

Reduced-Order Modeling and Data Assimilation of the El Niño–Southern Oscillation

by

Yusuf Aydogdu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2025

© Yusuf Aydogdu 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Nan Chen
Professor, Dept. of Mathematics, University of Wisconsin – Madison

Supervisor: Navaratnam Sri Namachchivaya
Professor, Dept. of Applied Mathematics, University of Waterloo

Internal Member: Marek Stastna
Professor, Dept. of Applied Mathematics, University of Waterloo

Internal Member: Hans De Sterck
Professor, Dept. of Applied Mathematics, University of Waterloo

Internal-External Member: Eihab Abdel-Rahman
Professor, Dept. of Systems Design Engineering, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Simulations of complex fluid dynamics problems or climate models take weeks to complete even when run parallel in state-of-the-art supercomputers. Given computational resource constraints and the need for adaptable simulation settings, cost-efficient and accurate algorithms are essential. In this thesis, we explore stable, efficient, and accurate methodologies when applied to the El Niño–Southern Oscillation (ENSO), which integrates coupled atmosphere, ocean, and sea surface temperature (SST) mechanisms in the equatorial Pacific. ENSO is one of the most influential and complex climate phenomena, affecting weather patterns across the globe.

ENSO consists of irregular oscillations between warm (El Niño) and cold (La Niña) phases in the Pacific Ocean, significantly impacting global weather patterns. Due to ENSO’s inherent complexity and uncertainties, it is particularly suited for stochastic modeling. By modeling these uncertainties, stochastic simulations offer a more accurate representation of ENSO’s variability, including its irregular periods and amplitudes. We first study the effects of stochastic perturbations on ENSO dynamics and introduce novel modeling and numerical schemes based on the Wiener Chaos Expansion (WCE). The key idea behind WCE is the explicit discretization of white noise through Fourier expansion. We also compare these methods with Monte Carlo (MC) simulations. Our findings demonstrate that the simulation of the linear stochastic ENSO model driven by the Ornstein-Uhlenbeck process using WCE requires far less computational resources and gives more accurate results compared to MC ensembles. This part of the thesis provides an alternative efficient approach for simulations of stochastic climate models and quantification of statistical moments, i.e., mean and variance.

In the next stage of this research, we explore a reduced-order modeling (ROM) framework based on the POD-Galerkin method when applied to a nonlinear ENSO model. POD-Galerkin reduced order modeling aims to reduce the computational complexity and present high-dimensional problems (usually PDEs) with reduced-order equations (ODEs). POD modes are optimal in capturing the system’s dominant features, making it particularly effective for reducing the dimensionality of systems governed by PDEs. By capturing the full-order ENSO (PDE) model with only four modes and four reduced-order equations, we achieve a substantial reduction in computational complexity without significant loss of accuracy. Due to the special properties of the model, we introduce a novel approach using different POD bases, but the same time coefficients for all model components. Moreover, we employ machine learning methods to explore different ROM and model discovery techniques in this part.

The final part of this thesis focuses on the data assimilation of the nonlinear stochastic ENSO model, which forms the core results of this research. We first demonstrate the validity of POD-Galerkin reduced order modeling for the stochastic ENSO driven by the Ornstein-Uhlenbeck process. We project SPDEs onto POD modes derived from the deterministic model and introduce the reduced-order stochastic equations (SDEs). After setting up the filtering framework, we combine these equations and artificial observations in the Pacific ocean, based on realistic experiments, to estimate the ENSO-related SST anomalies. We employ particle filters and test the efficiency using different number of particles and ensembles. From novel ENSO modeling to uncertainty quantification, from reduced order modeling to nonlinear filtering, this thesis provides a promising approach for accurate and efficient predictions of ENSO-related climate variables.

Acknowledgements

I would like to begin by expressing my sincerest gratitude to my supervisor, Prof. Navaratnam Sri Namachchivaya, who is not only an exceptional teacher but also an exemplary person and intellectual. My PhD under the supervision of Prof. Namachchivaya has been a growing experience, both in terms of scientific thinking and interpersonal skills. Working with Prof. Namachchivaya has always been a great pleasure and a deeply enriching learning journey.

I would like to thank Prof. Peter Baxendale for his invaluable help, guidance, and insightful discussions during my PhD. I especially thank Prof. Baxendale for sharing his notes on POD theory with me. I would also like to thank Prof. Marek Stastna, Prof. Hans De Sterck, Prof. Eihab Abdel-Rahman, and Prof. Nan Chen for being members of my committee and for their time spent reading my thesis. Additionally, I would like to thank the University of Waterloo and the Department of Applied Mathematics for providing research and teaching opportunities. From a summer school for high school students at Nesin Mathematics Village in Turkey to the completion of my PhD in Canada, I would like to thank all my teachers who have inspired me.

Words cannot fully express my gratitude to my mother and father. I would like to thank them for their unconditional love, support, encouragement, and for being wonderful parents. I would also like to thank my older brother and older sister for their endless support throughout my life. I would like to thank my nephew and niece for bringing joy to my life. I am always grateful to be a part of my family.

I would like to thank my friends across Europe and North America. I would especially like to thank my friends whom I met in Waterloo. Without their energy, support, friendship, and kindness, this journey would not have been as meaningful. I will definitely miss the chats during long lunches, home gatherings, parties, camping trips, and all the activities we enjoyed together. Finally, I would like to thank my friends from the ODTÜ Alumni in Canada for making being abroad feel like home and making the distance more tolerable for me.

Dedication

This thesis is dedicated to my mother, Şerife, and my father, Tuncay, whose unwavering love and support have been my foundation.

