf{H}_k \mathbf{H}^{\mathbf{b}} \mathbf{b}^f_k\right),$$

where

$$\mathbf{K}_k = \mathbf{P}^f_k \begin{bmatrix} \mathbf{H}_k & \left(\mathbf{F} - \mathbf{I}_{n_\mathbf{x}}\right) \mathbf{H}_k \mathbf{H}^{\mathbf{b}} \end{bmatrix}^T \left(\mathbf{R}_k + \begin{bmatrix} \mathbf{H}_k & \left(\mathbf{F} - \mathbf{I}_{n_\mathbf{x}}\right) \mathbf{H}_k \mathbf{H}^{\mathbf{b}} \end{bmatrix} \mathbf{P}^f_k \begin{bmatrix} \mathbf{H}_k & \left(\mathbf{F} - \mathbf{I}_{n_\mathbf{x}}\right) \mathbf{H}_k \mathbf{H}^{\mathbf{b}} \end{bmatrix}^T\right)^{-1}.$$

This presented formulation follows a KF approach, and it illustrates how to convert the colored noise process into a white one. However, in practice, the same augmented state process can be followed with a nonlinear forecast model, using an EnKF to perform the assimilation. This enables the filter to perform nonlinear updates (see Section 3.2.2) while keeping the colored noise modeling approach [66, 7].

3.4 Testing Models

3.4.1 Linear Harmonic Oscillator (LHO)

A simple Linear Harmonic Oscillator (LHO) will initially be used to test the filtering techniques. In simplified notation, the driving ODE associated to a (possibly damped) LHO is given by

$$\ddot{y}(t) - \theta_2 \dot{y}(t) - \theta_1 y(t) = 0, \tag{3.32}$$

where $\theta_1, \theta_2 < 0$. This formulation can be translated into a more typical state-space form

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{pmatrix} \mathbf{x}(t), \tag{3.33}$$

where $\mathbf{x}(t) = [y(t)\ \dot{y}(t)]^T$, and both states are assumed to be observable, i.e., $\mathbf{y}(t) = \mathbf{I}_2\mathbf{x}(t)$. Furthermore, this system's advantages are its simplicity and linearity. The latter allows us to guarantee that Kalman-based analysis produces optimal estimates and validate the implemented methods.

3.4.2 The Lorenz-96 Model

The Lorenz-96 model is a simple nonlinear dynamical system that simplifies atmospheric phenomena to study nonlinear interactions and chaotic behavior. The model was initially presented in 1996 by E. N. Lorenz during an European Centre for Medium-Range Weather Forecasts (ECMWF) workshop on predictability [142, 13], but officially published in 2006 [143] as part of a book in predictability of weather and climate.

The model was developed to study how the error grows along with prediction time in the context of meteorological and atmospheric phenomena. It has allowed researchers to gain insight into problems and nonlinear interactions that often arise with larger models closer to reality [13]. The Lorenz-96 has been broadly used in topics like chaos (e.g. [144, 145, 146, 13]), predictability (e.g. [147, 148, 143]), and data assimilation (e.g. [149, 99, 150, 107, 9]), the latter being the main focus of this work.

The Lorenz-96 belongs to a more general class of dynamical systems extensively studied by Lorenz, namely, the Forced Dissipative Systems (FDS) with quadratic terms [151, 13]. In general, these systems are in the form

$$\dot{x}_j(t) = \sum_{k,l=1}^{n_{\mathbf{x}}} a_{jkl} x_k(t) x_l(t) - \sum_{k=1}^{n_{\mathbf{x}}} b_{jk} x_k(t) + c_j, \quad j = 1, \dots, n_{\mathbf{x}}. \tag{3.34}$$

The Lorenz-96 model was first studied in four dimensions by Lorenz [144] when he was searching for the simplest FDS that exhibits chaotic behavior. The Lorenz-96 model is completely determined by the equation of the j-th state, namely,

$$\dot{x}_j(t) = x_{j-1}(t) \left(x_{j+1}(t) - x_{j-2}(t)\right) - x_j(t) + F, \quad j = 1, \dots, n_{\mathbf{x}} \tag{3.35}$$

with the 'boundary'/cyclic condition $x_{j-n_{\mathbf{x}}}(t) = x_{j+n_{\mathbf{x}}}(t) = x_j(t)$, $\forall j = 1, \dots, n_{\mathbf{x}}$, and where $F \in \mathbb{R}$ is a constant parameter (forcing) term. The Lorenz-96 model has been widely used in DA applications, mainly as a model to perform synthetic experiments [13], given its simplicity, and yet it exhibits chaotic behavior. Moreover, this model was also selected as it only includes one forcing parameter (F), and the nature of the dynamics depends heavily on it [142, 152, 13]. For instance, low values of F lead to asymptotically stable solutions, while $F = 8$ yields chaotic trajectories. In addition, Lorenz and Emanuel [152] and Sun, Miyoshi, and Richard [153] discuss that if the values of all states over a long period of time are considered, then it can be proved that their mean lies in the interval $[0, F]$, and their standard deviation in the interval $[0, {}^{F}\!/_{2}]$.

This thesis assumes $N = 20$ and observability of all Lorenz-96 states. Following the ideas presented by Erdal, Neuweiler, and Wollschläger [11], in the case of bias estimation in a model like Lorenz-96, if not all states are observable, the spatial correlation of the bias terms needs to be carefully chosen for the Kalman update actually to perform proper estimation of the bias. Thus, this situation is avoided simply by assuming that all states are observable.

4

Methods

This chapter bridges the theory presented in Chapter 3, and the implementation and experimentation. In particular, we first present an overview of details to be considered when developing a DA application. We then move to the current state of the code and its capabilities, and we finish the chapter by summarizing some parts of the experimental design followed to obtain the results.

4.1 The Data Assimilation Framework

This section will discuss some practical aspects and remarks relevant to implementing a *flexible* sequential (bias-aware) data assimilation framework. These considerations result from a combination of the theory and challenges encountered when implementing the code for this work. They will be described in a general manner, and at the end of the section, they will be summarized more explicitly, even though some are closely related or consequences of each other.

4.1.1 The Practical Setup

As a reference, we will consider that the DA framework acts on a nonlinear black-box system. For the illustrative purpose of this section, we assume that the dynamics of the state of the system can be reduced to a state-space in the form

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \mathbf{u}(t); \boldsymbol{\theta}), \tag{4.1}$$

where $\mathbf{x}$ is the state of the system and $\mathbf{u}$ is an external input. For instance, this formulation could result from a spatial discretization of a PDE.

This equation is then discretized and integrated in time through some numerical method, which we address as the *Integrator*, that takes the system from some initial state $\mathbf{x}(t_1)$ at the initial time t_1 to some new state $\mathbf{x}(t_2)$ at some specified final time t_2, using an integration time step of Δt (assumed constant for simplicity). This Integrator can be generally described by Algorithm 2, where **step** : $\mathbb{R}_+ \times \mathbb{R}^{n_\mathbf{x}} \to \mathbb{R}^{n_\mathbf{x}}$

is a single step of the numerical integration of f, and the notation $\mathbf{X}[:, i]$ represents the indexing of a tensor in the implementation using `NumPy` [154].

Algorithm 2 Explicit Integrator associated to (4.1).

```
Input: $t_1$                                         ▷ The initial time
Input: $\mathbf{x}(t_1)$                             ▷ The initial state
Input: $t_2$                                         ▷ The final time
Input: $\Delta t$                                    ▷ The integration time step
$n_{\mathbf{t}} \Leftarrow \lfloor \frac{t_2 - t_1}{\Delta t} \rfloor$
$\mathbf{X} \Leftarrow \mathbf{0} \in \mathbb{R}^{n_{\mathbf{x}} \times n_{\mathbf{t}}}$        ▷ Store states
$\mathbf{X}[:, 0] \Leftarrow \mathbf{x}(t_1)$
for $i = 0$ to $n_{\mathbf{t}} - 1$ do
    $\mathbf{X}[:, i+1] \Leftarrow$ step $(i\Delta t, \mathbf{X}[:, i])$
Output: $\mathbf{X}$                  ▷ The computed states between $t_1$ and $t_2$
```

For instance, in the case of a 4th order Runge Kutta (RK4) scheme, **step** would be given by

$$\mathbf{step}\,(t,\, \mathbf{x}) = \mathbf{x} + \Delta t \left(\frac{\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4}{6} \right), \tag{4.2}$$

where

$$\begin{aligned}
\mathbf{k}_1 &= f\,(t, \mathbf{x}, \mathbf{u}(t); \boldsymbol{\theta}),\\
\mathbf{k}_2 &= f\left(t + \tfrac{\Delta t}{2}, \mathbf{x} + \mathbf{k}_1 \tfrac{\Delta t}{2}, \mathbf{u}\left(t + \tfrac{\Delta t}{2}\right); \boldsymbol{\theta}\right),\\
\mathbf{k}_3 &= f\left(t + \tfrac{\Delta t}{2}, \mathbf{x} + \mathbf{k}_2 \tfrac{\Delta t}{2}, \mathbf{u}\left(t + \tfrac{\Delta t}{2}\right); \boldsymbol{\theta}\right),\\
\mathbf{k}_4 &= f\,(t + \Delta t, \mathbf{x} + \mathbf{k}_3 \Delta t, \mathbf{u}(t + \Delta t); \boldsymbol{\theta}),
\end{aligned}$$

and where f, $\mathbf{u}$ and $\boldsymbol{\theta}$ are as given in (4.1). Thus, in a DA context, the forward (background) model would run between assimilation steps, using the analysis step from the previous iteration as an initial condition. In this manner, the framework would have (in a base case) two time sequences: one for integration and one for assimilation (a subset). A simple diagram in Fig. 4.1 shows the difference between the assimilation and integration steps. Note also how the assimilation runs on the index k, while the model integration runs on a finer scale with a time step of Δt. In summary, this approach results in a discrete dynamical system like (3.17), where the forward model operator $\mathcal{M}_k$ represents the model integrator.

4.1.2 Considerations and Challenges

First, the framework should be able to perform automatic state augmentation. Whether the state is augmented with extra uncertain parameters, bias terms, or both, the size of the state vector can change depending on the setup desired by the

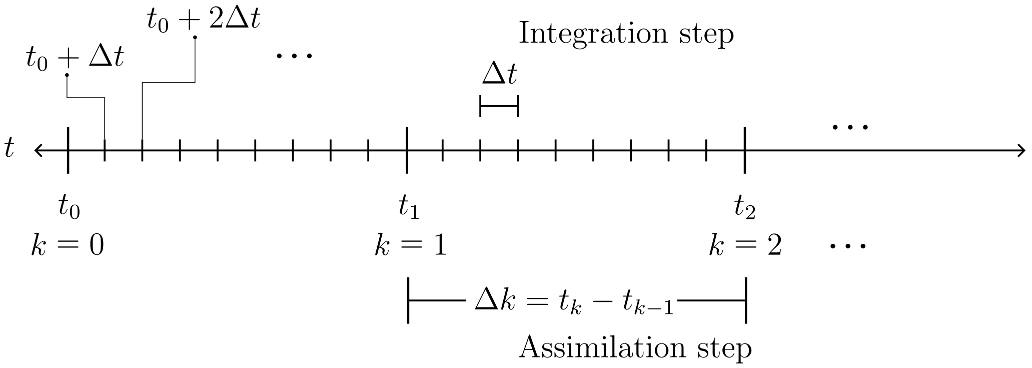


Figure 4.1: Integration vs. assimilation time steps.

end user. This has some consequences. First, all the objects related to the Kalman filter should be augmented as

$$\begin{aligned}
\mathbf{x}_k &\leftarrow \begin{bmatrix}\mathbf{x}_k^T & \mathbf{b}_k^T & \boldsymbol{\theta}^T\end{bmatrix}^T, \\
\mathcal{M}_k(\cdot) &\leftarrow \begin{bmatrix}\mathcal{M}_k(\cdot)^T & \mathcal{G}_k(\cdot)^T & \boldsymbol{\theta}_k^T\end{bmatrix}^T, \\
\mathbf{H}_k &\leftarrow \begin{bmatrix}\mathbf{H}_k & \mathbf{0} & \mathbf{0}\end{bmatrix}, \\
\mathbf{Q}_k &\leftarrow \operatorname{diag}\left(\mathbf{Q}_k, \mathbf{Q}_k^{\mathbf{b}}, \mathbf{Q}_k^{\boldsymbol{\theta}}\right).
\end{aligned}$$

Second, the estimated model parameters in the state vector $\hat{\boldsymbol{\theta}}_k$ should also update the parameters inside the forward model before the next prediction. Both of these remarks imply that the augmented state vector needs to be constructed at the beginning of each assimilation step, concatenating the real system's state, the current value of the bias state, and the current uncertain parameter values. Moreover, in theory, the forward model can be any dynamical system (black box), which ranges from an explicit ODE to any model product of an implementation in any software external to the DA framework. These models have their own "solver", responsible for discretizing time and solving the dynamics at every integration step (for instance, a Runge-Kutta method). Still, the dynamics for the parameters and the bias are inherently different. For example, in the case of a ColKF, the AR model is a discrete model that can easily be propagated, which means that the augmented system can have several "solvers" acting simultaneously to produce the augmented state forecast, and this also means that the models cannot be integrated one after the other, as possible coupling between them is possible, particularly, in the case of a feedback filter.

Regarding this last point, a feedback implementation of a filter translates to using the unbiased estimate of the state $\tilde{\mathbf{x}}_k$ (see Section 3.3.1) as input of the forward model. However, in a general case, implementing the bias back into the system can be done in several ways. We distinguish two main approaches: first, the formulation presented in Section 3.3.1, namely, equations (3.24a-3.24d) explicitly account for the bias in the discrete-time model, thus, after using an integrator. This approach is referred to as *discrete bias*. Second, motivated by the ideas presented by, for instance, Baek et al. [99] and Diab Montero [9], the bias term can be added as an extra forcing that affects the dynamics explicitly, namely $\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \mathbf{u}(t); \boldsymbol{\theta}) - \mathbf{H}^{\mathbf{b}}\mathbf{b}$. This implies that the bias correction acts before the discretization in time. This approach is referred to as *continuous bias*. Note that this last approach modifies the dynamic model explicitly, which may be a complication in the case of black-box forward models.

Moreover, the basic formulation of the EnKF, particularly (3.10), implies that a stochastic update of the forecast state is performed at every assimilation time instant k. However, in practice, these stochastic updates can be performed, for instance, at every integration step instead, namely, every Δt of the integrator (see Section 4.1.1). This situation may be desirable in cases where the ensemble spread is not large enough and the filter collapses too soon. In a flexible DA library, it should be up to the end user to decide which approach to take.

As a final remark, we should also consider how to estimate the state forecast covariance matrix in a *no-feedback* filter. In (3.24), in the case of no-feedback $\mathbf{F} = \mathbf{0}$ and it is clear that the state is corrected when observing the output $\mathbf{y}_k$. In a filter formulation, this translates to the correction in the analysis equation. However, the equation for the Kalman gain $\mathbf{K}_k$ in (3.7) uses $\mathbf{P}_k^f$. The literature on bias filters without feedback fails to specify if this covariance matrix should be computed using the state forecast $\mathbf{x}_k^f$ or if the bias-corrected state $\tilde{\mathbf{x}}_k$ can be used instead. One would expect the latter to estimate the background covariance better, although the formulation of bias-aware filters seems to favor the former approach [7]. This can be another possible option that the user can specify when working with a bias-aware filter without feedback.

Some additional considerations will now be described, though they were not dealt with in this project's scope. First, we assume the observation model is linear, which may not always be accurate. Evensen [41, Appendix A.1] addresses this issue for the case of the EnKF as long as the observation model is monotone in the state and not too nonlinear. Moreover, throughout the experiments carried out in this work, we also assume that the observation model is constant and observations come at a constant frequency. However, this can no longer be the case if, for instance, measurements are available at different times or rates. A possible approach to overcome this situation is performing asynchronous filtering, where the analysis is run at particular time steps and assimilating all the data that became available in one go [155].

4.2 Twin Experiments

A twin experiment is a controlled setup to evaluate and refine DA techniques. It involves creating one forecast model run, including realizations of the noise processes, that serves as the "truth" and then using it as observation measurements with a DA method. The observational data is generated from the truth model and then deliberately perturbed with noise to simulate real-world measurement errors. These perturbed observations are then assimilated into the guess model. The goal is to see how well the DA method can adjust the forecast model to recover the "truth" model [156]. Consequently, the performance of the data assimilation process can be assessed by comparing the output of the assimilated guess model with the truth model. In summary, twin experiments are essential for validating data assimilation techniques before applying them to actual observational data in real-world scenarios.

The procedure to conduct a twin experiment can be broken down into four general steps: (**1**) perform a run of the prediction model ("truth"), (**2**) sample synthetic observations (with added noise) from this simulation, (**3**) assimilate the sampled observations into a different run of the prediction model with varying realizations of noise (e.g. with different Random Number Generator (RNG) seed) or perturbed initial, forcing or boundary conditions, and (**4**) evaluate performance. Moreover, twin experiments can also be classified as identical and non-identical twin experiments. The former refers to a setup in which the synthetic true observations are generated by the same model to which the DA is applied, only perturbing the initial/boundary conditions or forcing terms [157]. The latter refers to the case of having different model types between the one that generates the observations and the one used in the DA loop [156], since in some cases studying how robust the methods are is desired.