Table of Contents

Examining Committee	ii
Author's Declaration	iii
Abstract	iv
Acknowledgements	v
Dedication	vi
List of Figures	x
List of Tables	xiii
1 Introduction	1
1.1 Climate Models and ENSO	4
1.1.1 Climate Modeling	4
1.1.2 El Niño-Southern Oscillation (ENSO)	5
1.2 Methodologies	6
1.2.1 Stochastic Numerical Methods	7
1.2.2 Reduced Order Modeling	8
1.2.3 Nonlinear Filtering	10
1.3 Thesis Layout	13
2 Modeling and Simulations of the El Niño-Southern Oscillation (ENSO)	14
2.1 Climate Modeling	14
2.2 Equatorial Circulation and the El Niño	15
2.2.1 El Niño and Southern Oscillation	15
2.2.2 Dynamics at Low Latitudes	17
2.2.3 Equatorial Waves	18

2.3	ENSO Modeling	19
2.3.1	A Meridionally Truncated ENSO Model	20
2.4	ENSO Simulations	23
2.5	Conclusion and Future Work	25
3	A Numerical Study of Stochastic El Niño-Southern Oscillation (ENSO) Using Wiener Chaos Expansion	27
3.1	Stochastic Numerical Methods	27
3.2	The El Niño-Southern Oscillation and Stochastic Simulations	28
3.3	Wiener Chaos Expansion	29
3.3.1	Wiener Chaos Expansions of Functionals of Brownian Motions	30
3.4	Simple Ocean Model Driven by White Noise	32
3.4.1	Discontinuities of the Homogeneous Model	33
3.4.2	Mean and Variance Through Characteristic Lines	34
3.4.3	Recursive Propagator Equations	35
3.4.4	Comparison of Theoretical and Numerical Results	37
3.5	Linear Coupled Stochastic ENSO Model	39
3.6	Results	42
3.6.1	Comparison of Wiener Chaos Expansion and Monte Carlo Simulations	44
3.7	Conclusion	47
4	Reduced-Order Modeling (ROM) of the El Niño-Southern Oscillation (ENSO)	48
4.1	Reduced Order Modeling (ROM)	48
4.2	Nonlinear El Niño-Southern Oscillation Model	51
4.3	Proper Orthogonal Decomposition (POD)	53
4.3.1	Proof of Main Theorem	54
4.4	POD-Galerkin ROM Applied to Nonlinear ENSO Model	57
4.4.1	Reconstruction (Accuracy) Results Using POD-Galerkin Reduced Order Modeling	62
4.5	Neural Networks	65
4.5.1	Autoencoders	66
4.6	Sparse Identification of Nonlinear Dynamics (SINDy)	67
4.6.1	Reconstruction (Accuracy) Results Using SINDy Reduced Order Modeling	68
4.7	Conclusion and Future Work	68

5	Data Assimilation of the El Niño-Southern Oscillation (ENSO)	70
5.1	Data Assimilation of the El Niño-Southern Oscillation (ENSO)	70
5.2	Stochastic Nonlinear ENSO Model	71
5.3	Reduced Order Modeling of Stochastic Nonlinear ENSO Model	72
5.4	Nonlinear Filtering Theory	76
5.4.1	Nonlinear Filtering Problem	77
5.4.2	Particle Filter	78
5.5	Prediction of ENSO-Related SST Using a Particle Filter	80
5.6	Conclusion and Future Directions	85
6	Conclusion and Future Work	86
6.1	Conclusions and Future Work	86
6.1.1	Concluding Remarks	86
6.1.2	Future Work	87
	References	90
	APPENDICES	98
A	A Complete El Niño-Southern Oscillation Model	99
A.1	A Complete Coupled Model	99
A.2	Parabolic Cylinder Functions	100
B	Hermite Polynomials and Wiener Chaos Expansion	101
B.1	Hermite Polynomials	101
C	Reduced-Order Equations of the El Niño-Southern Oscillation Model	104
C.1	Invariance of the POD Modes	104
C.2	Reduced Order Equations	105
C.2.1	POD-Galerkin Reduced Order Equations	105
C.2.2	POD-SINDy Reduced Order Equations	106
C.2.3	Autoencoder-SINDy Reduced Order Equations	107
C.3	Hopf Bifurcation	109
D	Nonlinear Filtering of the El Niño-Southern Oscillation (ENSO)	110
D.1	Validation of POD-Galerkin ROMs of the Stochastic ENSO Model	110
D.2	Particle Filter Algorithm	113
D.3	A Spectral Approach to Nonlinear Filtering : Revisiting Wiener Chaos Expansion .	114
D.3.1	Recursive Equations	115

List of Figures

1.1	Thesis flowchart.	13
2.1	The sea-surface temperature anomalies during a strong El Niño year (December, 1997) [156].	16
2.2	Upper: SST time series in the Eastern Equatorial Pacific (Niño 3) region. Lower: Negative of the Southern Oscillation Index (SOI), the anomalous pressure difference between Tahiti and Darwin [156].	16
2.3	Sketch of the key members of the ENSO cycle [156].	18
2.4	Oceanic waves induced by a wind anomaly on the equator [156].	19
2.5	Zonal thermocline feedback profile $\eta(x)$	22
2.6	Simulation of the equatorial atmospheric Kelvin and Rossby waves, oceanic Kelvin and Rossby waves and sea surface temperature.	24
3.1	Initial conditions for Kelvin and Rossby waves at time $t = 0$ that satisfy the boundary conditions (left), Kelvin wave at time $t = L/2c$ (middle) and Rossby wave at time $t = 2L/c$ (right).	33
3.2	Kelvin and Rossby waves through characteristics.	34
3.3	Mean and variance of Kelvin and Rossby waves over time (left) and at the terminal time $T = 4L/c$ (right).	38
3.4	Random field solutions of oceanic Kelvin and Rossby waves generated using different stochastic bases.	39
3.5	Deterministic and stochastic wind activities (left) and the noise profile confined to the western Pacific (right).	41
3.6	Response of the SST to different atmosphere-ocean coupling κ at the mid-ocean. . .	41
3.7	Variance of OU-process obtained from WCE with different number of modes (with $\sigma = 1.2$, $d_p = 1.0$ and $\hat{\tau} = -0.25$) (left) and $L1$ errors relative to the exact (analytical) solution (right).	42
3.8	Comparison of WCE and MC variance results of $T(t, x)$ with different number of ensembles at a fixed space point $x = 0.6$	45
3.9	(a) Mean of oceanic Kelvin wave $K^O(t, x)$ obtained from WCE (1st propagator), MC (300 ensembles), and the corresponding $L1$ error. (b) Variance of oceanic Kelvin wave obtained from WCE (50 propagators), MC (300 ensembles).	46