Moreover, the procedure for estimating bias in a twin experiment is similar. The only difference is that the model is perturbed with the predefined bias before the synthetic observations are generated. For instance, assume we have a discrete model that evolves according to (3.17a) and that we have a discrete bias function $\mathbf{b}_k$ that we would like to estimate (or account for) using the bias-aware method. Then, depending on whether it is a feedback ($\mathbf{F} \equiv \mathbf{I}_{n_\mathbf{x}}$) or a no feedback ($\mathbf{F} \equiv \mathbf{0}$) filter, the model used to generate the synthetic observations is

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}) + \mathbf{F}\mathbf{H}^{\mathbf{b}}\mathbf{b}_k, \\ \mathbf{y}_k &= \mathbf{H}_k(\mathbf{x}_k + (\mathbf{I}_{n_\mathbf{x}} - \mathbf{F})\mathbf{H}^{\mathbf{b}}\mathbf{b}_k)\end{aligned}$$

whereas the forward model used within the DA experiment would be

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}), \\ \mathbf{y}_k &= \mathbf{H}_k\mathbf{x}_k\end{aligned}$$

so that the effect of $\mathbf{b}_k$ is recovered using the DA method.

4.3 Implementation Details

The code was developed in `Python 3.12`, and it is publicly accessible via the `GitHub` repository `https://github.com/Daples/master-thesis`.

One of the main advantages of the current implementation is its versatility in supporting multiple filtering algorithms. Currently, it includes the KF, the EnKF (with optional covariance localization), and the ColKF with an underlying EnKF for the augmented system, and with a flag for feedback or no feedback. Thus, the implementation can be customized to specific application needs, and the architecture allows for more convenient implementation of other filtering techniques.

The implementation is designed to work smoothly with any nonlinear ODE model, provided the observation model is linear (see Section 3.2.2). Moreover, the architecture enables any solver for the forward model to be connected without significant reconfiguration. Another advantage is the flexibility to modify this model in various ways, such as allowing inputs $\mathbf{u}$ and adjusting continuous offsets s^c, discrete offsets s^d, and observation offsets s^o. More explicitly, the code extends any system like (4.1) to the modified form

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), \mathbf{u}(t); \boldsymbol{\theta}) + s^c(t), \tag{4.3a}$$

$$\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k, \mathbf{u}_k; \boldsymbol{\theta}) + s^d(k) + \mathbf{w}_k, \tag{4.3b}$$

$$\mathbf{y}_k = \mathbf{H}_k\left(\mathbf{x}_k + s^o(k)\right) + \mathbf{v}_k, \tag{4.3c}$$

where $s^c : \mathbb{R}_+ \to \mathbb{R}^{n_\mathbf{x}}$ represent a continuous offset to the state-space formulation of the system, $s^d : \mathbb{Z}_+ \to \mathbb{R}^{n_\mathbf{x}}$ is an offset after the integrator performs discretization in time, and $s^o : \mathbb{Z}_+ \to \mathbb{R}^{n_\mathbf{x}}$ is an offset that is applied to the state before being observed. In our case, for bias-aware filters, these offsets represent different forms of bias, and, in particular, s^c and s^d correspond to *feedback* forms of bias, and s^o to a *no-feedback* form.

Note also that system (4.3) includes the ODE and the discretization of the dynamics of the state vector $\mathbf{x}$, i.e. $\mathcal{M}_k(\cdot)$ is (4.1) after performing time discretization, following the concepts presented in Section 4.1.1. Both equations are shown only for explanatory purposes. In this manner, the model can be easily modified to meet our needs. For instance, the formulation of the ColKF presented in (3.24) can be easily achieved by adding the dynamics for the bias $\mathbf{b}_k$, and setting $s^c(t) = 0$, $s^d(k) = -\mathbf{F}\mathbf{H}^{\mathbf{b}}\mathbf{b}_k$, and $s^o(k) = (\mathbf{F} - \mathbf{I}_{n_{\mathbf{x}}})\mathbf{H}^{\mathbf{b}}\mathbf{b}_k$. Whereas, if the bias is desired to affect the ODE directly, then $s^c(t) = -\mathbf{F}\mathbf{H}^{\mathbf{b}}\mathbf{b}_t$ and $s^d(k) = 0$ instead.

In addition, the code supports estimating any number of parameters specified in the model. The model can be built explicitly with parameter instances, each with a flag that tells the framework to estimate it or not. The state is automatically augmented accordingly. Furthermore, the implementation offers different ways to manage uncertainty within the model, parameters, and bias. In particular, for every dynamical model, the user can specify whether stochastic updates should be performed and with which frequency. This can be done at either the integration step, the assimilation step, or none. Finally, the framework also includes functionality for standardized results gathering and plot generation for single or multiple filtering runs.

5

Results

This chapter presents the general results of this project. All numerical integration was performed using a RK4 method. In addition, the RNG is seeded, meaning that all experiments are reproducible under the same conditions. Finally, unless otherwise specified, we assume that the EnKF by default performs stochastic updates to the state only at each assimilation step.

5.1 Twin Experiments

The procedure to follow to perform a twin experiment was explained in a general setting in Section 4.2. This section presents the results for the twin experiments run on the LHO and Lorenz-96 models. These experiments were conducted for two main reasons: first, they help validate the implementation developed, and second, they also serve as a reference for later experiments.

5.1.1 Linear Harmonic Oscillator

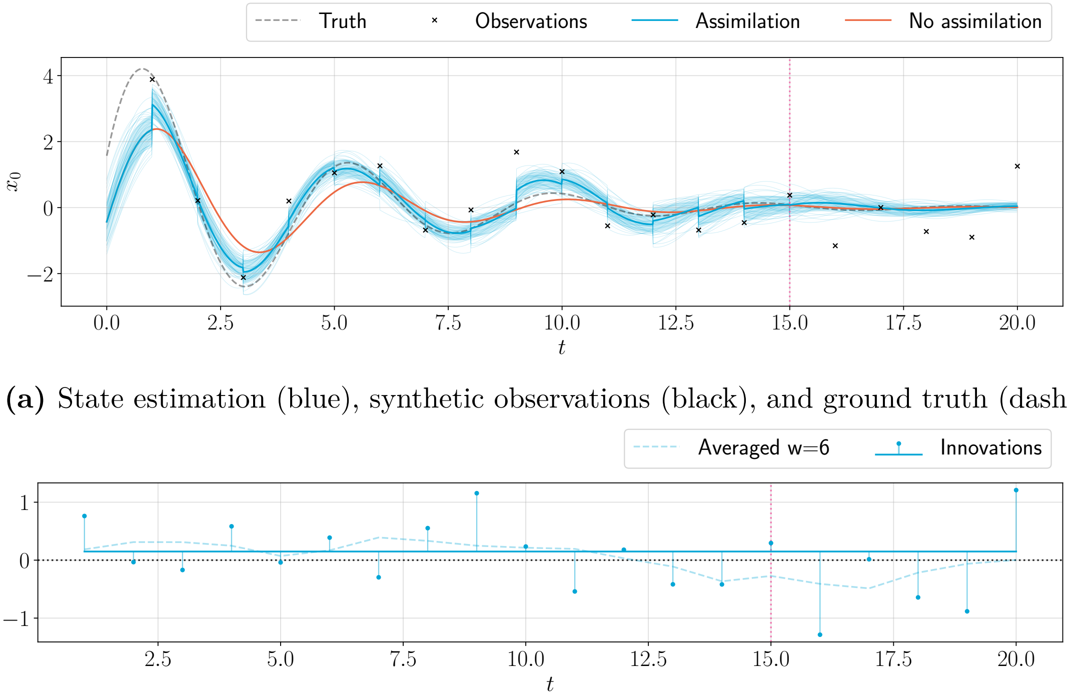
We first tested the EnKF for the LHO model by running a twin experiment. To this end, we used the experimental parameters presented in Table 5.1, where θ_1 and θ_2 are as specified in (3.33), T is the final simulation time, K is the final assimilation time, Δt is the integration time step, and Δk is the assimilation time step. Moreover, note that $\mathbf{x}_0$ refers to the initial state used by the "true" run, while $\boldsymbol{\mu}_0$ is the expected value of that initial state, used to initialize the EnKF. Moreover, note that this application of the EnKF does not require covariance localization. The reason for this is two-fold: first, as the state vector is only comprised of two variables, these are expected to be correlated (also given the explicit formulation of the LHO). Second, covariance localization is applied only when the state variables have a spatial connotation, which this model does not have.

For illustration purposes, assume that the model is described in the international system of units (time in seconds). This setup means that this experiment runs on the time interval from 0 to 20 seconds, integrating the model every 0.01 seconds and performing assimilation every 1 second until 15 seconds. In the last 5 seconds, the model is run without Kalman update. Finally, note we assume that all states are observable $\mathbf{y}_k = \mathbf{x}_k$.

Table 5.1: Parameters for EnKF twin experiment with LHO model.

$n_{\mathbf{x}}$	θ_1	θ_2	$\mathbf{H}_k$	$\boldsymbol{\mu}_0 \equiv \mathbf{x}_0$	$\mathbf{P}_0$	$\mathbf{Q}_k$	$\mathbf{R}_k$	T	K	Δt	Δk	N
2	-2	-0.5	$\mathbf{I}_2$	$[\pi/2\ \ 6.5]^T$	$0.5\mathbf{I}_2$	$0.1\mathbf{I}_2$	$0.3\mathbf{I}_2$	20	15	0.01	1	40

We then perturb the mean of the initial state $\boldsymbol{\mu}_0$ by adding $-2 \cdot \mathbf{1}_2$ and apply the EnKF. Fig. 5.1 shows the assimilation results for x_0 in the twin experiment. In particular, Fig. 5.1a presents the estimated state, ground truth, and synthetic observations (an additional run without DA was also added for reference). In addition, Fig. 5.1b shows the innovations at each analysis step (stems), the baseline (horizontal continuous line) corresponds to the average across all innovations in the interval $[0, K]$, and the dashed line corresponds to a moving average with the specified window. Note that both figures include a vertical dashed pink line indicating the end of the assimilation window, i.e., only the forward model is run afterward. This is done to showcase the model's prediction capabilities without assimilation. The same scheme is used for the remaining results.



(a) State estimation (blue), synthetic observations (black), and ground truth (dashed).

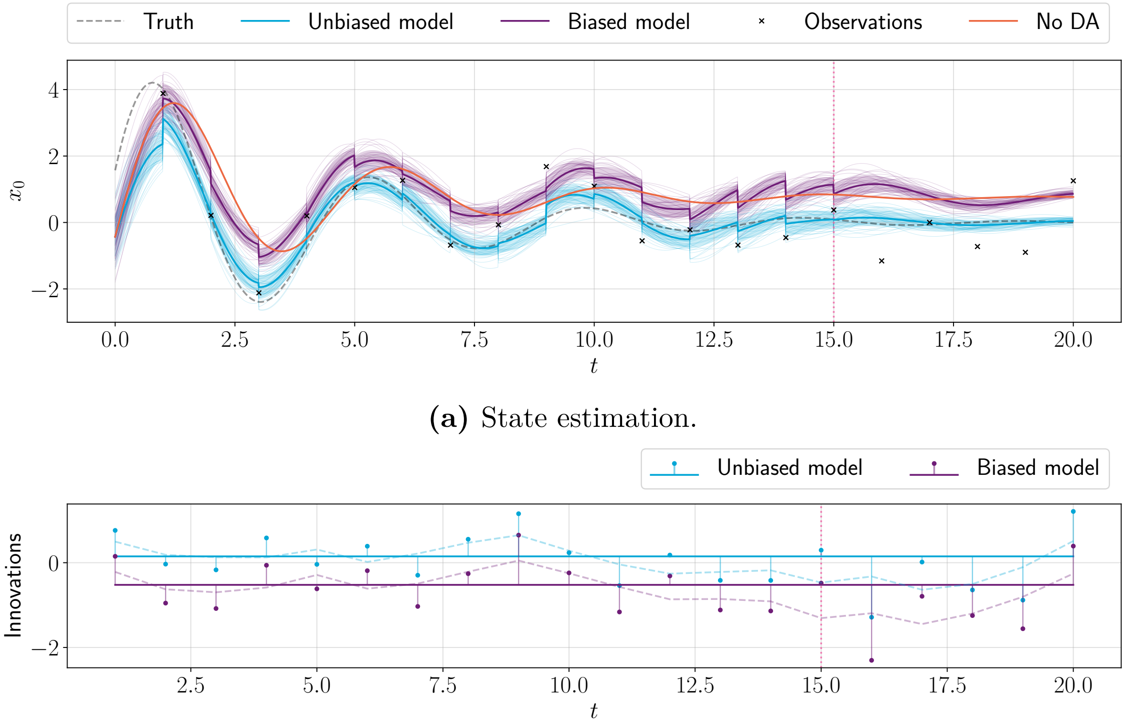
(b) Innovation process.

Figure 5.1: Twin experiment results with EnKF for x_0 in the LHO model.

Now, suppose we introduce a simple bias $[1\ 1]^T$ into the forward model dynamics as

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} 0 & 1 \\ -2 & -0.5 \end{pmatrix} \mathbf{x}(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$
$$\mathbf{y}(t) = \mathbf{x}(t),$$

and use the same synthetic observations from the last twin experiment to correct the misspecified model. In that case, the EnKF struggles to estimate the states between assimilation steps, as presented in Fig. 5.2a, again, for the state x_0. Note how this perturbation in the dynamics makes the state consistently higher than the ground truth (and even the original run) and how the Kalman update always performs a correction in the same direction. In addition, from 5.2b, it can be observed that the average innovations for the biased run are not particularly close to its anticipated value of zero. In this case, we explicitly added model bias to the LHO; however, both of these outcomes are known to be strong evidence of model bias in a general DA experiment [4, 101].



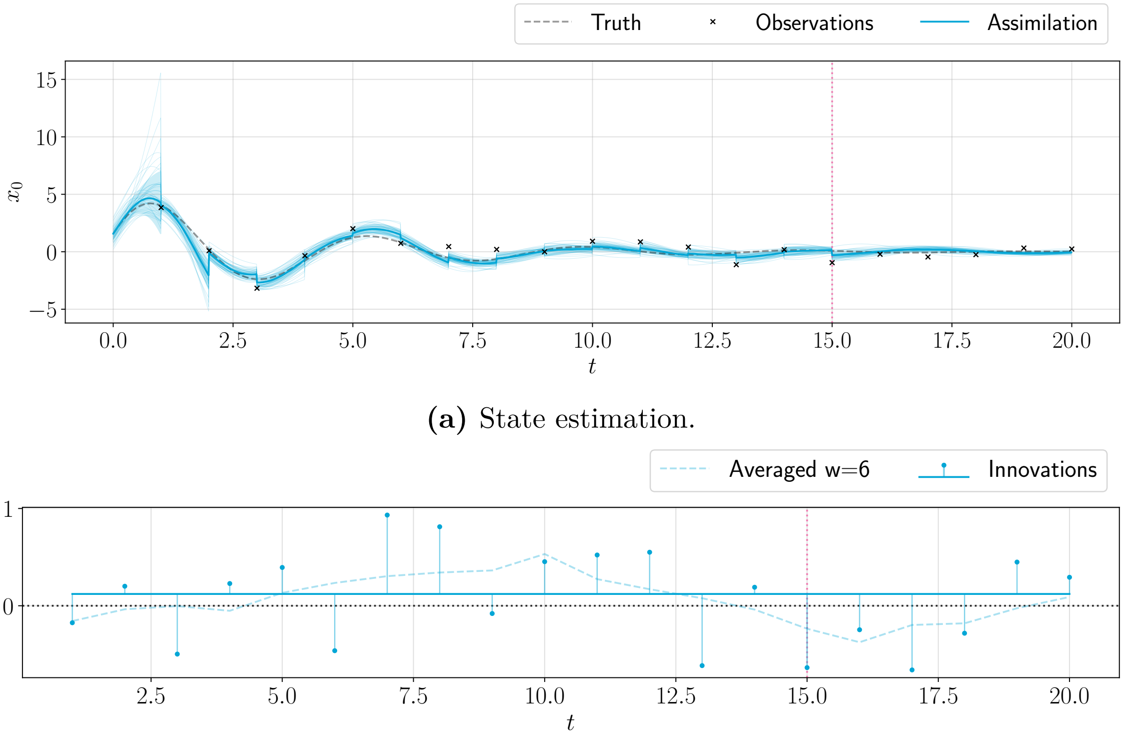
(a) State estimation.

(b) Innovation process.

Figure 5.2: EnKF results for x_0 with bias (purple) against the original model (blue) in the LHO model.

Furthermore, a simple parameter estimation experiment was performed for the linear model. The same setup presented in Table 5.1 is used for this experiment, and the same set of observations is assimilated. However, the parameters are initialized as $\theta_1 = -3$ and $\theta_2 = 0.5$, and included in the augmented state vector following the methodology presented in Section 3.2.4, with constant forecast error standard deviation of $\sigma_k^{\theta_1} = \sigma_k^{\theta_2} = 0.7$.

The state estimation results of this experiment are presented in Fig. 5.3. Note how the ensemble quickly spreads out with respect to the true model, given that each ensemble is initialized with a different set of parameters, but as new observations are assimilated, the deviation of the ensemble from the expected behavior decreases, showing that the DA cycle is successfully correcting the model parameters on each ensemble. This can be better visualized in Fig. 5.4, where the estimated parameters through time are presented. This figure shows that both parameters reach a close estimated value to the truth (given in the dotted horizontal lines).



(a) State estimation.

(b) Innovation processes.

Figure 5.3: Estimation filtering results for x_0 in parameter estimation experiment for the LHO model.

5.1.2 Lorenz-96

Similarly, a twin experiment was also constructed for the Lorenz-96 model. Table 5.2 presents the parameters used for this experiment. These parameters are motivated

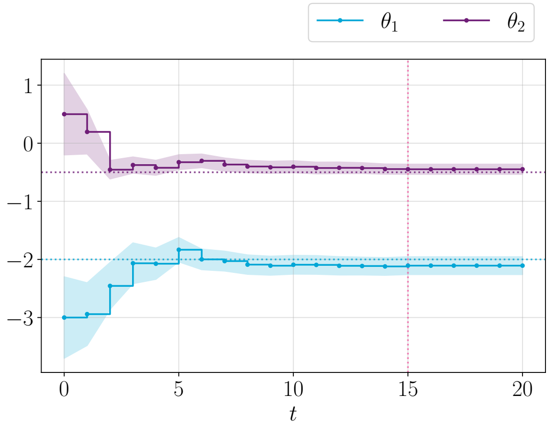


Figure 5.4: Estimated parameters through time for the LHO model. The dotted line represents the true value.

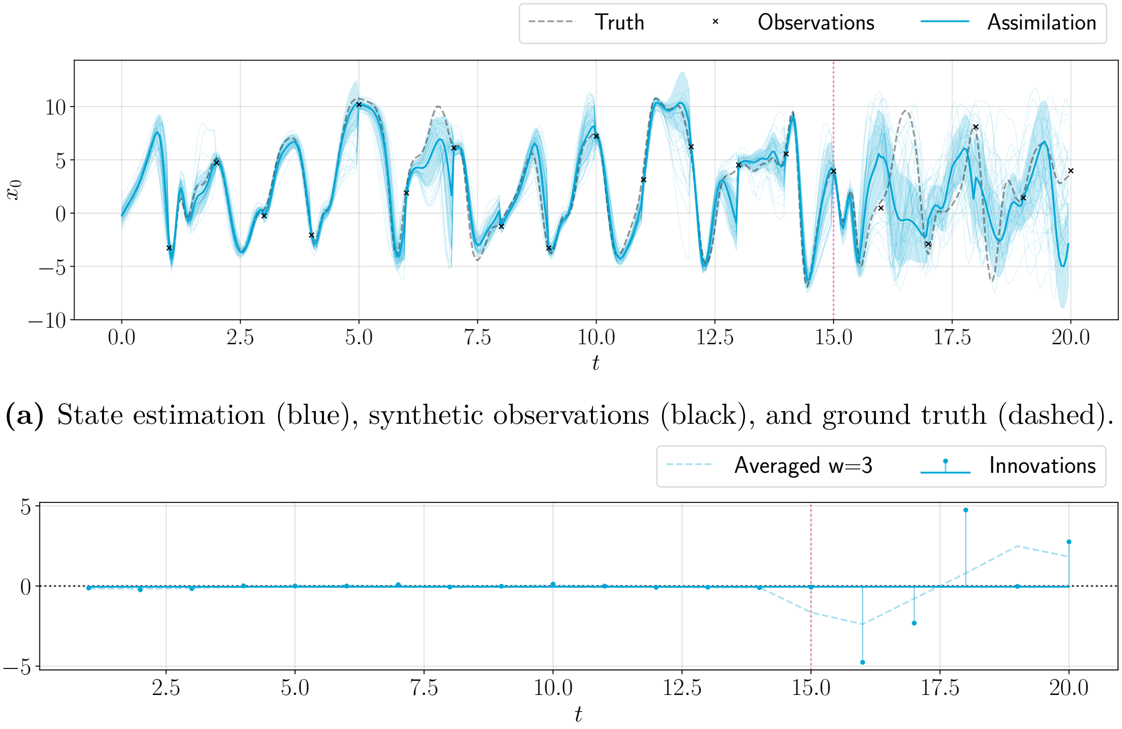
Table 5.2: Parameters for EnKF twin experiment with Lorenz-96 model.

$n_{\mathbf{x}}$	F	$\mathbf{H}_k$	$\boldsymbol{\mu}_0 \equiv \mathbf{x}_0$	$\mathbf{P}_0$	$\mathbf{Q}_k$	$\mathbf{R}_k$	T	K	Δt	Δk	N
20	8	$\mathbf{I}_{n_{\mathbf{x}}}$	$G(0, \mathbf{I}_{n_{\mathbf{x}}})$	$0.5\mathbf{I}_{n_{\mathbf{x}}}$	$0.5\mathbf{I}_{n_{\mathbf{x}}}$	$0.3\mathbf{I}_{n_{\mathbf{x}}}$	20	15	0.05	1	40

by the ones used by Baek et al. [99] for their model setup and experiments. Note that these parameters cause the system to exhibit chaotic behavior [13], specifically given the value of F. Moreover, the expression for $\boldsymbol{\mu}_0$ means that the expected value of the initial state was just drawn from a standard Gaussian distribution. Finally, this initial experiment does not use covariance localization since the validation of the EnKF's performance is the main goal of this experiment. The impact of covariance localization will be addressed jointly with the estimation of parameters later in this section.

The assimilation results for x_0 are presented in Fig. 5.5, where the state estimation results and innovation process are presented. In general, good results are obtained. However, it can also be observed that the state forecast between analysis times does not always match the truth. The reason for this is twofold: first, chaotic systems are known to be highly sensitive to initial conditions [151, 142, 13], and since the analysis is not always a perfect match of the truth state, the dynamics cause the system to deviate quickly. Second, the assimilation time step Δk seems too high for the current experimental setup. In addition, the model's predictive capability, in this case, is none. When the assimilation ends at $t = 15$, all ensemble members succumb to the chaotic behavior of the system, deviating significantly from the truth. This is also the reason why the innovation process increases heavily in scale in the prediction stage.

Furthermore, we also add a bias to the dynamics of the Lorenz-96 of the prediction



(a) State estimation (blue), synthetic observations (black), and ground truth (dashed).

(b) Innovation process.

Figure 5.5: EnKF results for x_0 on the Lorenz-96 model.

model, following the pattern also from the LHO experiments. Nevertheless, the bias added in this case is somewhat different. We follow the perturbation model showcased by Baek et al. [99], where the bias varies in space but stays constant over time. Formally, each state follows the perturbed equation

$$\dot{x}_j(t) = x_{j-1}(t)\left(x_{j+1}(t) - x_{j-2}(t)\right) - x_j(t) + F + \beta_j,$$

where

$$\beta_j = \frac{F}{2} \sin\left(\frac{2\pi(j-1)}{N}\right).$$

This form of bias has the shape presented in Fig. 5.6. This means that states corresponding to $j = 5$ and $j = 15$ are the ones affected the most by the bias. The results of this experiment are presented in Fig. 5.7. It can be observed that even though the bias does not directly affect x_0, the biased model shows different results inside the assimilation window. In this case, it is not as straightforward as in the LHO model, but there are still significant corrections that Kalman update needs to account for (see, e.g., around $t = 5$ and $t = 12$).

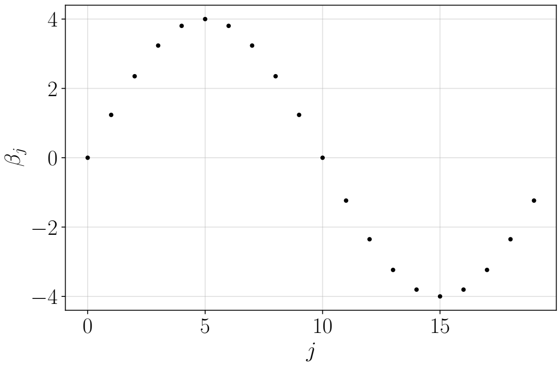
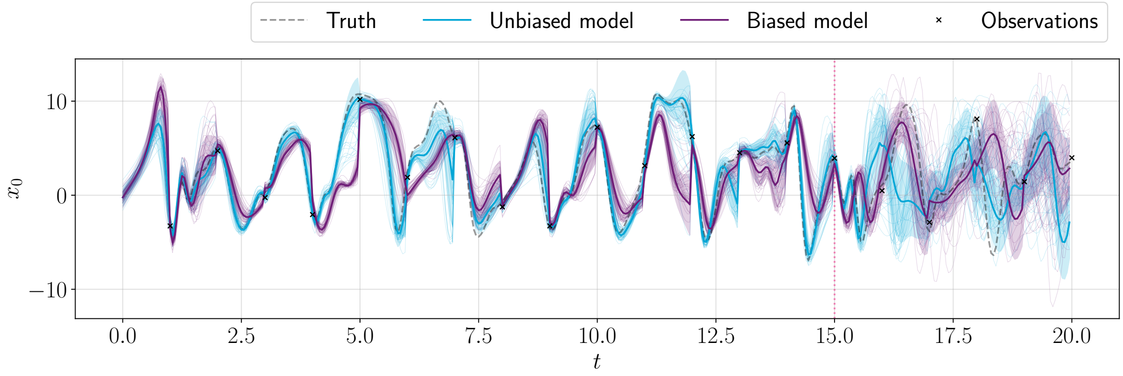
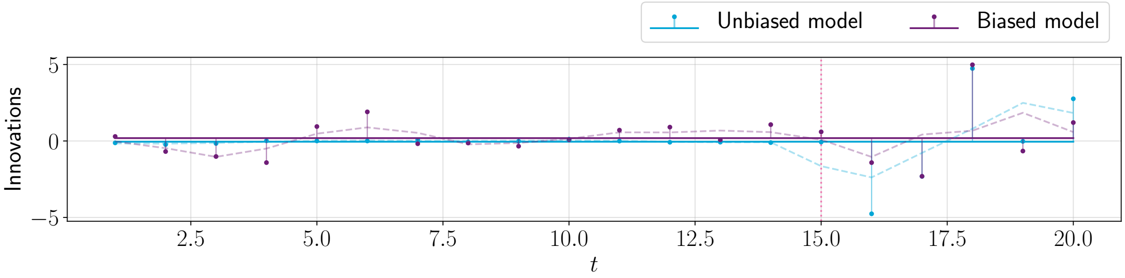


Figure 5.6: Spatial variation of the bias β_j.



(a) State estimation.

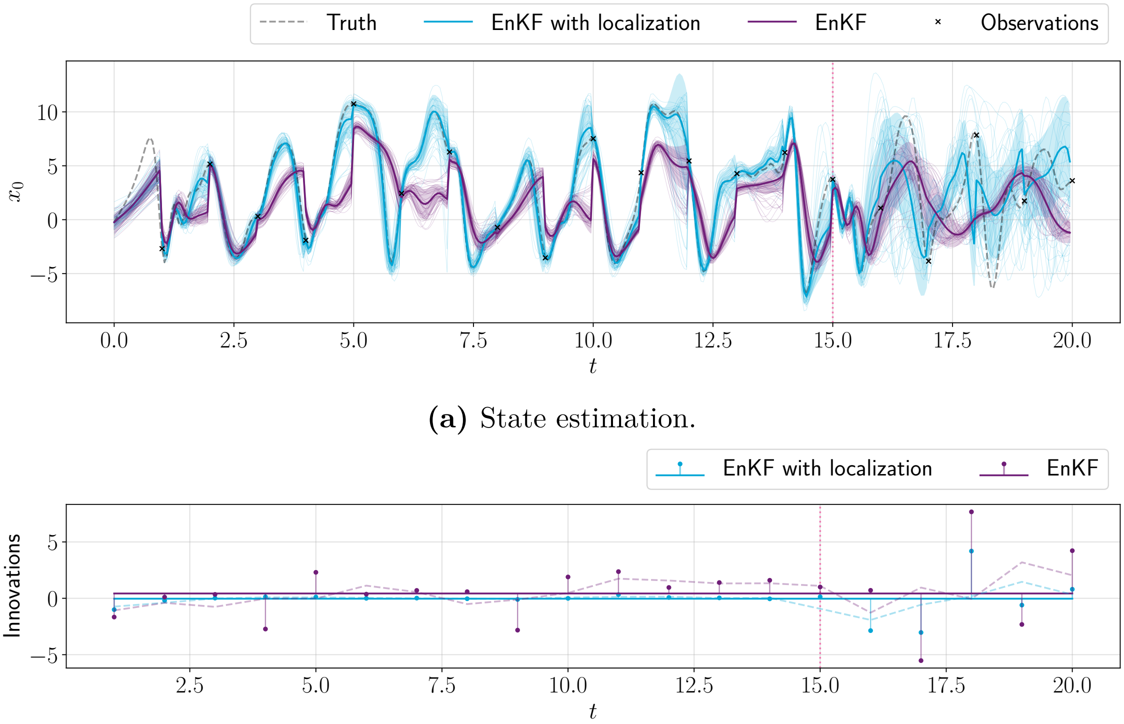


(b) Innovation process.

Figure 5.7: EnKF results for x_0 with bias (purple) against the original model (blue) in the Lorenz-96 model.

In the case of the Lorenz-96 model, we also performed a simple parameter estimation experiment to estimate the forcing F. Moreover, this experiment was carried out with and without localization to illustrate its impact on the assimilation with the Lorenz-96 model. We now present these results.

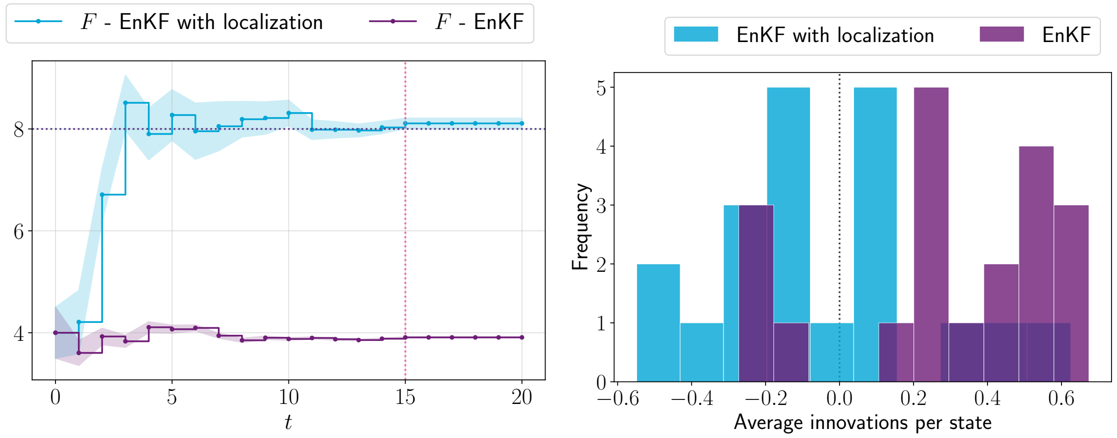
For this experiment, we initialized the parameter at $F_0 = 4$ with a constant standard deviation $\sigma_k^F = 0.5$, which represents the parameter uncertainty. The ensemble size was purposely reduced to $N = 30$ to enhance the effects caused by small ensemble sizes and showcase the relevance of covariance localization. We set the localization radius $r = 3$. The results are presented in Figs. 5.8 and 5.9.



(a) State estimation.

(b) Innovation processes.

Figure 5.8: EnKF results for x_0 with (blue) and without covariance localization (purple) for Lorenz-96 model.



(a) Estimated parameters.

(b) Average innovations for all states.

Figure 5.9: Results for Lorenz-96 experiments with parameter estimation, with and without localization.

The filtering and parameter estimation results are significantly better when using localization. Note that the state estimates in Fig. 5.8a using the EnKF without localization (purple line) cannot replicate the same nature as the truth trajectory (gray dashed line). This is closely linked to the forcing F estimation presented in Fig. 5.9a since it is clear that the estimated forcing without localization cannot reach the reference value. Therefore, the prediction model being used in the assimilation is inherently different from the one used to generate the synthetic observations.

Finally, for this example, an illustration of the covariance localization at the first assimilation step is presented in Fig. 5.10. The first panel (a) shows the covariance of the forecast states before localization and the resulting one after on the second panel (b). The associated mask, generated with Gaussian decay (see Section 3.2.3), is presented in the bottom panel (c). As expected, only the local covariance is preserved after applying the localization mask.