3.10	(a) Mean of oceanic Rossby wave $R^O(t, x)$ obtained from WCE (1st propagator), MC (300 ensembles), and the corresponding $L1$ error. (b) Variance of oceanic Rossby wave obtained from WCE (50 propagators), MC (300 ensembles).	46
3.11	(a) Mean of sea surface temperature (SST) $T(t, x)$ obtained from WCE (1st propagator), MC (300 ensembles), and the corresponding $L1$ error. (b) Variance of sea surface temperature (SST) obtained from WCE (50 propagators), MC (300 ensembles).	46
4.1	(a) POD modes $\tilde{\Psi}(x)$ and (b) projected initial conditions used to simulate the reduced-order equations.	60
4.2	Comparison of initial conditions of PDE $\langle(u_3(x, 0), u_4(x, 0), u_5(x, 0))\rangle$ and reconstructed initial conditions using four POD modes $\langle(\bar{u}_3(x, 0), \bar{u}_4(x, 0), \bar{u}_5(x, 0))\rangle$	61
4.3	Comparison of POD time coefficients $q_\alpha(t)$ and simulations of reduced order equations $a_\alpha(t)$ for $\alpha = 1, \dots, 4$	62
4.4	Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of atmospheric Kelvin $K^A(t, x)$ together with reconstruction (L1) errors.	63
4.5	Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of atmospheric Rossby $R^A(t, x)$ together with reconstruction (L1) errors.	64
4.6	Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of oceanic Kelvin $K^O(t, x)$ together with reconstruction (L1) errors.	64
4.7	Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of oceanic Rossby $R^O(t, x)$ together with reconstruction (L1) errors.	64
4.8	Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of SST $T(t, x)$ together with reconstruction (L1) errors.	65
4.9	Autoencoder architecture for ENSO data with $\tanh()$ activations function	66
5.1	Ensembles of Ornstein-Uhlenbeck process $\tau(t, \omega)$ (ten ensembles).	72
5.2	POD modes of deterministic model (left) and noise profile $\epsilon\left(\chi_O \frac{c\kappa}{2} s_p(x), -\chi_O \frac{c\kappa}{3} s_p(x), 0\right)$ as a function of space (right).	73
5.3	Scaled noise amplitude ensembles of the Ornstein-Uhlenbeck process with σ values.	74
5.4	Comparison of POD time coefficients $q^S(t, \omega)$ and simulation of stochastic reduced order equations $a^S(t, \omega)$	74
5.5	Stochastic full order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of SST $T(t, x)$, each obtained using the average of ten ensemble simulations and four modes.	75
5.6	L1 errors between POD and POD-G reconstructions, each obtained using the average of ten ensemble simulations.	75
5.7	The big picture for DA methods and algorithms [2].	76
5.8	Sequential assimilation [2].	77
5.9	Evolution of signal density using particles [169].	79
5.10	Four (artificial) sensors located on the SST modes $(\psi_1^S(x), \psi_2^S(x), \psi_3^S(x), \psi_4^S(x))$	81

5.11	Four (artificial) sensors located between Australia (west) and South America (east) for ENSO-related SST prediction.	81
5.12	Average of truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 2 particles and 10 ensembles. . .	84
5.13	Average of truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 10 particles and 10 ensembles. . .	84
5.14	Average of truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 100 particles and 10 ensembles. .	84
A.1	Meridional profiles of atmosphere ($\phi_{0,2}(y)$) and ocean ($\Psi_{0,2}(Y)$) parabolic cylinder functions.	100
C.1	POD modes $\hat{\Psi}_\alpha(x)$ for $\alpha = 1, \dots, 4$ obtained from simulations with different initial conditions.	104
C.2	POD time coefficients $q_\alpha(t)$ for $\alpha = 1, \dots, 4$ obtained from simulations with different initial conditions.	105
C.3	Comparison of POD time coefficients $q_\alpha(t)$ and the simulation of the POD-SINDy equations $s_\alpha(t)$	107
C.4	Comparison of autoencoder latent space $r_\alpha(t)$ and the simulation of the AE-SINDy equations $s_\alpha(t)$	108
C.5	Dynamics of reduced-order model (C.1) as $t \rightarrow \infty$	109
D.1	Ten ensembles of POD time coefficients $q_\alpha^s(t)$ using Equation (5.8).	110
D.2	Ten ensembles of POD-Galerkin reduced-order stochastic equations $a_\alpha^s(t)$ using Equations (5.9)-(5.10).	111
D.3	First three ensembles of the SST component of full-order stochastic (5.1) (left), POD reconstruction of full-order model using $q_\alpha^s(t, \omega)$ and $\Psi_\alpha(x)$ (middle) and L1 error between them (right).	112
D.4	First three ensembles of the SST component of full-order stochastic (5.1) (left), POD reconstruction of full-order model using $a_\alpha^s(t)$ and $\Psi_\alpha(x)$ (middle) and L1 error between them (right).	113
D.5	Evolution of probability densities over time.	116

List of Tables

1.1	Hierarchy of ENSO Models	5
1.2	Methodologies.	6
2.1	Numerical scheme for the the El Niño-Southern Oscillation (ENSO).	24
2.2	Descriptions of parameters in the operator $\mathcal{A}(\kappa)$ and nonlinearity $\mathcal{N}(u, u, \mu)$	25
3.1	Gaussian random variables multiplied by the corresponding propagator coefficients for different sets.	38
3.2	Numerical scheme for the solution of deterministic propagators combined with the appropriate random basis to obtain the WCE solution of SPDEs.	43
3.3	Comparison of relative mean and variance errors for MC ensembles and WCE method for sea surface temperature $T(t, x)$, including computation times on a common laptop.	45
4.1	Reconstruction errors of model PDEs (root mean square) K^A, R^A, K^O, R^O, T , using different number of modes with $\alpha = 1, \dots, 4$	61
4.2	Root mean square error between FOM and used methods for K^O, R^O , and T using four modes.	68
5.1	Values of projected noise intensities σ	73
5.2	Root mean square (RMSE) between truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 2 particles where α is the corresponding mode and N is the ensemble.	82
5.3	Root mean square (RMSE) between truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 10 particles where α is the corresponding mode and N is the ensemble.	83
5.4	Root mean square (RMSE) between truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 100 particles where α is the corresponding mode and N is the ensemble.	83
C.1	Coefficients of linear and quadratic terms in the system of the POD-SINDy equations.	106
C.2	Coefficients of linear and quadratic terms in the system of the AE-SINDy equations.	108
D.1	Steps of the Particle Filter Algorithm.	114