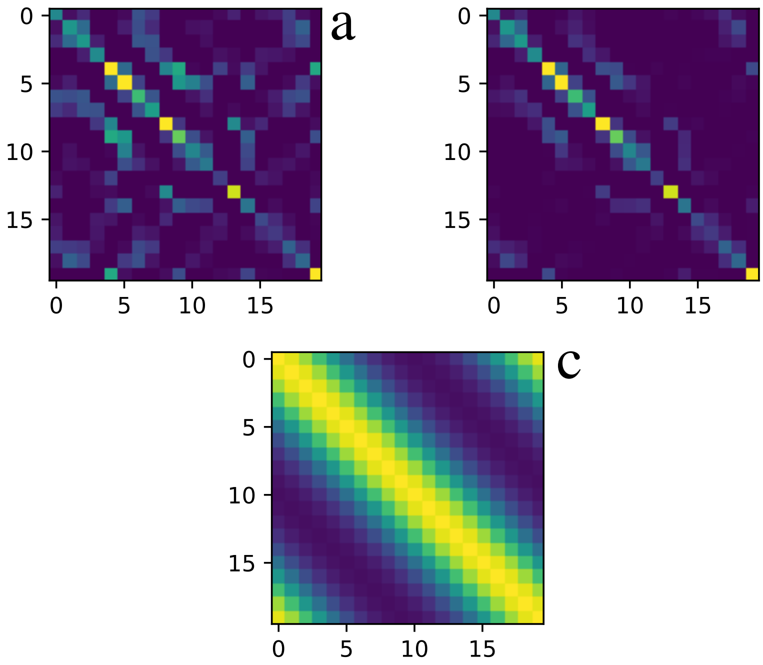


Figure 5.10: Covariance of the forecast states, before (a) and after (b) localization, using the mask presented in (c).

In summary, this last experiment showcased the importance of covariance localization when performing DA and parameter estimation using the EnKF. From this point, all DA experiments using the Lorenz-96 are performed using covariance localization.

5.2 Bias Estimation using ColKF

This section presents the results of applying the ColKF to estimate constant discrete bias while showcasing the different approaches to update the model and bias stochastically. In particular, we run two different experiments for each model:

- In the first one, the stochastic update of the forward model and bias (AR model) happens at every analysis step (every Δk). We refer to this experiment as *Analysis-Analysis update*.
- In the second one, both the forward model and the AR process are stochastically updated at every integration step (every Δt). We refer to this experiment as *Integration-Integration update*.

The two remaining experiment combinations, namely *Integration-Analysis update* and *Analysis-Integration update*, were moved to Appendix A.

All experiments in this section follow the twin experiment for bias estimation methodology described in Section 4.2. Moreover, we made all the biases "reachable" in these experiments. In other words, if the model used to generate the synthetic observations was perturbed using discrete bias, then the AR process in the ColKF is also set to estimate discrete bias, as $s^d(k)$ in (4.3b). This allows us to establish a reference value for the estimated bias state. Furthermore, a *feedback* formulation, that is $\mathbf{F} \equiv \mathbf{I}_{n_\mathbf{x}}$, was used for all the experiments in this section. Finally, we assume in both models that the bias is given by $\mathbf{b}_k = b^{\text{true}} \cdot \mathbf{1}_{n_\mathbf{x}}$, $\forall k \in \mathbb{Z}_+$, namely, it is constant over time and space, and with magnitude b^{true}. Note also that this implies that $n_\mathbf{x} = n_\mathbf{b}$ and $\mathbf{H}^\mathbf{b} = \mathbf{I}_{n_\mathbf{x}}$.

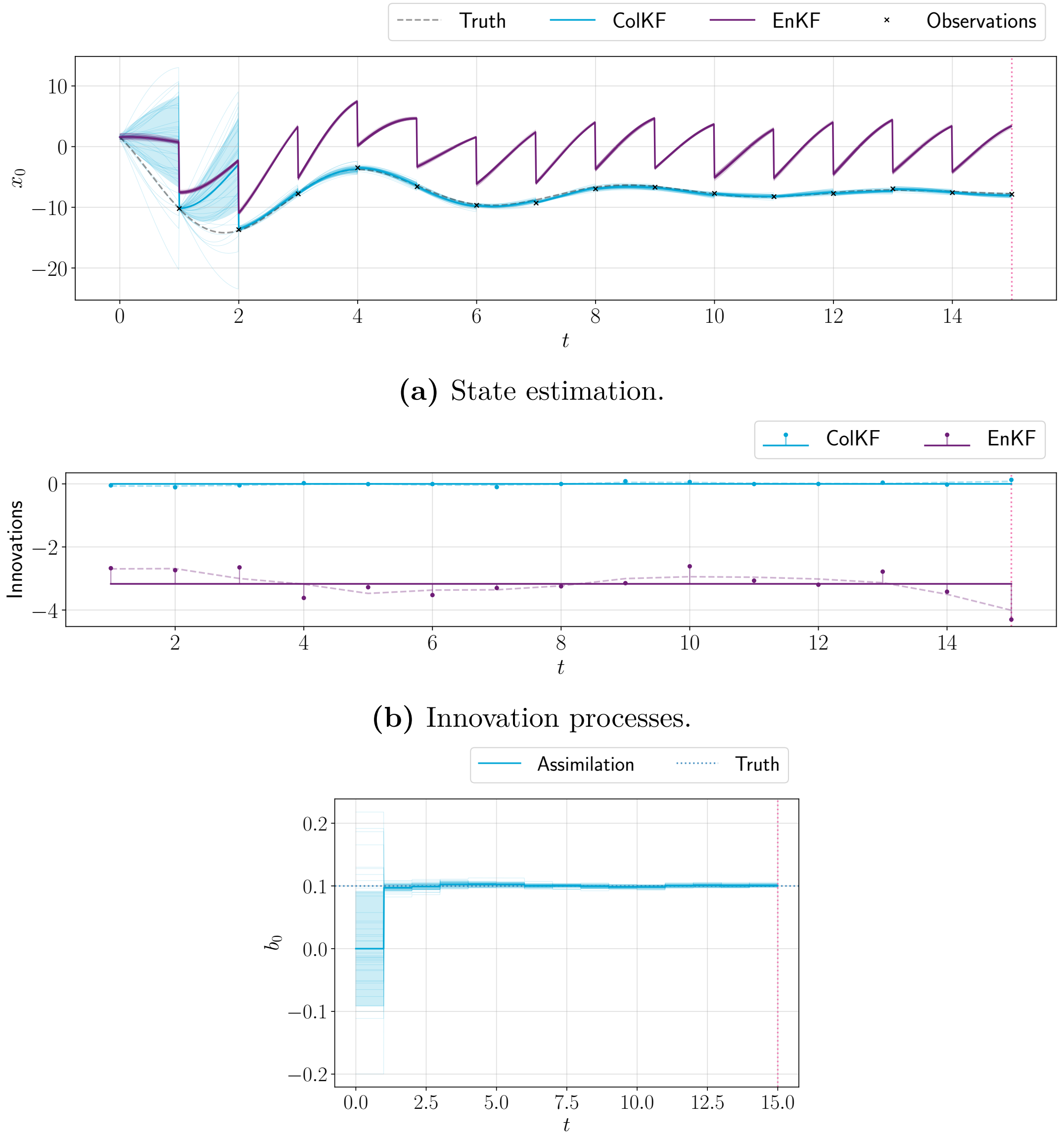
5.2.1 Linear Harmonic Oscillator

For both the *Analysis-Analysis update* and *Integration-Integration update* experiments with the LHO, we used the parameters presented in Table 5.3.

Table 5.3: Parameters for ColKF bias estimation twin experiments with the LHO.

$n_\mathbf{x}$	θ_1	θ_2	$\mathbf{H}_k$	$\boldsymbol{\mu}_0 \equiv \mathbf{x}_0$	$\mathbf{P}_0$	$\mathbf{Q}_k$	$\mathbf{R}_k$	T	K	Δt	Δk	N	$\mathbf{A}_k$	b^{true}	$\mathbb{E}[\mathbf{b}_0]$	$\mathbb{V}[\mathbf{b}_0]$	$\mathbf{Q}_k^\mathbf{b}$
2	-2	-0.5	$\mathbf{I}_2$	$[\pi/2 \;\; 0.5]^T$	$0.1\mathbf{I}_2$	$0.1\mathbf{I}_2$	$0.05\mathbf{I}_2$	15	15	0.01	1	40	$\mathbf{I}_2$	-0.1	$\mathbf{0}$	$0.01\mathbf{I}_2$	$10^{-6}\mathbf{I}_2$

The results for the *Analysis-Analysis update* experiment are presented in Fig. 5.11. From the state estimation plot, it is clear that the EnKF cannot properly estimate the observations from the biased model using only the unbiased forecast model, while the ColKF can properly account for the bias after two analysis steps. In addition, it is clear that the ensemble spread is quite low in the EnKF run, whereas the ColKF run shows a large spread before the bias is properly learned. This is expected since every bias component is initialized randomly for all the ensemble members, and thus, the trajectories start with different forcings; also, since the stochastic updates only happen during the analysis, the uncertainty does not grow significantly in between assimilation times. Moreover, from Fig. 5.11c, it is clear that the ensemble of bias for x_0 quickly converges to the truth value.



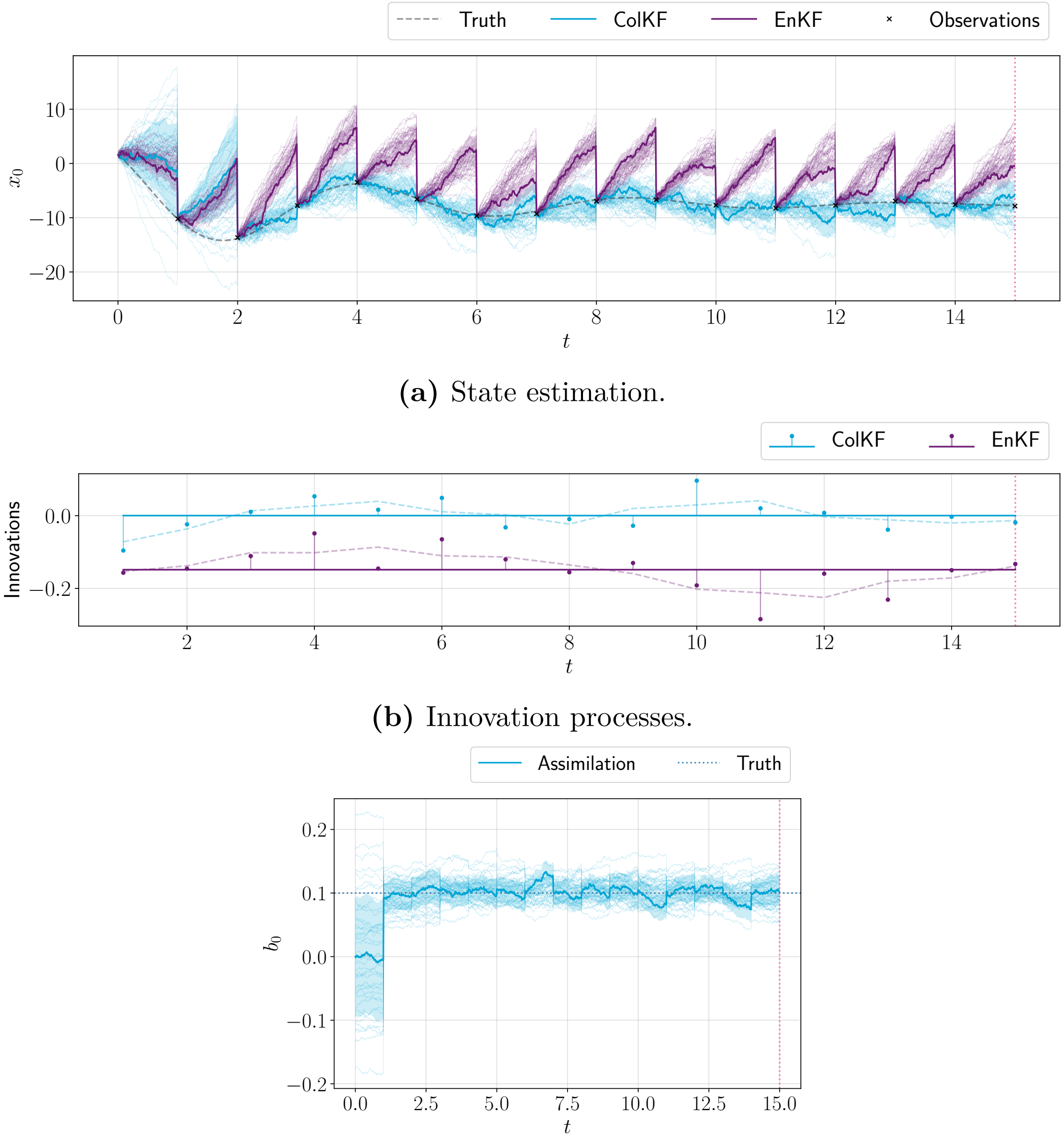
(a) State estimation.

(b) Innovation processes.

(c) Estimated AR process (proxy for bias), corresponding to x_0.

Figure 5.11: Results for x_0 with the EnKF (purple) and with the ColKF (blue) for the LHO model with added constant bias in *Analysis-Analysis update* experiment.

Furthermore, for the *Integration-Integration update* experiment, results are presented in Fig. 5.12. In this case, the uncertainty clearly grows in between analysis steps, as both the model and AR process are being stochastically updated. Moreover, the EnKF performs arguably better in this experiment than the *Analysis-Analysis update*. A quick glimpse of Fig. 5.12b and comparing it to 5.11b shows that the average of the innovations is lower. However, the state estimation shows that this improvement is only in the analysis since the same “saw tooth” kind of behavior is still present. In the case of the ColKF, the filter shows good state estimation results and proper estimation of bias (see Fig. 5.12c). It is clear that the AR process, in this case, cannot converge exactly to the desired value since the stochastic updates happen too often, but it stays around the true value.



(a) State estimation.

(b) Innovation processes.

(c) Estimated AR process (proxy for bias), corresponding to x_0.

Figure 5.12: Results for x_0 with the EnKF (purple) and with the ColKF (blue) for the LHO with added constant bias in *Integration-Integration update* experiment.

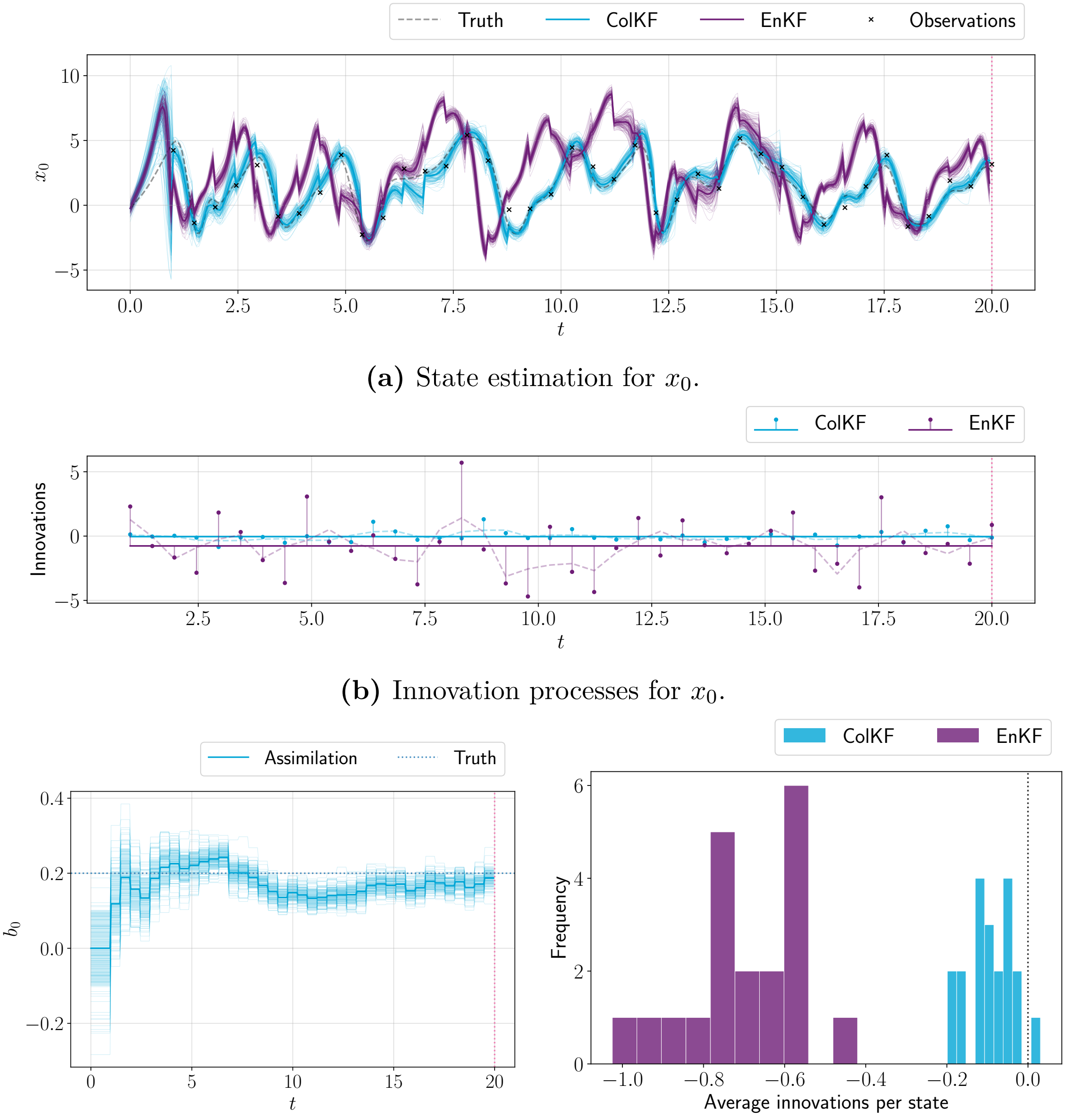
5.2.2 Lorenz-96

For both the *Analysis-Analysis update* and *Integration-Integration update* experiments with the Lorenz-96, we used the parameters presented in Table 5.4. Moreover, both the EnKF and ColKF used covariance localization with a localization radius of $r = 3$, the same as presented in Section 5.1.2.

Table 5.4: Parameters for ColKF bias estimation twin experiment with the Lorenz-96 model.

$n_{\mathbf{x}}$	F	$\mathbf{H}_k$	$\boldsymbol{\mu}_0 \equiv \mathbf{x}_0$	$\mathbf{P}_0$	$\mathbf{Q}_k$	$\mathbf{R}_k$	T	K	Δt	Δk	N	$\mathbf{A}_k$	b^{true}	$\mathbb{E}[\mathbf{b}_0]$	$\mathbb{V}[\mathbf{b}_0]$	$\mathbf{Q}_k^{\mathbf{b}}$
20	8	$\mathbf{I}_{n_{\mathbf{x}}}$	$G(0, \mathbf{I}_{n_{\mathbf{x}}})$	$0.1\mathbf{I}_{n_{\mathbf{x}}}$	$0.05\mathbf{I}_{n_{\mathbf{x}}}$	$0.3\mathbf{I}_{n_{\mathbf{x}}}$	15	15	0.05	0.5	80	$\mathbf{I}_{n_{\mathbf{x}}}$	-0.2	$\mathbf{0}$	$0.01\mathbf{I}_{n_{\mathbf{x}}}$	$10^{-4}\mathbf{I}_{n_{\mathbf{x}}}$

The results for the *Analysis-Analysis update* experiment are presented in Fig. 5.13. Similar to the LHO experiments, the EnKF underperforms in the presence of model bias. This is better summarized in Fig. 5.13d, where the frequency histogram of average innovations for the EnKF (purple) does not include values as close to zero as the results obtained with the ColKF (blue). It is clear how well the ColKF performs in this case. Moreover, in contrast to the LHO *Analysis-Analysis update* experiment, the ensemble spread is higher in the Lorenz-96 experiment. This is due to the high nonlinearity inherent in the system. Finally, note that the estimated bias for x_0 does not exactly converge to the true value, even when stochastic updates are happening only at the analysis times. In this case, given the high dimension of the state (in comparison to the LHO), and given that the dynamics of a state depend nonlinearly on the nearby states, it is possible that the effects are collectively mitigated through corrections on the nearby states and thus, the bias is not precisely estimated. Nevertheless, these results show that the ColKF can correctly account for bias in this experiment. Moreover, the *Integration-Integration update* experiment results are showcased in Fig. 5.14. Again, a quick look into the frequency histogram of average innovations (in Fig. 5.14d) shows that the performance of the EnKF actually improves when the stochastic updates happened at every integration step when comparing with its counterpart in Fig. 5.13d. Furthermore, the estimated bias is not as stable around the true value in comparison with the equivalent experiment in the LHO.



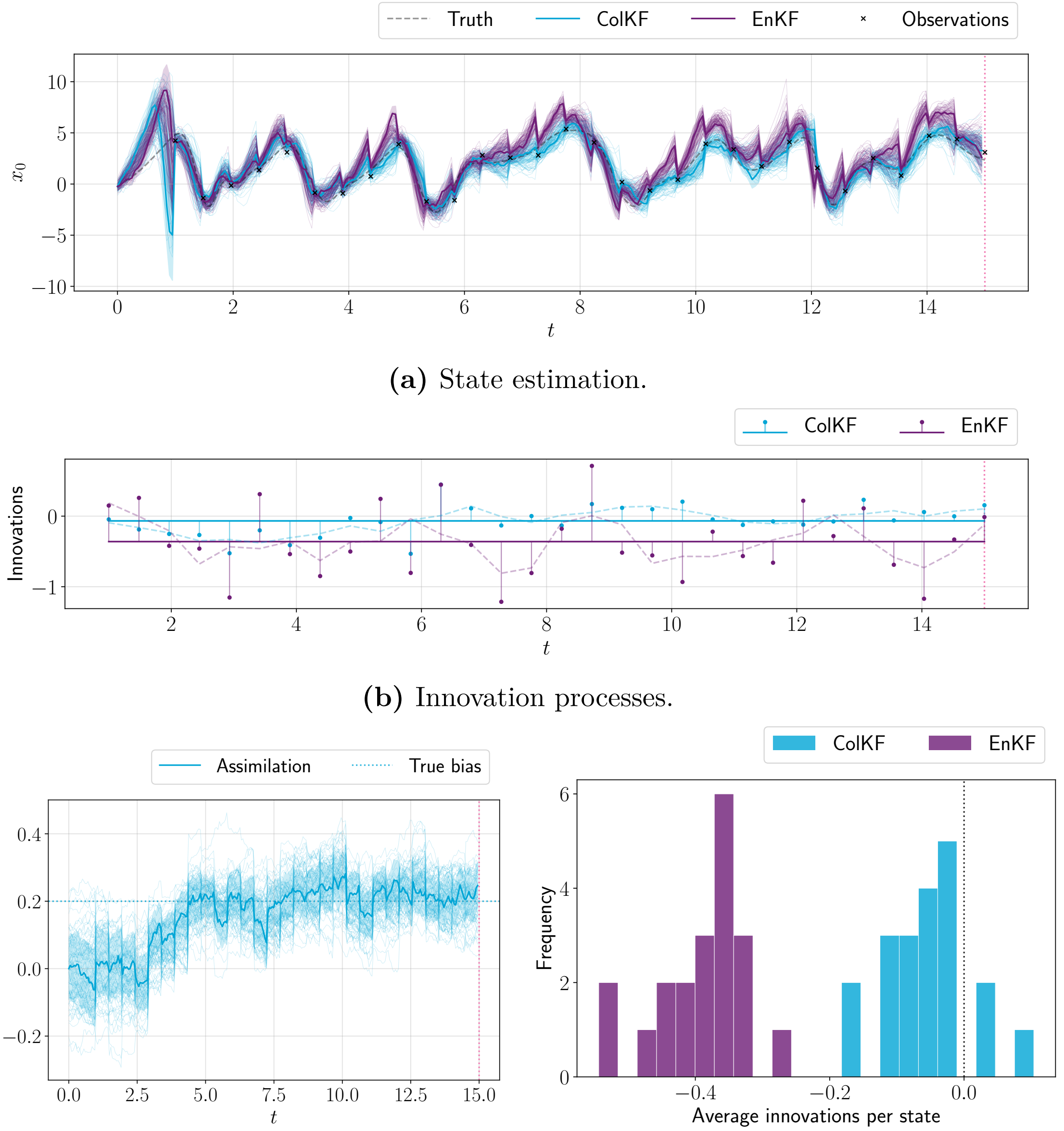
(a) State estimation for x_0.

(b) Innovation processes for x_0.

(c) Estimated AR process (proxy for bias), corresponding to x_0.

(d) Frequency histogram of average innovation for each state.

Figure 5.13: Results with the EnKF (purple) and with the ColKF (blue) for the Lorenz-96 model with added constant bias in *Analysis-Analysis update* experiment.



(a) State estimation.

(b) Innovation processes.

(c) Estimated AR process (proxy for bias), corresponding to x_0.

(d) Frequency histogram of average innovation for each state.

Figure 5.14: Results for x_0 with the EnKF (purple) and with the ColKF (blue) for the Lorenz-96 model with added constant bias in *Integration-Integration update* experiment.

5.3 Joint Parameter and Bias Estimation

This section presents the central results of this work. We start by showing the straightforward case of the model parameter not being distinguishable from bias in a continuous feedback formulation in the Lorenz-96 model. We begin by presenting the perturbed model used to generate the synthetic observations. The ODE of the j-th state, $j = 1, \dots, n_{\mathbf{x}}$ is given by

$$\dot{x}_j(t) = x_{j-1}(t)\left(x_{j+1}(t) - x_{j-2}(t)\right) - x_j(t) + F + b, \tag{5.1}$$

where $b \in \mathbb{R}$ is the added bias. It is clear from this formulation that any attempt to estimate the bias b and the forcing F will lead to a trade-off between the two quantities. This is tested using a bias estimation twin experiment (see Section 4.2), with model setup parameters shown in Tables 5.5, and the bias and forcing setup shown in Table 5.6.

Table 5.5: Model setup for parameter and bias estimation experiment on the Lorenz-96 model using a feedback ColKF.

$n_{\mathbf{x}}$	$\mathbf{H}_k$	$\boldsymbol{\mu}_0 \equiv \mathbf{x}_0$	$\mathbf{P}_0$	$\mathbf{Q}_k$	$\mathbf{R}_k$	T	K	Δt	Δk	N
20	$\mathbf{I}_{n_{\mathbf{x}}}$	$G(0, \mathbf{I}_{n_{\mathbf{x}}})$	$0.1\mathbf{I}_{n_{\mathbf{x}}}$	$0.05\mathbf{I}_{n_{\mathbf{x}}}$	$0.5\mathbf{I}_{n_{\mathbf{x}}}$	50	50	0.01	0.5	1000

Table 5.6: Bias and parameters setup for combined estimation on the Lorenz-96 model using a feedback ColKF.

$\mathbf{A}_k$	F^{true}	b^{true}	$\mathbb{E}[F_0]$	$\mathbb{V}[F_0]$	$\mathbb{E}[b_0]$	$\mathbb{V}[b_0]$	$\mathbf{Q}_k^{\mathbf{b}}$
1	7	1	9	4	2	4	0

Additionally, to make this effect more evident, we assume that the ColKF only adds one additional state to estimate the bias. This means that all equations are modified using the same AR process state estimate, and therefore, $n_{\mathbf{b}} = 1$ with $\mathbf{H}^{\mathbf{b}} = \mathbf{1}_{n_{\mathbf{x}}}$

The state estimation results were omitted since this experiment focuses only on combined parameter and bias estimation. The performance of the ColKF was already showcased in Section 5.2.2. The results of this experiment are summarized in Fig. 5.15. In particular, Figs. 5.15a and 5.15b exhibit the evolution of the estimated bias b and forcing F through the assimilation period, and Fig. 5.15c shows their sum. Just as expected, their sum is properly estimated as it converges to the anticipated value of 8, and it can also be observed that the uncertainty of this sum decreases significantly once the expected value is achieved (note that the standard deviation bands collapse around the estimated values after 10 steps). However, the individual bias and forcing estimation show that these quantities do not particularly converge to the true values. The uncertainty bands are still quite large compared to those of the sum, showing that, in this case, the ColKF is not estimating bias b and forcing F, but rather just their sum $F + b$ combined.

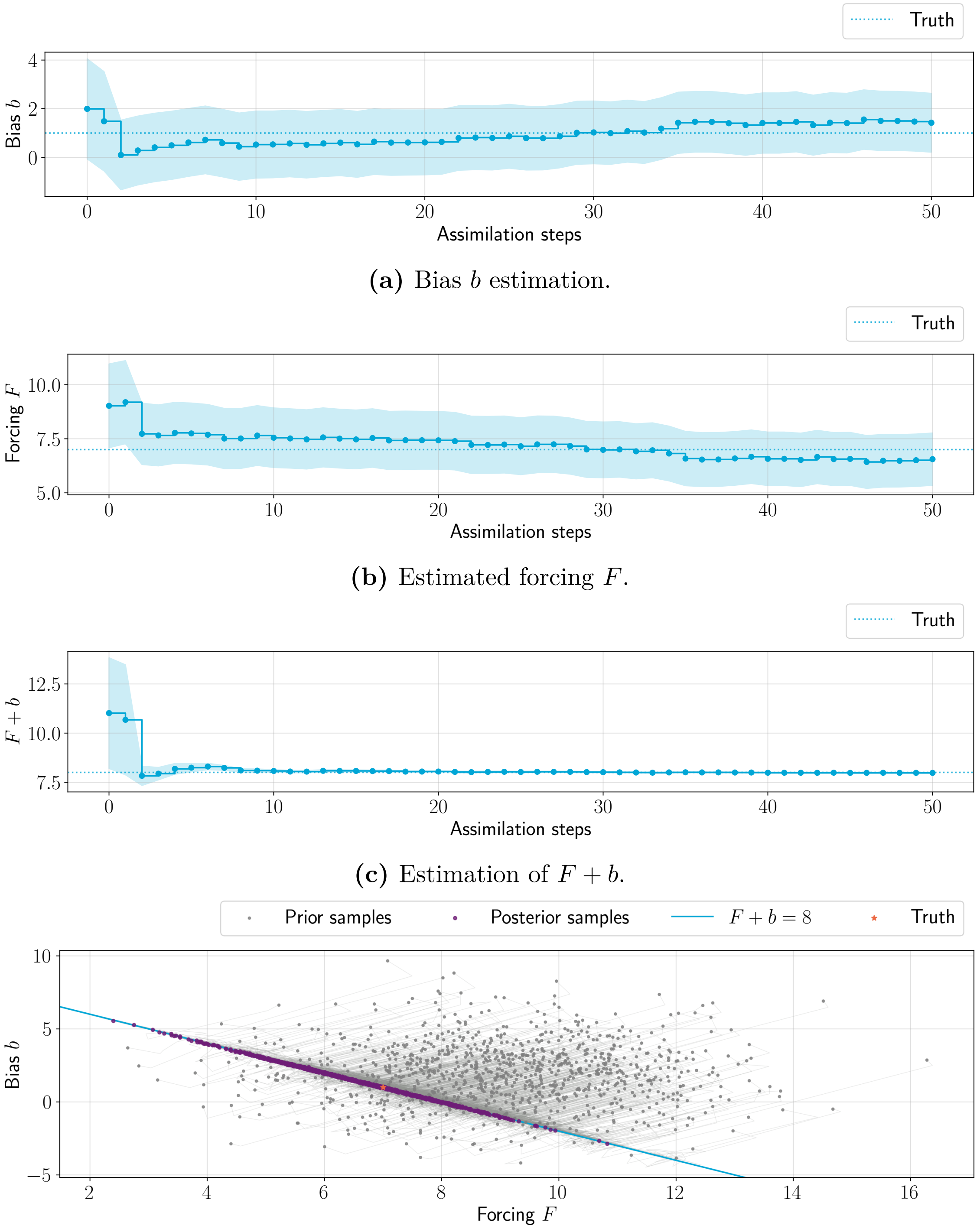
The final panel, namely Fig. 5.15d, is a better visualization of how the sum is estimated rather than each quantity separately. The figure shows the initial bias and forcings that were used for each ensemble member (priors, in gray dots) and the evolution of these over time (path showed in thin gray lines). The final estimates of these quantities (posteriors) are shown in the purple dots, and note how they all lie in the "optimal" line $F + b = 8$. The truth value $F = 7$ and $b = 1$ is plotted in orange for reference. In summary, this plot shows that the ColKF with feedback can only estimate the sum of both quantities and generates the presence of Pareto-like frontiers in the bias-parameter phase portrait.

Finally, we performed one last experiment, in this case with a no feedback version of the ColKF. Note that in this case, we also followed the same methodology for bias estimation twin experiment described in Section 4.2, but in this case $\mathbf{F} \equiv \mathbf{0}$ for a no feedback filter. The experiment uses the same model setup parameters presented in Table 5.6, but the bias and forcing setup is presented in Table 5.7. In addition, we use the bias-corrected forecast states from the ensemble to compute the forecast covariance matrix $\mathbf{P}_k^f$. This last remark is in line with the comment made in Section 4.1.2, regarding which state ensemble should be used to estimate $\mathbf{P}_k^f$ (and ultimately the Kalman gain) from the ensemble.

Table 5.7: Bias and parameters setup for combined estimation on the Lorenz-96 model using a feedback ColKF.

$\mathbf{A}_k$	F^{true}	b^{true}	$\mathbb{E}[F_0]$	$\mathbb{V}[F_0]$	$\mathbb{E}[b_0]$	$\mathbb{V}[b_0]$	$\mathbf{Q}_k^{\mathbf{b}}$
1	8	1	10	4	2	4	0

The obtained results are summarized in Fig. 5.16. In this case, the estimation of $F + b$ is no longer relevant, and this plot was omitted. Note how both the bias and the forcing are estimated correctly (Figs. 5.16a and 5.16b) and with high confidence (as the one standard deviation band is not visible). Moreover, from Fig. 5.16c, it can be observed that no Pareto-like front is obtained, and all the posterior estimates (purple dots) in the ensemble concentrate around the true value $F = 8$ and $b = 1$ (orange dot). This effectively implies that no feedback formulation of the ColKF can correctly estimate both the model parameter F and the bias b that was added to the synthetic observations.