Chapter 1

Introduction

Numerical climate prediction (NCP) refers to the simulation of the Earth’s climate system through mathematical models and computational methods to forecast climate behavior over various timescales, from months to decades or even centuries. These simulations are critical for understanding and managing the impacts of climate variability on natural and human systems. By simulating the Earth’s climate system, NCP provides insights into long-term trends in temperature, precipitation, and extreme weather events. These insights can inform policy decisions on climate adaptation and mitigation, support sustainable development, and guide the design of infrastructure resilient to future climate conditions. As a result, NCP can help assess risks associated with phenomena like tropical cyclones and polar ice melt, which have global socioeconomic and environmental consequences.

NCP is also critical for advancing our understanding of complex feedback mechanisms within the Earth’s climate system. These models help track the interactions between various climate components, such as the atmosphere, oceans, ice sheets, and vegetation, providing a comprehensive view of how changes in one system can influence others. By improving the accuracy of these simulations, NCP plays a crucial role in predicting the potential impacts of climate change on biodiversity, food security, and water resources. Furthermore, NCP is instrumental in assessing the effectiveness of climate interventions, such as carbon sequestration and geoengineering strategies, enabling informed decisions on how to mitigate adverse climate outcomes [143]. As climate challenges become more pressing, the ability to make well-informed predictions will be indispensable for safeguarding the planet’s ecosystems and human societies.

NCP on a global scale provides significant advantages in understanding and forecasting climate variability and change. High-resolution models substantially enhance the representation of key processes such as convection, ocean-atmosphere interactions, and the hydrological cycle, leading to improved accuracy in predicting extreme weather events and long-term climate trends [136]. These models have been significantly improved by advances in computational power and data assimilation, allowing forecasts today to achieve a level of accuracy that was unreachable decades ago [60]. Large-ensemble simulations further guide climate risk assessments by providing probabilistic insights into rare and effective events [115]. Moreover, the integration of NCP and artificial intelligence promises to further increase forecast efficiency and skill by leveraging vast datasets and complex pattern recognition [144].

While predictions remain central to NCP, the uncertainty surrounding them, particularly due to incomplete data from remote regions, such as the deep ocean and polar ice caps, limits their reliability. Moreover, the uncertainty in initial and boundary conditions, which particularly stems

from such incomplete observational data and the inherent long-term variability of the climate system, makes prediction more challenging. Additionally, integrating diverse observational datasets into climate models through data assimilation is computationally demanding, especially when considering complex feedback mechanisms [161]. The inherent unpredictability of the climate system complicates distinguishing natural variability from anthropogenic influences [93].

While uncertainty in initial and boundary conditions challenges model accuracy, additional limitations arise from the models themselves, particularly in simulating small-scale processes and coupling system components. Current models often struggle to adequately simulate small-scale processes, such as cloud dynamics, ocean-atmosphere interactions, and ice-sheet behavior, which are critical for projecting future climates. These processes are typically parameterized due to computational constraints, introducing biases into simulations [8]. Furthermore, coupling different components of the Earth system, such as the atmosphere, oceans, and biosphere, remains a computationally intensive task, requiring vast resources and advanced algorithms. Addressing these challenges calls for increased investment in high-performance computing, the expansion of global observational networks, and the development of more sophisticated numerical methods [70].

These model limitations are compounded by the high dimensionality of climate systems, which require models to account for numerous variables across fine temporal and spatial scales. Climate models must represent interactions across the atmosphere, oceans, land surface, and cryosphere, resulting in a vast number of variables. Solving the governing equations for these variables requires substantial computational resources, especially as higher resolutions demand smaller time steps to maintain stability. Moreover, the representation of small-scale processes, such as cloud microphysics, ocean eddies, and ice dynamics, often relies on parameterizations, which introduce biases and limit model accuracy [134]. The sheer size of the data generated by these models also complicates storage, retrieval, and analysis, making it challenging to extract actionable insights [143].

Data assimilation and uncertainty quantification in high-dimensional climate systems further amplify these challenges. Integrating diverse observational datasets from satellites, ground stations, and other sources into large-scale models requires advanced data assimilation techniques, like 4D-Var or ensemble Kalman filters, which struggle to scale effectively with increasing dimensions. Similarly, ensemble-based methods used to assess uncertainty involve running multiple high-dimensional simulations, significantly increasing computational demands [125]. Additionally, the chaotic nature of the climate system complicates the separation of natural variability from anthropogenic forcing, challenging the accuracy of long-term projections.

To address these challenges, the development of fast algorithms and reduced-order models (ROMs) has become crucial. These methods aim to reduce computational costs while maintaining the accuracy of simulations, which is vital given the high-dimensional nature of the climate system. As climate models become more complex and require higher resolutions to capture fine-scale processes, the computational demand increases significantly, making long-term simulations and frequent model updates impractical. Fast algorithms, such as those based on parallel computing and advanced numerical techniques, enable the efficient processing of large datasets and the rapid execution of simulations. Reduced-order models (ROMs) offer an effective approach to mitigate the computational burden by simplifying the system while retaining its essential dynamics. By focusing on the most influential modes and reducing the complexity of the model, ROMs maintain accuracy while dramatically reducing the number of calculations required. These advances are essential for scaling NCP to the global level, providing timely predictions, and enabling more frequent updates that are critical for climate monitoring, policy-making, and adaptation strategies. The integration of fast algorithms and ROMs into climate models facilitates more efficient exploration of future

climate scenarios, ultimately enhancing our ability to mitigate and adapt to the impacts of climate change.