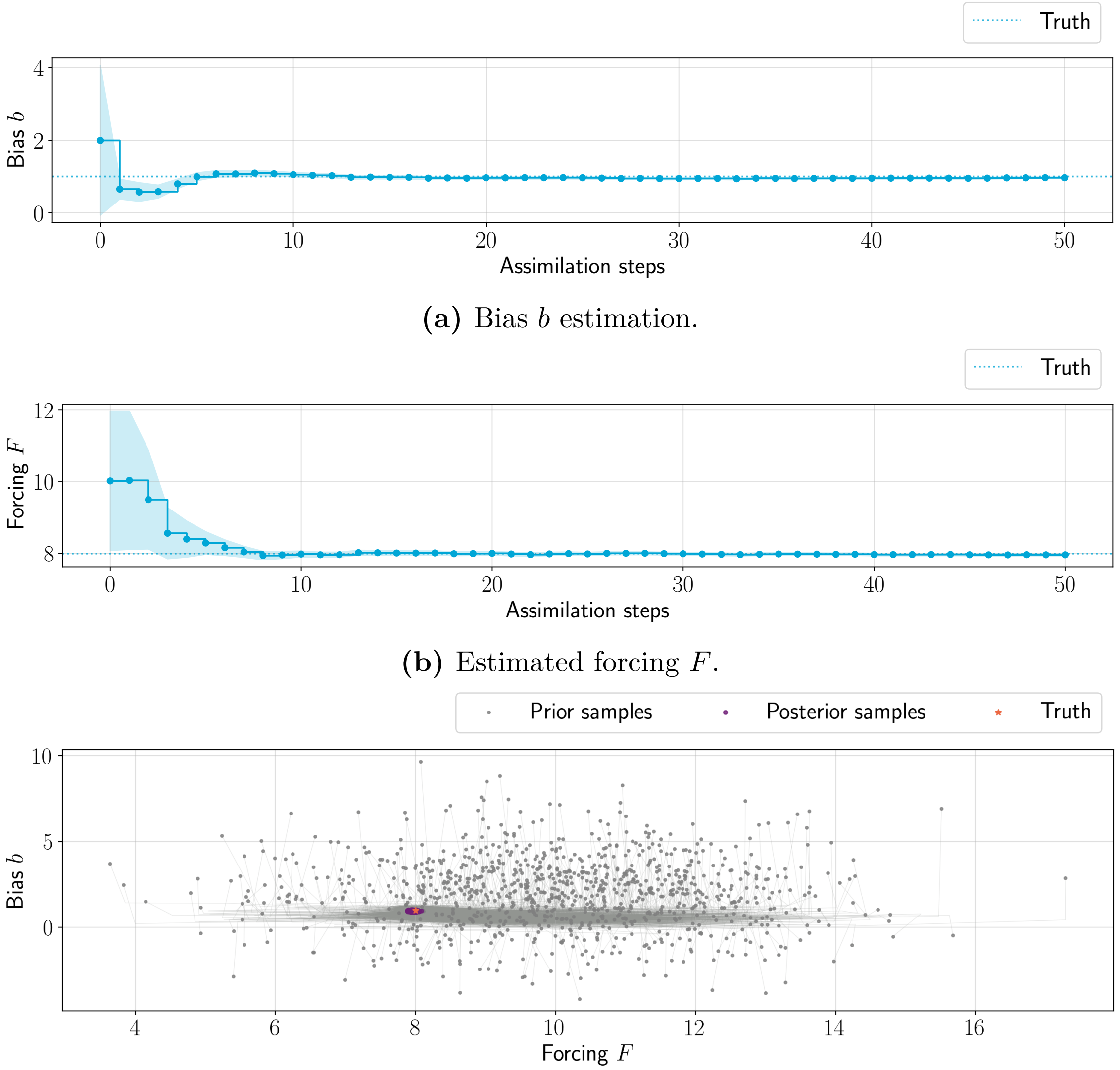
(a) Bias b estimation.

(b) Estimated forcing F.

(c) Estimation of $F + b$.

(d) *Prior* to *posterior* of bias b vs. forcing F. The initial ensemble of b and F is presented (gray dots) and their evolution in time until reaching the posterior estimates (purple dots), all laying on the $F + b = 8$ optimal line.

Figure 5.15: Joint forcing F and bias b estimation results for Lorenz-96 using a feedback ColKF approach.



(a) Bias b estimation.

(b) Estimated forcing F.

(c) *Prior* to *posterior* of bias b vs. forcing F. The initial ensemble of b and F is presented (gray dots) and their evolution in time until reaching the posterior estimates (purple dots), all laying on the $F + b = 8$ optimal line.

Figure 5.16: Joint forcing F and bias b estimation results for Lorenz-96 using a no feedback ColKF approach.

6

Discussion and Conclusion

In summary, this work implemented and applied a sequential bias-aware method, the ColKF, along with the EnKF, to several synthetic test scenarios in Python. In particular, twin experiments were conducted for the LHO and the Lorenz-96 model. These experiments included state estimation, state and parameter estimation, state and bias estimation, and all of these combined. The latter test addressed the objectives of this work in more detail, where the effect of combined parameters and bias estimation could (or not) be distinguished depending on the type of bias filter used (feedback or no feedback). Moreover, having worked on the implementation of the DA framework from scratch, several alternatives and challenges were identified during this process and included in this work (see Section 4.1.2).

From a practical point of view, the implementation developed offers a robust and flexible approach to state estimation for nonlinear dynamic systems. Its support for multiple filters, adaptability to various nonlinear ODE models, capability to explicitly modify forward models, parameter estimation, handling of temporal correlations, management of uncertainty, and unified result generation collectively contribute to its versatility in practical DA applications (see Section 4.3).

However, multiple challenges were encountered when working on this framework. Of course, the question about what needs to be considered for a general implementation remains open. New challenges will always arise when working with other DA techniques, new data types, new models, new computational requirements, and beyond. However, in Section 4.1.2, we presented some of the technical aspects that should be considered when implementing a general library for (bias-aware) DA. A detailed discussion on how to address these issues in a programming language is outside the scope of this work, and the solution to these problems often leads to increasing the complexity of the code base. However, I hope this can kick-start a new open-source DA framework in Python that can be used for future research endeavors or possibly incorporate these ideas into existing libraries. There are not a lot of libraries in Python for DA, and to the best of our knowledge, none allow for bias estimation explicitly, making this a step towards a first implementation or the inclusion within existing DA packages.

The framework was validated first through the use of twin experiments involving the EnKF for a LHO and the Lorenz-96 models. This process involved standard state estimation tests using synthetic observations, assimilation of these same observations but using a perturbed (biased) forecast model, and finally, using these observations to perform parameter estimation. In particular, for the Lorenz-96 model, the parameter estimation experiment also included a comparison between performing this experiment without and with covariance localization. This technique is found to be essential when performing DA experiments for the Lorenz-96 model.

This project successfully constructs an experimental setup for performing joint state, parameter, and bias estimation in one go. From a practical standpoint, it is hard to distinguish between model errors and misspecified (uncertain) parameters in the model, and we showed how they cannot always be confidently estimated together. In particular, for the Lorenz-96 model, we showed how the forcing parameter F cannot be distinguished from the bias in a feedback filter formulation, where the bias affects the ODE system explicitly. Whereas, in the case of a filter that **does not** use feedback, the parameter and the bias can be estimated independently. In addition, the emergence of "optimal curves" (Pareto-like frontiers) in the posterior phase portrait between parameters and bias (for all ensemble members) is evidence of these quantities being indistinguishable. Although this situation was tested assuming a constant and observable model bias, this outcome is an indication of the applicability of the methodology to a wider set of problems.

Moreover, this Pareto front that appeared in the feedback ColKF is clearly identifiable mainly since we assume that the bias state is the same for all states of the Lorenz-96, following the same behavior as the forcing F. However, in the case of a larger bias state, these curves may not be as easy to notice. One possible path forward is to repeat this same setup with a larger number of bias states. This would allow for a more compelling conclusion regarding more general bias-aware filters and possibly larger models.

In summary, this thesis showcased how both parameters and bias can be incorporated into a filter to be estimated using DA. Both practices share a lot of similarities in their formulation, and they even show similar effects on the resulting estimates. Still, they cannot always be properly addressed simultaneously, as we presented in the case of a feedback filter in the Lorenz-96 model. However, they were found to be distinguishable in the case of a non-feedback filter. This shows that how the bias estimate is fed back into the DA loop significantly impacts the results. This is a key finding since the literature often resorts to feedback formulations when accounting for bias, and it has been shown that this can lead to combinations of parameters and bias that can fit the biased observations. Yet, they do not correspond to the actual values that were expected.

An additional remark needs to be made about how the Kalman gain in a non-feedback version of a filter can be calculated in two different ways while still satisfying the “requirements' for it to be a non-feedback filter. In particular, the covariance matrix can be estimated using the bias-corrected ensemble of forecast states. This approach is still a no-feedback version, as the biased state would still be used for the forward model. Along this same line, the feedback version of the filter also leaves room for modifications/different approaches. As described in Section 4.1.2, the bias that gets fed back into the state can be applied before or after time discretization. Namely, the bias can directly affect the continuous dynamics or be added after every integrator step is performed. On top of other factors like frequency for stochastic updates, length of the bias state, parameterization of the bias model, initialization, and uncertainty of bias, these variations contribute to the already myriad of possible strategies in bias-aware filtering, like the ones discussed by De Lannoy et al. [122]. In this sense, the bias-aware technique is highly dependent on the model's quality and the estimates' use.

Furthermore, the observations are taken to be “far” apart in time for the experiments in this work, primarily to showcase the effect of bias and parameter estimation and to mimic a configuration in an operational setting. However, it is known that if observations are taken with higher frequency, the impact of having a biased prediction model is mitigated on the resulting analysis [4]. This could be a path that future research can investigate: how sensitive the results presented here are to the frequency of observations assimilated.

Looking ahead, there are several more directions for future research. One potential avenue is to explore the methodology developed in this thesis in practical, real-world settings, which would help validate the results and conclusions of this work for other applications in DA. Additionally, it would be beneficial to replicate the results for more complex structures of bias, such as time-varying and state-dependent biases, and to explore model errors that include missing components of the model, e.g., missing states or connectedness of cells in the Lorenz-96 model. Furthermore, investigating the efficacy of the SepKF and other bias-aware methods could offer valuable insights into performing joint parameters and bias estimation in a non-feedback formulation.

References

[1] G. Evensen, F. Vossepoel, and P. van Leeuwen. *Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem.* Springer Textbooks in Earth Sciences, Geography and Environment. Springer International Publishing, 2022. ISBN: 9783030967093.

[2] M. Asch, M. Bocquet, and M. Nodet. *Data Assimilation: Methods, Algorithms, and Applications.* Fundamentals of Algorithms. Society for Industrial and Applied Mathematics, 2016. ISBN: 9781611974546.

[3] S. Fletcher. *Data Assimilation for the Geosciences: From Theory to Application.* Elsevier Science, 2022. ISBN: 9780323972536.

[4] D. P. Dee. "Bias and Data Assimilation". In: *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography* 131.613 (2005), pp. 3323–3343.

[5] A. Nóvoa, A. Racca, and L. Magri. "Inferring Unknown Unknowns: Regularized Bias-aware Ensemble Kalman Filter". In: *Computer Methods in Applied Mechanics and Engineering* 418 (2024), p. 116502.

[6] D. P. Dee and A. M. Da Silva. "Data Assimilation in the Presence of Forecast Bias". In: *Quarterly Journal of the Royal Meteorological Society* 124.545 (1998), pp. 269–295.

[7] J. P. Drécourt, H. Madsen, and D. Rosbjerg. "Bias Aware Kalman Filters: Comparison and Improvements". In: *Advances in Water Resources* 29.5 (2006), pp. 707–718.

[8] J. J. Ruiz, M. Pulido, and T. Miyoshi. "Estimating Model Parameters with Ensemble-based Data Assimilation: A Review". In: *Journal of the Meteorological Society of Japan. Ser. II* 91.2 (2013), pp. 79–99.

[9] H. Diab Montero. "Ensemble Data Assimilation Methods for Estimating Fault Slip and Future Earthquake Occurrences". English. Dissertation (TU Delft). Delft University of Technology, 2024. ISBN: 978-94-6384-568-7. DOI: 10.4233/uuid:92ebf293-d057-4ab3-b11a-ba2204f377e8.

[10] G. Evensen, D. P. Dee, and J. Schröter. "Parameter Estimation in Dynamical Models". In: *Ocean Modeling and Parameterization* (1998), pp. 373–398.

[11] D. Erdal, I. Neuweiler, and U. Wollschläger. "Using a Bias Aware EnKF to Account for Unresolved Structure in an Unsaturated Zone Model". In: *Water Resources Research* 50.1 (2014), pp. 132–147.

[12] C. Chui and G. Chen. *Kalman Filtering: with Real-Time Applications.* Springer International Publishing, 2017. ISBN: 9783319476100.

[13] D. van Kekem. "Dynamics of the Lorenz-96 model: Bifurcations, Symmetries and Waves". English. PhD thesis. University of Groningen, 2018. ISBN: 978-94-034-0979-5.

[14] R. H. Reichle. "Data Assimilation Methods in the Earth Sciences". In: *Advances in Water Resources* 31.11 (2008), pp. 1411–1418.

[15] B. Wang et al. "Kalman Filter and Its Application in Data Assimilation". In: *Atmosphere* 14.8 (2023), p. 1319.

[16] R. N. Bannister. "A Review of Operational Methods of Variational and Ensemble-variational Data Assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 143.703 (2017), pp. 607–633.

[17] F. X. Le Dimet et al. "Data Assimilation in Hydrology: Variational Approach". In: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications* (2009), pp. 367–405.

[18] M. D'Elia et al. "Applications of Variational Data Assimilation in Computational Hemodynamics". In: *Modeling of Physiological Flows* (2012), pp. 363–394.

[19] J. A. Cummings and O. M. Smedstad. "Variational Data Assimilation for the Global Ocean". In: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II).* Springer, 2013, pp. 303–343.

[20] I. Hoteit et al. "Data Assimilation in Oceanography: Current Status and New Directions". In: *New Frontiers in Operational Oceanography* (2018), pp. 465–512.

[21] J. Jin et al. "Spatially Varying Parameter Estimation for Dust Emissions using Reduced-tangent-linearization 4DVar". In: *Atmospheric Environment* 187 (2018), pp. 358–373.

[22] J. Jin et al. "Machine Learning for Observation Bias Correction with Application to Dust Storm Data Assimilation". In: *Atmospheric Chemistry and Physics* 19.15 (2019), pp. 10009–10026.

[23] A. Yarce Botero et al. “Estimating NOx LOTOS-EUROS CTM Emission Parameters over the Northwest of South America through 4DEnVar TROPOMI NO2 Assimilation”. In: *Atmosphere* 12.12 (2021), p. 1633.

[24] A. Yarce Botero. “Data Assimilation in a LOTOS-EUROS Chemical Transport Model for Colombia Using Satellite Measurements”. English. Dissertation (TU Delft). Delft University of Technology, 2024. ISBN: 978-90-834024-2-0. DOI: 10.4233/uuid:e202a1f3-9c73-42d1-b7f6-d45f9631df74.

[25] M. Kano et al. “Adjoint-based Direct Data Assimilation of GNSS Time Series for Optimizing Frictional Parameters and Predicting Postseismic Deformation Following the 2003 Tokachi-oki Earthquake”. In: *Earth, Planets and Space* 72 (2020), pp. 1–24.

[26] O. M. Smedstad and J. J. O’Brien. “Variational Data Assimilation and Parameter Estimation in an Equatorial Pacific Ocean Model”. In: *Progress in Oceanography* 26.2 (1991), pp. 179–241.

[27] F. C. Vossepoel et al. “Adjustment of Near-equatorial Wind Stress with Four-dimensional Variational Data Assimilation in a Model of the Pacific Ocean”. In: *Monthly Weather Review* 132.8 (2004), pp. 2070–2083.

[28] Q. Huang, F. Primeau, and T. DeVries. “CYCLOCIM: A 4-D Variational Assimilation System for the Climatological Mean Seasonal Cycle of the Ocean Circulation”. In: *Ocean Modelling* 159 (2021), p. 101762.

[29] A. Carrassi et al. “Data Assimilation in the Geosciences: An Overview of Methods, Issues, and Perspectives”. In: *Wiley Interdisciplinary Reviews: Climate Change* 9.5 (2018), e535.

[30] L. Bertino, G. Evensen, and H. Wackernagel. “Sequential Data Assimilation Techniques in Oceanography”. In: *International Statistical Review* 71.2 (2003), pp. 223–241. DOI: https://doi.org/10.1111/j.1751-5823.2003.tb00194.x.

[31] S. I. Aanonsen et al. “The ensemble Kalman filter in reservoir engineering—a review”. In: *Spe Journal* 14.03 (2009), pp. 393–412.