This thesis aims to enhance climate prediction by developing efficient, stable, and accurate algorithms, focusing specifically on improving predictions for the El Niño-Southern Oscillation (ENSO), a critical climate phenomenon influencing global weather patterns. Our work consists of four main components. First, we employ the ENSO model developed by Majda and collaborators [23, 155] and propose a slightly novel variation by adjusting the boundary conditions. Unlike the original model, our proposed model is confined to the equatorial Pacific Ocean rather than the entire equatorial belt.

Building on this new modeling formulation, we use Wiener Chaos Expansion (WCE) and Monte Carlo (MC) ensemble approaches to simulate a linear stochastic ENSO model as the second step. In this part, we demonstrate that the WCE method can be more efficient and accurate compared to the MC technique. However, applying the WCE method to a nonlinear stochastic ENSO model introduces significant challenges. WCE suffers from rapid growth in the number of terms required to represent solutions as the dimensionality of the system increases, making computations expensive and complex. As a result of nonlinear interactions, non-Gaussian dynamics in stochastic models lead to intricate coupling between chaos coefficients, complicating truncation strategies and introducing errors in the approximation of higher-order terms. These challenges in the nonlinear model prompted us to explore more efficient methods.

To address the high dimensionality and nonlinearity of the problem, we employed reduced-order modeling techniques based on Proper Orthogonal Decomposition (POD) combined with Galerkin projection in the fourth chapter of this thesis, which constitutes the core of our research. These reduction techniques allows us to approximate the high-dimensional nonlinear ENSO model using a small number of reduced-order equations, enabling efficient, accurate, and cost-effective computations. In addition to POD-Galerkin reduction, we explore neural network and sparse identification algorithms to investigate alternative reduced-order approaches. In the fifth chapter of this thesis, we apply nonlinear filtering algorithms to predict sea surface temperatures using reduced-order models. This work represents an initial and promising step toward achieving accurate predictions for real-world systems.

The aim of this chapter is to provide a comprehensive overview of the thesis by outlining its key components. We begin by exploring various Earth climate systems in Section (1.1), narrowing our focus to the El Niño-Southern Oscillation (ENSO) in Section (1.1.2), which serves as the experimental foundation of this research. Section (1.2) highlights the methodologies developed to enable efficient and accurate simulations of ENSO. In Section (1.2.1), we delve into stochastic numerical techniques, including the Wiener Chaos Expansion (WCE) and Monte Carlo simulations. This is followed by reduced-order modeling strategies in Section (1.2.2), with an emphasis on the POD-Galerkin method and complementary data-driven approaches. Finally, Section (1.2.3) focuses on nonlinear filtering methods, which constitute the central objective of this research. We also provide the thesis layout scheme in Section (1.3). The following chapters will detail the methodologies employed to achieve these improvements, beginning with an exploration of the ENSO model and progressing through the advanced techniques developed in this research.

1.1 Climate Models and ENSO

1.1.1 Climate Modeling

Climate modeling involves using mathematical and computational techniques to simulate the behavior of Earth’s climate system. These models are designed to capture the interactions among key components of the climate—such as the atmosphere, oceans, land surface, and ice—to study past, present, and future climate conditions. By incorporating physical laws governing phenomena like fluid dynamics, thermodynamics, and radiative transfer, climate models enable scientists to predict changes in temperature, precipitation, wind patterns, and other climate variables. They are essential tools for understanding how natural factors (e.g., volcanic eruptions) and human activities (e.g., greenhouse gas emissions) influence the climate across different time scales, from seasons to centuries [119].

Climate models vary in complexity. For instance, Energy Balance Models (EBMs) are simplified models that estimate global temperatures by focusing on the balance between incoming solar radiation and outgoing heat. More advanced models like General Circulation Models (GCMs) simulate atmospheric and oceanic circulation, capturing the intricate flow of air and water across the globe. The most comprehensive models, known as Earth System Models (ESMs), integrate additional components like the biosphere and carbon cycle [107]. These climate models are used to predict future climate conditions, understand the role of human activities in climate change, and assess risks to ecosystems, agriculture, and society. Despite uncertainties, such models have become essential tools in addressing climate change, and are continuously refined by comparing their outputs with historical data and observations.

There are several major climate phenomena that significantly influence weather patterns and global climate. These events involve interactions between the ocean and atmosphere and can lead to shifts in temperature, precipitation, and storm activity worldwide. For example, The North Atlantic Oscillation (NAO) is a climate phenomenon that describes fluctuations in atmospheric pressure differences between the Icelandic Low and the Azores High, two semi-permanent pressure systems in the North Atlantic Ocean. These pressure variations significantly affect the weather patterns across North America, Europe, and parts of North Africa, especially during the winter months [58, 59]. Moreover, the Madden-Julian Oscillation (MJO) is a tropical disturbance that travels eastward around the globe and affects atmospheric circulation, particularly in the Indian and Pacific Oceans [99, 172]. It influences monsoon activity and tropical storm development. In this thesis, we study a natural climate phenomenon characterized by periodic variations in sea surface temperatures and atmospheric pressure in the tropical Pacific Ocean: the El Niño-Southern Oscillation (ENSO).

The hierarchy of ENSO models ranges from simple conceptual frameworks to complex coupled general circulation models (GCMs) as shown in Table (1.1). At the basic level, conceptual models like the delayed oscillator or recharge-discharge models capture key feedback mechanisms with minimal complexity [151]. Intermediate complexity models, such as intermediate coupled models (ICMs), include simplified ocean and atmospheric dynamics to balance realism and computational efficiency [171]. At the highest level, fully coupled GCMs simulate detailed interactions between the ocean, atmosphere, and land surface, providing the most comprehensive—but computationally intensive—representations of ENSO dynamics [83].

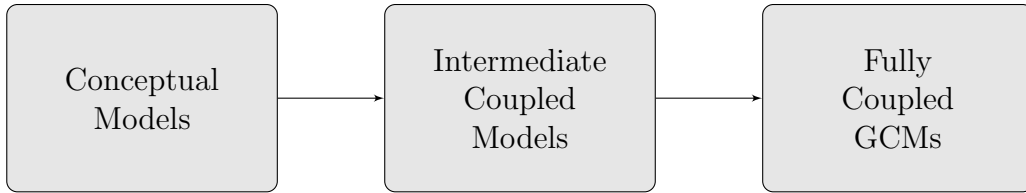


Table 1.1: Hierarchy of ENSO Models

1.1.2 El Niño-Southern Oscillation (ENSO)

The El Niño-Southern Oscillation (ENSO) is a recurring and intricate climate phenomenon originating in the tropical Pacific Ocean. It involves the dynamic interplay between the ocean and atmosphere, causing periodic fluctuations in sea surface temperatures (SSTs). These changes influence atmospheric conditions over the equatorial Pacific and other low-latitude regions. ENSO significantly impacts global weather patterns, making its accurate modeling and simulation crucial for climate prediction.

Initial observations of climatic anomalies in low-latitude regions were made by Bjerknes [12] in the late 1960s, who noted the correlation between sea surface temperature (SST) and equatorial atmospheric circulations. These findings established the link between El Niño and the Southern Oscillation. Building on empirical observations, Matsuno and Gill [42] used simplified one-layer shallow-water models with damping to provide analytical solutions, representing the atmospheric response through equatorial Kelvin and Rossby waves. Since the 1970s, experimental forecasts of ENSO have been conducted. Among the intermediate atmosphere-ocean coupled models, the Cane-Zebiak model [171] has been particularly influential, accurately capturing the key dynamics of the coupled system. Additionally, Neelin, Dijkstra, and their colleagues proposed a comprehensive deterministic ENSO theory [29, 120], thoroughly analyzing the system components.

Another important contribution to ENSO studies was made by Fei Fei Jin, who introduced the influential recharge-discharge oscillator conceptual model that highlights the basin-scale heat content buildup as central to the oscillatory behavior of ENSO [66]. Jin also provided a comprehensive synthesis of ENSO modeling that unifies delayed oscillator and western Pacific oscillator while accounting both internal dynamics and external stochastic forcing [67].

ENSO is driven by complex ocean-atmosphere interactions, where changes in SST significantly impact atmospheric pressure and wind patterns, which in turn influence oceanic currents and upwelling. A key mechanism is the Bjerknes feedback, a positive loop where warm SST anomalies in the eastern Pacific weaken trade winds, further reducing upwelling and enhancing warming. The depth and tilt of the thermocline, a layer with a steep temperature gradient, play crucial roles in the development of ENSO events, as they affect the distribution of SST anomalies [121]. Additionally, the propagation of equatorial Kelvin and Rossby waves spreads these anomalies across the equatorial Pacific, influencing the onset and decay of ENSO events. These interconnected processes make ENSO a highly dynamic and complex phenomenon, challenging to model and predict accurately.

In this thesis, we study the nonlinear ENSO model developed by Majda and colleagues [23, 155] (Majda’s model), where the ocean and atmosphere are both modeled as non-dissipative linear shallow water equations, with a dissipative nonlinear thermodynamics budget. Similar to the Zebiak-Cane model, it is an anomaly model, where the anomalies are calculated relative to a mean annual cycle specified from observations, and thus all state variables in subsequent sections are

anomalies from an equilibrium state and are nondimensional. Building on the foundation of the Cane-Zebiak model, Majda’s model addresses the limitations of linearity and determinism, making it a more versatile tool for studying the complexities of ENSO phenomena. This model provides a more realistic framework for studying ENSO variability, extremes, and predictability in the presence of noise. Such models can then be further simplified by meridional truncation that reduces the dimensionality of the problem by approximating the solution in the meridional (north-south) direction using a predefined set of modes. The model introduced by Majda and co-workers in [23, 155] are defined as:

$$\mathbf{U} \stackrel{\text{def}}{=} [u(t, x, y), v(t, x, y), \theta(t, x, y), U(t, x, Y), V(t, x, Y), H(t, x, Y), T(t, x, y)]$$

The differential form, model components, boundary conditions, and modeling parameters are provided in the Appendix (A).

It is shown that the ENSO model can be projected onto parabolic cylinder functions, resulting in nonlinear coupled PDEs in five variables $\mathbf{u}(t, x) \stackrel{\text{def}}{=} [K^A(t, x), R^A(t, x), K^O(t, x), R^O(t, x), T(t, x)]^T$ that can be expressed in differential equation form as

$$\mathcal{M} \frac{\partial \mathbf{u}}{\partial t}(t, x) = \mathcal{A}(\kappa) \mathbf{u}(t, x) + \mathcal{N}(\mathbf{u}, \mathbf{u}, \mu)(t, x) + \dot{W}(t) \quad (1.1)$$

where $\mathcal{M}$ is a 5×5 constant matrix with diagonals $[0, 0, 1, 1, 1]$, the linear operator $\mathcal{A}(\kappa)$ is a linear operator dependent on the wind stress coefficient κ and nonlinearity $\mathcal{N}(\mathbf{u}, \mathbf{u}, \mu)$ with advection coefficient μ , respectively. The dissipative steady atmosphere model is driven by atmospheric Kelvin $K^A(t, x)$ and Rossby $R^A(t, x)$ waves, respectively. Moreover, a dissipative shallow water ocean model forced by atmospheric waves is composed of oceanic Kelvin $K^O(t, x)$ and Rossby $R^O(t, x)$ waves. The final component of the coupled model is the sea surface temperature (SST) model $T(t, x)$ that characterizes the temperature anomalies. The randomness $W(t)$ will be discussed in more detail in the thesis. The equation (1.1) serves as the central component of this work.

1.2 Methodologies

In this thesis, we employ numerical and reduced-order modeling techniques to analyze the deterministic and stochastic versions of Equation (1.1). The techniques used in this work fall into three main categories. First, stochastic numerical methods include Wiener chaos expansion (WCE) and Monte Carlo simulations. Second, reduced-order modeling approaches involve POD-Galerkin reduction and neural networks combined with sparse identification methods. Finally, nonlinear filtering methods include WCE and particle filtering. Table (1.2) summarizes the methods used in this work.