[32] N. Schutgens et al. “Applying an Ensemble Kalman Filter to the Assimilation of AERONET Observations in a Global Aerosol Transport Model”. In: *Atmospheric Chemistry and Physics* 10.5 (2010), pp. 2561–2576.

[33] M. Pang et al. “Dust storm forecasting through coupling LOTOS-EUROS with localized ensemble Kalman filter”. In: *Atmospheric Environment* 306 (2023), p. 119831.

[34] S. Lopez-Restrepo et al. “Forecasting PM10 and PM2. 5 in the Aburrá Valley (Medellín, Colombia) via EnKF based data assimilation”. In: *Atmospheric Environment* 232 (2020), p. 117507.

[35] G. Evensen. “Sequential Data Assimilation for Nonlinear Dynamics: The Ensemble Kalman Filter”. In: *Ocean Forecasting: Conceptual Basis and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 97–116. ISBN: 978-3-662-22648-3. DOI: 10.1007/978-3-662-22648-3_6.

[36] I. M. Navon. “Data Assimilation for Numerical Weather Prediction: A Review”. In: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications* (2009), pp. 21–65.

[37] A. Sandu and T. Chai. “Chemical Data Assimilation — An overview”. In: *Atmosphere* 2.3 (2011), pp. 426–463.

[38] C. Montzka et al. “Multivariate and Multiscale Data Assimilation in Terrestrial Systems: A Review”. In: *Sensors* 12.12 (2012), pp. 16291–16333.

[39] W. Lahoz, B. Khattatov, and R. Menard. *Data Assimilation: Making Sense of Observations*. Springer Berlin Heidelberg, 2010. ISBN: 9783540747031.

[40] S. Park and L. Xu. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*. Earth and Environmental Science. Springer Berlin Heidelberg, 2009. ISBN: 9783540710561.

[41] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Earth and Environmental Science. Springer Berlin Heidelberg, 2009. ISBN: 9783642037115.

[42] R. E. Kalman. “A New Approach to Linear Filtering and Prediction Problems”. In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. ISSN: 0021-9223. DOI: 10.1115/1.3662552.

[43] R. E. Kalman and R. S. Bucy. “New Results in Linear Filtering and Prediction Theory”. In: *Journal of Basic Engineering* 83.1 (Mar. 1961), pp. 95–108. ISSN: 0021-9223. DOI: 10.1115/1.3658902.

[44] M. S. Grewal, V. D. Henderson, and R. S. Miyasako. “Application of Kalman Filtering to the Calibration and Alignment of Inertial Navigation Systems”. In: *29th IEEE Conference on Decision and Control*. IEEE. 1990, pp. 3325–3334.

[45] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2003. ISBN: 9781580538510.

[46] A. Barrau and S. Bonnabel. “Invariant Kalman Filtering”. In: *Annual Review of Control, Robotics, and Autonomous Systems* 1.1 (2018), pp. 237–257.

[47] J. Sasiadek and P. Hartana. “Sensor Data Fusion using Kalman Filter”. In: *Proceedings of the Third International Conference on Information Fusion.* Vol. 2. IEEE. 2000, WED5–19.

[48] F. Caron et al. “GPS/IMU Data Fusion using Multisensor Kalman Filtering: Introduction of Contextual Aspects”. In: *Information fusion* 7.2 (2006), pp. 221–230.

[49] V. P. Oikonomou et al. “The Use of Kalman Filter in Biomedical Signal Processing”. In: *Kalman Filter: Recent Advances and Applications* (2009).

[50] M. Roth et al. “The Ensemble Kalman Filter: A Signal Processing Perspective”. In: *EURASIP Journal on Advances in Signal Processing* 2017 (2017), pp. 1–16.

[51] C. Urrea and R. Agramonte. “Kalman Filter: Historical Overview and Review of its Use in Robotics 60 Years After its Creation”. In: *Journal of Sensors* 2021.1 (2021), p. 9674015.

[52] F. Auger et al. “Industrial Applications of the Kalman Filter: A Review”. In: *IEEE Transactions on Industrial Electronics* 60.12 (2013), pp. 5458–5471.

[53] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* ISSN. Elsevier Science, 1970. ISBN: 9780080960906.

[54] K. Reif et al. “Stochastic Stability of the Discrete-time Extended Kalman Filter”. In: *IEEE Transactions on Automatic control* 44.4 (1999), pp. 714–728.

[55] S. J. Julier and J. K. Uhlmann. “New Extension of the Kalman Filter to Nonlinear Systems”. In: *Signal Processing, Sensor Fusion, and Target Recognition VI.* Vol. 3068. Spie. 1997, pp. 182–193.

[56] S. J. Julier and J. K. Uhlmann. “Unscented Filtering and Nonlinear Estimation”. In: *Proceedings of the IEEE* 92.3 (2004), pp. 401–422.

[57] M. Verlaan and A. W. Heemink. “Tidal Flow Forecasting using Reduced Rank Square Root Filters”. In: *Stochastic hydrology and Hydraulics* 11 (1997), pp. 349–368.

[58] S. E. Cohn. “An Introduction to Estimation Theory (Special Issue – Data Assimilation in Meteorology and Oceanography: Theory and Practice)”. In: *Journal of the Meteorological Society of Japan. Ser. II* 75.1B (1997), pp. 257–288.

[59] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering.* Vol. 3. Springer, 2009.

[60] D. Catlin. *Estimation, Control, and the Discrete Kalman Filter.* Applied Mathematical Sciences. Springer New York, 2012. ISBN: 9781461245285.

[61] M. Grewal and A. Andrews. *Kalman Filtering: Theory and Practice with MATLAB*. IEEE Press. Wiley, 2014. ISBN: 9781118851210.

[62] F. Govaers. *Introduction and Implementations of the Kalman Filter*. IntechOpen, 2019. ISBN: 9781838805364.

[63] J. H. van Schuppen. *Control and System Theory of Discrete-Time Stochastic systems*. Springer, 2021.

[64] G. Evensen. "The Ensemble Kalman Filter for Combined State and Parameter Estimation". In: *IEEE Control Systems Magazine* 29.3 (2009), pp. 83–104.

[65] G. Evensen. "Sequential Data Assimilation with a Nonlinear Quasi-geostrophic Model using Monte Carlo Methods to Forecast Error Statistics". In: *Journal of Geophysical Research: Oceans* 99.C5 (1994), pp. 10143–10162.

[66] G. Evensen. "The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation". In: *Ocean dynamics* 53 (2003), pp. 343–367.

[67] J. Mandel. *Efficient Implementation of the Ensemble Kalman Filter*. University of Colorado at Denver and Health Sciences Center, Center for Computational Mathematics, 2006.

[68] Y. C. and C. S. "Assimilating Vortex Position with an Ensemble Kalman Filter". In: *Monthly Weather Review* 135.5 (2007), pp. 1828–1845. DOI: `10.1175/MWR3351.1`.

[69] J. Mandel. "A Brief Tutorial on the Ensemble Kalman Filter". In: *arXiv preprint arXiv:0901.3725* (2009).

[70] M. van Loon, P. J. Builtjes, and A. J. Segers. "Data Aassimilation of Ozone in the Atmospheric Transport Chemistry Model LOTOS". In: *Environmental Modelling & Software* 15.6-7 (2000), pp. 603–609.

[71] M. Pang et al. "Neighbouring Time Ensemble Kalman Filter (NTEnKF) Data Assimilation for Dust Storm Forecasting". In: *Geoscientific Model Development Discussions* 2023 (2023), pp. 1–28.

[72] C. Huang and X. Li. "Experiments of Soil Moisture Data Assimilation System based on Ensemble Kalman Filter". In: *Plateau Meteorol* 4 (2006), p. 013.

[73] Z. Deng, Y. Tang, and H. J. Freeland. "Evaluation of Several Model Error Schemes in the EnKF Assimilation: Applied to Argo Profiles in the Pacific Ocean". In: *Journal of Geophysical Research: Oceans* 116.C9 (2011).

[74] V. Kitsios et al. "Ensemble Kalman Filter Parameter Estimation of Ocean Optical Properties for Reduced Biases in a Coupled General Circulation Model". In: *Journal of Advances in Modeling Earth Systems* 13.2 (2021), e2020MS002252.

[75] S. Ohishi, T. Miyoshi, and M. Kachi. “An Ensemble Kalman Filter-based Ocean Data Assimilation System Improved by Adaptive Observation Error Inflation (AOEI)”. In: *Geoscientific Model Development* 15.24 (2022), pp. 9057–9073.

[76] *Using the EnKF for Assisted History Matching of a North Sea Reservoir Model.* SPE Reservoir Simulation Conference. Feb. 2007, SPE-106184–MS. DOI: 10.2118/106184-MS.

[77] M. Krymskaya, R. Hanea, and M. Verlaan. “An Iterative Ensemble Kalman Filter for Reservoir Engineering Applications”. In: *Computational Geosciences* 13 (2009), pp. 235–244.

[78] F. Akter et al. “Modified Ensemble Kalman Filter for Reservoir Parameter and State Estimation in the Presence of Model Uncertainty”. In: *Journal of Petroleum Science and Engineering* 199 (2021), p. 108323.

[79] P. L. Houtekamer et al. “Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations”. In: *Monthly Weather Review* 133.3 (2005), pp. 604–620.

[80] M. Wei et al. “Ensemble Transform Kalman Filter-based Ensemble Perturbations in an Operational Global Prediction System at NCEP”. In: *Tellus A: Dynamic Meteorology and Oceanography* 58.1 (2006), pp. 28–44.

[81] M. Bonavita, L. Torrisi, and F. Marcucci. “The Ensemble Kalman Filter in an Operational Regional NWP System: Preliminary Results with Real Observations”. In: *Quarterly Journal of the Royal Meteorological Society* 134.636 (2008), pp. 1733–1744.

[82] G. Y. El Serafy and A. E. Mynett. “Improving the Operational Forecasting System of the Stratified Flow in Osaka Bay using an Ensemble Kalman Filter–based Steady State Kalman Filter”. In: *Water Resources Research* 44.6 (2008).

[83] P. L. Houtekamer, H. L. Mitchell, and X. Deng. “Model Error Representation in an Operational Ensemble Kalman Filter”. In: *Monthly Weather Review* 137.7 (2009), pp. 2126–2143.

[84] H. McMillan et al. “Operational Hydrological Data Assimilation with the Recursive Ensemble Kalman Filter”. In: *Hydrology and Earth System Sciences* 17.1 (2013), pp. 21–38.

[85] P. L. Houtekamer et al. “Higher Resolution in an Operational Ensemble Kalman Filter”. In: *Monthly Weather Review* 142.3 (2014), pp. 1143–1162.

[86] A. Rafieeinasab et al. “Comparative Evaluation of Maximum Likelihood Ensemble Filter and Ensemble Kalman Filter for Real-time Assimilation of

Streamflow Data into Operational Hydrologic Models". In: *Journal of Hydrology* 519 (2014), pp. 2663–2675.

[87] K. Sueki et al. "Precision and Convergence Speed of the Ensemble Kalman Filter-based Parameter Estimation: Setting Parameter Uncertainty for Reliable and Efficient Estimation". In: *Progress in Earth and Planetary Science* 9.1 (2022), p. 47.

[88] D. Canuto et al. "An Ensemble Kalman Filter Approach to Parameter Estimation for Patient-Specific Cardiovascular Flow Modeling". In: *Theoretical and Computational Fluid Dynamics* 34 (2020), pp. 521–544.

[89] J. Annan et al. "Parameter Estimation in an Intermediate Complexity Earth System Model Using an Ensemble Kalman Filter". In: *Ocean Modelling* 8.1-2 (2005), pp. 135–154.

[90] H. Moradkhani et al. "Dual State–Parameter Estimation of Hydrological Models Using Ensemble Kalman Filter". In: *Advances in Water Resources* 28.2 (2005), pp. 135–147.

[91] S. Zhang et al. "Coupled Data Assimilation and Parameter Estimation in Coupled Ocean–Atmosphere Models: A Review". In: *Climate Dynamics* 54 (2020), pp. 5127–5144.

[92] B. Friedland. "Treatment of Bias in Recursive Filtering". In: *IEEE Transactions on Automatic Control* 14.4 (1969), pp. 359–367.

[93] E. Tacker and C. Lee. "Linear Filtering in the Presence of Time-varying Bias". In: *IEEE Transactions on Automatic Control* 17.6 (1972), pp. 828–829.

[94] J. Mendel. "Extension of Friedland's Bias Filtering Technique to a Class of Nonlinear Systems". In: *IEEE Transactions on Automatic Control* 21.2 (1976), pp. 296–298.

[95] B. Friedland. "Notes on Separate-bias Estimation". In: *IEEE Transactions on Automatic Control* 23.4 (1978), pp. 735–738.

[96] M. B. Ignagni. "Separate Bias Kalman Estimator with Bias State Noise". In: *IEEE Transactions on Automatic Control* 35.3 (1990), pp. 338–341.

[97] D. H. Zhou et al. "Extension of Friedland's Separate-bias Estimation to Randomly Time-varying Bias of Nonlinear Systems". In: *IEEE Transactions on Automatic Control* 38.8 (1993), pp. 1270–1273. DOI: 10.1109/9.233167.

[98] J. Rasmussen et al. "Data Assimilation in Integrated Hydrological Modelling in the Presence of Observation Bias". In: *Hydrology and Earth System Sciences* 20.5 (2016), pp. 2103–2118.

[99] S. J. Baek et al. "Local Ensemble Kalman Filtering in the Presence of Model Bias". In: *Tellus A: Dynamic Meteorology and Oceanography* 58.3 (2006), pp. 293–306.

[100] R. Ménard. "Bias Estimation". In: *Data Assimilation: Making Sense of Observations* (2010), pp. 113–135.

[101] M. Glegola, R. Hanea, and G. Kaleta. "Bias Aware Data Assimilation in Reservoir Characterization". In: *ECMOR XIV-14th European Conference on the Mathematics of Oil Recovery*. Vol. 2014. 1. European Association of Geoscientists & Engineers. 2014, pp. 1–18.

[102] R. Lorente-Plazas and J. P. Hacker. "Observation and Model Bias Estimation in the Presence of Either or Both Sources of Error". In: *Monthly Weather Review* 145.7 (2017), pp. 2683–2696.

[103] C. Aggarwal. *Neural Networks and Deep Learning: A Textbook*. Springer International Publishing, 2023. ISBN: 9783031296420.

[104] L. Grigoryeva and J. Ortega. "Echo State Networks are Universal". In: *Neural Networks* 108 (2018), pp. 495–508.

[105] A. Nóvoa and L. Magri. "Real-time Thermoacoustic Data Assimilation". In: *Journal of Fluid Mechanics* 948 (2022), A35.

[106] A. Novoa, A. Racca, and L. Magri. "Bias-aware Thermoacoustic Data Assimilation". In: *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*. Vol. 265. 6. Institute of Noise Control Engineering. 2023, pp. 1924–1931.

[107] S. K. Tamang et al. "Ensemble Riemannian Data Assimilation: Towards Large-scale Dynamical Systems". In: *Nonlinear Processes in Geophysics* 29.1 (2022), pp. 77–92.

[108] T. G. Asher et al. "Low Frequency Water Level Correction in Storm Surge Models using Data Assimilation". In: *Ocean Modelling* 144 (2019), p. 101483.

[109] J. C. Derber and W. S. Wu. "The Use of TOVS Cloud-cleared Radiances in the NCEP SSI Analysis System". In: *Monthly Weather Review* 126.8 (1998), pp. 2287–2299.

[110] E. Fertig et al. "Observation Bias Correction with an Ensemble Kalman Filter". In: *Tellus A: Dynamic Meteorology and Oceanography* 61.2 (2009), pp. 210–226.

[111] J. R. Eyre. "Observation Bias Correction Schemes in Data Assimilation Systems: A Theoretical Study of Some of their Properties". In: *Quarterly Journal of the Royal Meteorological Society* 142.699 (2016), pp. 2284–2291.

[112] M. Ridler et al. “Bias-Aware Data Assimilation in Integrated Hydrological Modelling”. In: *Hydrology Research* 49.4 (2018), pp. 989–1004.

[113] J. V. T. Sørensen and H. Madsen. “Parameter Sensitivity of Three Kalman Filter Schemes for Assimilation of Water Levels in Shelf Sea Models”. In: *Ocean Modelling* 11.3-4 (2006), pp. 441–463.

[114] A. F. Da Silva and T. Colonius. “A Bias-Aware EnKF Estimator for Aerodynamic Flows”. In: *2018 Fluid Dynamics Conference*. 2018, p. 3225.

[115] S. Chumchean, A. Seed, and A. Sharma. “Correcting of Real-time Radar Rainfall Bias using a Kalman Filtering Approach”. In: *Journal of Hydrology* 317.1-2 (2006), pp. 123–137.

[116] N. F. Raboudi et al. “Ensemble Kalman Filtering with Coloured Observation Noise”. In: *Quarterly Journal of the Royal Meteorological Society* 147.741 (2021), pp. 4408–4424.