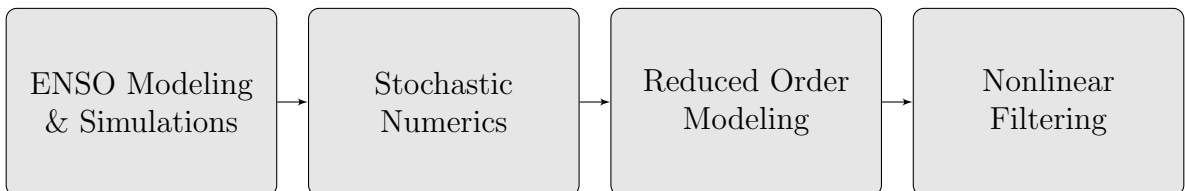


Table 1.2: Methodologies.

1.2.1 Stochastic Numerical Methods

Stochastic numerics is a subfield of numerical analysis that focuses on developing and applying numerical methods to solve problems involving randomness or uncertainty. Such problems are often modeled by stochastic processes, stochastic differential equations (SDEs), stochastic partial differential equations (SPDEs), or random fields. The primary challenge in stochastic numerics is to accurately and efficiently simulate, approximate, and analyze systems where the behavior is influenced by random factors, often leading to non-deterministic outcomes. SPDEs generalize partial differential equations (PDEs) by including stochastic terms, which may account for random inputs, boundary conditions, or coefficients. These equations are used in fields like fluid dynamics, climate modeling, economics, astrophysics and quantum mechanics where randomness is inherent in the system being modeled [50].

The development of stochastic numerics has been shaped by the evolution of probability theory, stochastic processes, and numerical analysis. Its history is marked by key contributions from mathematicians, physicists, and engineers, as well as the growing importance of computational methods in science and industry. The foundations of stochastic numerics lie in the development of probability theory, which began in the 17th century with the work of Blaise Pascal and Pierre de Fermat on problems of gambling and chance. This was further formalized by figures like Jakob Bernoulli (law of large numbers) and Pierre-Simon Laplace (central limit theorem) [82] in the 18th century. The concept of Brownian motion, observed by botanist Robert Brown in 1827, laid the groundwork for stochastic processes. Albert Einstein’s mathematical description of Brownian motion in 1905 [36], and later Norbert Wiener’s formalization of it as a stochastic process, were crucial in the development of stochastic differential equations (SDEs) [163, 165].

The development of SDEs began in the early 20th century with the work of Andrey Kolmogorov, who formulated the Kolmogorov forward and backward equations [77], and Norbert Wiener, who provided a rigorous mathematical foundation for Brownian motion. Kiyoshi Itô further advanced this field in the 1940s by developing Itô calculus [64], which became essential for solving SDEs. In the 1960s and 1970s, with the rise of digital computing, researchers began developing numerical methods to solve SDEs. The Euler-Maruyama method, an extension of the Euler method for ODEs to stochastic cases, became one of the first and most widely used techniques [105].

Despite many achievements, stochastic numerics often deals with high-dimensional spaces, where traditional numerical methods can suffer from the “curse of dimensionality”. Developing efficient algorithms that scale well with dimension is an ongoing challenge. Techniques like variance reduction, adaptive methods, and parallel computing are crucial in addressing these challenges.

In this thesis, we are interested in a class of stochastic partial differential equations (SPDEs)

$$u_t = \mathcal{L}(u) + \sigma \dot{W}(t) \tag{1.2}$$

where $\mathcal{L}(u)$ is a general linear or nonlinear differential operator in spatial variables, and $W(t)$ is a Brownian motion. The solution of (1.2) not only depends on the spatial-time variable (t, x) but also the Brownian motion path.

To solve Equation (1.2), we focus on two methods of stochastic numerics: Wiener Chaos Expansion (WCE) and Monte Carlo (MC) methods. WCE is a technique used to represent random variables or processes as an infinite series of orthogonal polynomials (Hermite polynomials) with respect to a Gaussian measure. This method is particularly useful in the context of solving SDEs or analyzing systems with Gaussian random inputs. The expansion is based on the Wiener-Itô

theorem, which states that any square-integrable random variable can be decomposed into an orthogonal series of Hermite polynomials of Gaussian random variables [17, 71, 94, 95, 98]. Each term in the series represents a different “order of chaos”, with the first term being deterministic and higher-order terms capturing increasing levels of randomness. WCE is powerful for uncertainty quantification, sensitivity analysis, and solving stochastic partial differential equations (SPDEs) by transforming a stochastic problem into a deterministic one in the space of the coefficients of the expansion. This transformation is represented as follows:

$$u(t, x, W^t) = \sum_{\alpha \in \mathcal{J}} u_{\alpha}(t, x) V_{\alpha}(\xi) \quad (1.3)$$

Here, $V_{\alpha}(\xi)$ are multi-variable Hermite polynomials of Gaussian random variables. The functions $u_{\alpha}(t, x)$ represent deterministic Hermite-Fourier coefficients in the Cameron-Martin orthogonal decomposition of $u(t, x, W^t)$. The Gaussian random variables are constructed by discretizing the Brownian motion W^t as a Fourier expansion with respect to a set of deterministic basis functions in $L^2([0, t])$. We revisit the WCE method for nonlinear filtering calculations in the final stage of this work.

On the other hand, MC methods involve simulating a large number of random samples to estimate expectations, integrals, or the solutions to stochastic problems. They are widely used due to their general applicability, particularly for high-dimensional problems. For instance, the mean of the random solution is estimated using the MC ensemble average:

$$u_M^{MC}(t, x) = \frac{1}{M} \sum_{k=1}^M u(t, x, \omega_k) \quad (1.4)$$

where $u(t, x, \omega_k)$ is a realization computed by the MC simulations, ω_k is the sample path of Brownian motion and M is the total number of realizations. In this thesis, we show that the WCE method can be more efficient than the MC method for the linear stochastic ENSO model. However, when studying the nonlinear model, we needed reduced-order modeling techniques, which led to the next stage of our work.