[117] B. Cao et al. “Assessing the performance of separate bias Kalman filter in correcting the model bias for estimation of soil moisture profiles”. In: *Journal of Meteorological Research* 33.3 (2019), pp. 519–527.

[118] L. Nerger and W. W. Gregg. “Improving Assimilation of SeaWiFS Data by the Application of Bias Correction with a Local SEIK Filter”. In: *Journal of Marine Systems* 73.1-2 (2008), pp. 87–102.

[119] G. A. Chepurin, J. A. Carton, and D. Dee. “Forecast Model Bias Correction in Ocean Data Assimilation”. In: *Monthly Weather Review* 133.5 (2005), pp. 1328–1342.

[120] Z. Deng, Y. Tang, and G. Wang. “Assimilation of Argo Temperature and Salinity Profiles using a Bias-aware Localized EnKF System for the Pacific Ocean”. In: *Ocean Modelling* 35.3 (2010), pp. 187–205.

[121] G. J. De Lannoy et al. “State and Bias Estimation for Soil Moisture Profiles by an Ensemble Kalman Filter: Effect of Assimilation Depth and Frequency”. In: *Water Resources Research* 43.6 (2007).

[122] G. J. De Lannoy et al. “Correcting for Forecast Bias in Soil Moisture Assimilation with the Ensemble Kalman Filter”. In: *Water Resources Research* 43.9 (2007).

[123] V. R. Pauwels et al. “Simultaneous Estimation of Model State Variables and Observation and Forecast Biases using a Two-stage Hybrid Kalman Filter”. In: *Hydrology and Earth System Sciences* 17.9 (2013), pp. 3499–3521.

[124] C. Draper et al. “A Dynamic Approach to Addressing Observation-Minus-Forecast Bias in a Land Surface Skin Temperature Data Assimilation System”. In: *Journal of Hydrometeorology* 16.1 (2015), pp. 449–464.

[125] X. Yang and T. Delsole. “Using the Ensemble Kalman Filter to Estimate Multiplicative Model Parameters”. In: *Tellus A: Dynamic Meteorology and Oceanography* 61.5 (2008), pp. 601–609.

[126] J. M. Lewis, S. Lakshmivarahan, and S. Dhall. *Dynamic Data Assimilation: A Least Squares Approach.* Vol. 13. Cambridge University Press, 2006.

[127] I. Reid and H. Term. “Estimation II”. In: *University of Oxford, Lecture Notes* (2001).

[128] G. Burgers, P. J. van Leeuwen, and G. Evensen. “Analysis Scheme in the Ensemble Kalman Filter”. In: *Monthly Weather Review* 126.6 (1998), pp. 1719–1724.

[129] C. H. Bishop, B. J. Etherton, and S. J. Majumdar. “Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects”. In: *Monthly Weather Review* 129.3 (2001), pp. 420–436.

[130] J. S. Whitaker and T. M. Hamill. “Ensemble data assimilation without perturbed observations”. In: *Monthly weather review* 130.7 (2002), pp. 1913–1924.

[131] J. L. Anderson. “An Ensemble Adjustment Kalman Filter for Data Assimilation”. In: *Monthly Weather Review* 129.12 (2001), pp. 2884–2903.

[132] M. Zupanski. “Maximum Likelihood Ensemble Filter: Theoretical Aspects”. In: *Monthly Weather Review* 133.6 (2005), pp. 1710–1726.

[133] P. Sakov and P. R. Oke. “A Deterministic Formulation of the Ensemble Kalman Filter: An Alternative to Ensemble Square Root Filters”. In: *Tellus A: Dynamic Meteorology and Oceanography* 60.2 (2008), pp. 361–371.

[134] E. D. Nino-Ruiz and A. Sandu. “Ensemble Kalman Filter Implementations Based on Shrinkage Covariance Matrix Estimation”. In: *Ocean Dynamics* 65 (2015), pp. 1423–1439.

[135] S. Lopez-Restrepo et al. “An Efficient Ensemble Kalman Filter Implementation via Shrinkage Covariance Matrix Estimation: Exploiting Prior Knowledge”. In: *Computational Geosciences* 25 (2021), pp. 985–1003.

[136] S. Lopez-Restrepo et al. “A Robust Ensemble-based Data Assimilation Method using Shrinkage Estimator and Adaptive Inflation”. In: *Authorea Preprints* (2022).

[137] R. Petrie. “Localization in the Ensemble Kalman Filter”. In: *MSc Atmosphere, Ocean and Climate University of Reading* 460 (2008).

[138] P. L. Houtekamer and H. L. Mitchell. “A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation”. In: *Monthly Weather Review* 129.1 (2001), pp. 123–137.

[139] T. M. Hamill, J. S. Whitaker, and C. Snyder. “Distance-dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter”. In: *Monthly Weather Review* 129.11 (2001), pp. 2776–2790.

[140] G. Gaspari and S. E. Cohn. “Construction of Correlation Functions in Two and Three Dimensions”. In: *Quarterly Journal of the Royal Meteorological Society* 125.554 (1999), pp. 723–757.

[141] R. Shumway and D. Stoffer. *Time Series Analysis and Its Applications: With R Examples*. Springer Texts in Statistics. Springer International Publishing, 2017. ISBN: 9783319524528.

[142] E. N. Lorenz. “Predictability: A Problem Partly Solved”. In: *Proc. Seminar on Predictability*. Vol. 1. 1. Reading. ECMWF, 1996, pp. 1–18.

[143] E. N. Lorenz. “Predictability – A Problem Partly Solved”. In: *Predictability of Weather and Climate*. Ed. by T. Palmer and R. Hagedorn. Cambridge University Press, 2006, pp. 40–58.

[144] E. N. Lorenz. “The Local Structure of a Chaotic Attractor in Four Dimensions”. In: *Physica D: Nonlinear Phenomena* 13.1 (1984), pp. 90–104. ISSN: 0167-2789. DOI: `https://doi.org/10.1016/0167-2789(84)90272-0`.

[145] L. Dieci, M. S. Jolly, and E. S. Van Vleck. “Numerical Techniques for Approximating Lyapunov Exponents and Their Implementation”. In: *Journal of Computational and Nonlinear Dynamics* 6.1 (Sept. 2010), p. 011003. ISSN: 1555-1415. DOI: `10.1115/1.4002088`.

[146] A. Karimi and M. R. Paul. “Extensive Chaos in the Lorenz-96 Model”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 20.4 (2010).

[147] K. Haven, A. Majda, and R. Abramov. “Quantifying Predictability through Information Theory: Small Sample Estimation in a Non-Gaussian Framework”. In: *Journal of Computational Physics* 206.1 (2005), pp. 334–362.

[148] L. Basnarkov and L. Kocarev. “Forecast Improvement in Lorenz 96 System”. In: *Nonlinear Processes in Geophysics* 19.5 (2012), pp. 569–575.

[149] E. Ott et al. “A Local Ensemble Kalman Filter for Atmospheric Data Assimilation”. In: *Tellus A: Dynamic Meteorology and Oceanography* 56.5 (2004), pp. 415–428.

[150] A. Trevisan and L. Palatella. “On the Kalman Filter Error Covariance Collapse into the Unstable Subspace”. In: *Nonlinear Processes in Geophysics* 18.2 (2011), pp. 243–250.

[151] E. N. Lorenz. “Deterministic Nonperiodic Flow”. In: *Journal of Atmospheric Sciences* 20.2 (1963), pp. 130–141. DOI: `10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2`.

[152] E. N. Lorenz and K. A. Emanuel. “Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model”. In: *Journal of the Atmospheric Sciences* 55.3 (1998), pp. 399–414.

[153] Q. Sun, T. Miyoshi, and S. Richard. “Control Simulation Experiments of Extreme Events with the Lorenz-96 Model”. In: *Nonlinear Processes in Geophysics Discussions* 2022 (2022), pp. 1–18.

[154] C. R. Harris et al. “Array Programming with NumPy”. In: *Nature* 585.7825 (2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

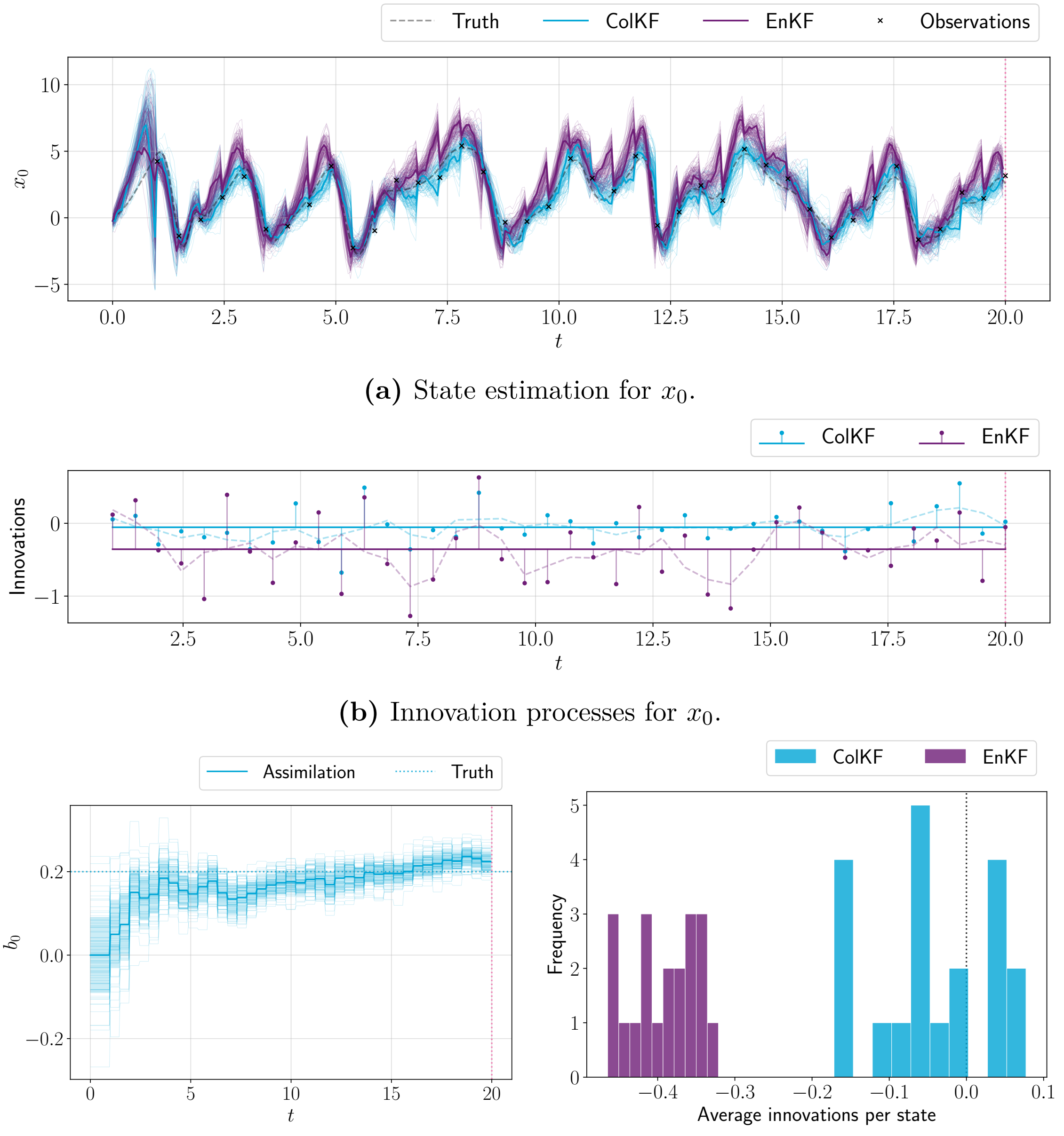
[155] P. Sakov, G. Evensen, and L. Bertino. “Asynchronous Data Assimilation with the EnKF”. In: *Tellus A: Dynamic Meteorology and Oceanography* 62.1 (2010), pp. 24–29. DOI: 10.1111/j.1600-0870.2009.00417.x.

[156] L. Yu et al. “Evaluation of Nonidentical Versus Identical Twin Approaches for Observation Impact Assessments: an Ensemble-Kalman-filter-based Ocean Assimilation Application for the Gulf of Mexico”. In: *Ocean Science* 15.6 (2019), pp. 1801–1814.

[157] J. Kao et al. “Data Assimilation with an Extended Kalman Filter for Impact-produced Shock-wave Dynamics”. In: *Journal of Computational Physics* 196.2 (2004), pp. 705–723.

A

Additional Bias Twin Experiment Results for Lorenz-96



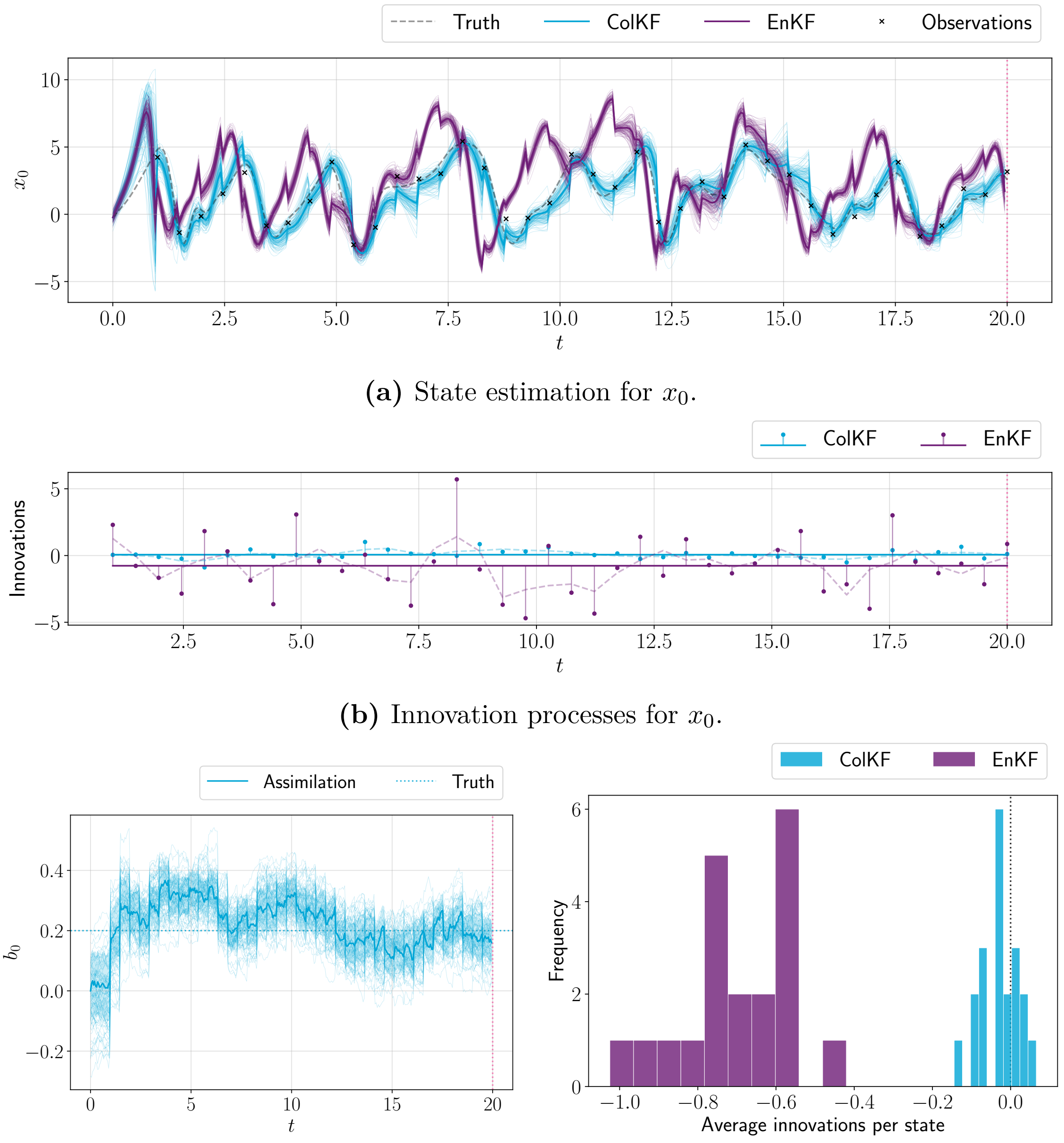
(a) State estimation for x_0.

(b) Innovation processes for x_0.

(c) Estimated AR process (proxy for bias), corresponding to x_0.

(d) Frequency histogram of average innovation for each state.

Figure A.1: Results with the EnKF (purple) and with the ColKF (blue) for the Lorenz-96 model with added constant bias in *Integration-Analysis update* experiment.



(a) State estimation for x_0.

(b) Innovation processes for x_0.

(c) Estimated AR process (proxy for bias), corresponding to x_0.

(d) Frequency histogram of average innovation for each state.

Figure A.2: Results with the EnKF (purple) and with the ColKF (blue) for the Lorenz-96 model with added constant bias in *Analysis-Integration update* experiment.

FACULTY OF ELECTRICAL ENGINEERING, MATHEMATICS AND COMPUTER SCIENCE
Delft University of Technology
Delft, Netherlands
www.tudelft.nl