1.2.2 Reduced Order Modeling

In the second part of the methodology, we focus on reduced-order modeling (ROM), aiming to simplify high-dimensional systems while preserving their key dynamics and essential features. High-dimensional systems, such as those governed by partial differential equations (PDEs) or coupled ordinary differential equations (ODEs), often involve a large number of variables. Therefore, it is important to explore low-dimensional systems that effectively represent these high-dimensional models. To achieve this, reduced order modeling (ROM) methods have been an important tool.

ROM is a mathematical and computational strategy aimed at simplifying complex, high-dimensional models by reducing the number of variables and equations while maintaining the essential characteristics and accuracy of the original system. ROM is particularly valuable in fields where computational resources are limited, and the full model is too complex or time-consuming to solve directly. It seeks to identify a lower-dimensional subspace where the most critical dynamics of the system occur. Once the key modes are identified, the original governing equations are projected onto the reduced space spanned by these modes. This process results in a set of reduced-order equations

that are much simpler to solve compared to the original full-order model.

In this thesis, we study the proper orthogonal decomposition (POD) method alongside Galerkin projection. POD, also known as Karhunen-Loève decomposition in the stochastic context in the 1960s and 1970s [72, 92], emerged as a powerful method for identifying dominant patterns in data. POD was originally used in the analysis of turbulent flows, where it helped to capture the most energetic structures in fluid dynamics [97, 147]. This laid the groundwork for applying ROM in various engineering fields [27, 111].

POD involves decomposing the original data or solution fields into orthogonal modes, ranked by their significance (e.g., energy content). The top modes are retained to form the reduced basis, and the remaining, less significant modes are discarded. This method ensures that the residual (the difference between the full and reduced models) is minimized in the least-squares sense. The vector field $u(t, x)$ can be approximated by spatial function $\psi_\alpha(x)$ and time coefficients $a_\alpha(t)$ such that:

$$u(t, x) \approx u_n(t, x) = \sum_{\alpha=1}^n a_\alpha(t) \psi_\alpha(x) \quad (1.5)$$

Along with the POD, Galerkin projection is used to derive ROM equations, which are typically systems of ordinary differential equations (ODEs). Once the POD decomposition is applied to the full-order model, the governing equations are projected onto the reduced subspace spanned by the retained POD modes. This involves substituting the POD expansion of the solution into the original partial differential equations (PDEs) and performing an inner product with each basis function to obtain the reduced dynamics. The resulting equations describe the temporal evolution of the coefficients associated with the POD modes and are significantly lower in dimension than the original system, making them computationally efficient while retaining the essential dynamics of the full-order model. After applying POD-Galerkin, the reduced order equations have the following form:

$$\dot{\hat{a}}_\gamma(t) = \sum_{\alpha=1}^n \hat{a}_\alpha(t) \langle \mathcal{A} \psi_\alpha, \psi_\gamma \rangle + \sum_{\alpha, \beta=1}^n \hat{a}_\alpha(t) \hat{a}_\beta(t) \langle \mathcal{N}(\psi_\alpha, \psi_\beta), \psi_\gamma \rangle \quad \gamma = 1, \dots, n \quad (1.6)$$

A key challenge in POD-Galerkin is balancing accuracy with computational efficiency. Reducing the model's order may introduce errors or loss of fidelity compared to the original model. Ensuring that the reduced model accurately represents the system's essential dynamics requires careful validation. Applying POD-Galerkin to nonlinear systems can be particularly challenging. Nonlinearities may not be well-captured by linear ROM techniques, requiring more sophisticated methods like nonlinear model reduction or machine learning-based approaches. In this work, we show that full order nonlinear (both deterministic and stochastic) ENSO models can be efficiently and accurately approximated using few POD modes and reduced order equations. Such equations will be used as the dynamical models for nonlinear filtering, which serves as the final component of this thesis.

In addition, we also utilize neural networks (autoencoders) for ROM purposes. Autoencoders play a crucial role in modern machine learning and have a wide range of applications, particularly in scenarios where data dimensionality needs to be reduced or where feature learning is essential [53, 157]. They are a type of artificial neural network designed to learn efficient representations of data, typically for the purpose of dimensionality reduction, denoising, or feature learning. The key idea behind an autoencoder is to compress input data into a lower-dimensional form (encoding) and then reconstruct the data from this compressed form (decoding). The encoder is a function that maps the input data $\mathbf{x}$ to a lower-dimensional representation or latent vector $\mathbf{z}$. This process

condenses the input data into a more compact form, which captures its essential features. The decoder is the reverse process of the encoder. It takes the latent representation $\mathbf{z}$ and attempts to reconstruct the original input $\mathbf{x}$ (denoted as $\hat{\mathbf{x}}$). The loss function measures how well the decoder reconstructs the input data. Specifically, it computes the difference between the original input $\mathbf{x}$ and the reconstructed input $\hat{\mathbf{x}}$:

$$\mathcal{L}(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|^2$$

This difference is typically calculated using the squared error, which penalizes large deviations between the original and reconstructed inputs. The overall objective of the encoder-decoder architecture is to minimize the loss function as following:

$$\text{Minimize } \mathcal{L}(\mathbf{x}, g(f(\mathbf{x})))$$

where $f(\mathbf{x})$ is the encoder function that maps input $\mathbf{x}$ to a latent representation, and $g(f(\mathbf{x}))$ is the decoder function that reconstructs or transforms it into the desired output. A smaller loss indicates a better reconstruction. To find governing equations, we employ sparse identification of nonlinear dynamics (SINDy) that is a method used for discovering governing equations or dynamical systems directly from data [16, 19, 139]. It is particularly valuable in systems where the underlying dynamics are unknown or too complex to model explicitly but where data is available. SINDy aims to find a sparse model that describes the dynamics of a system using a small number of active terms from a larger library of possible functions. The idea is that many physical systems can be described by a few key interactions or governing equations, and SINDy helps identify these from observational data. The SINDy algorithm takes m -time measurements of $x \in \mathbb{R}^n$, $X = [x(t_1), \dots, x(t_m)]^T \in \mathbb{R}^{m \times n}$ and attempts to discover the structure of a nonlinear differential equation of the form

$$\dot{X} = f(X(t)) \approx \Theta(X)\Xi \quad (1.7)$$

where $\Theta(X) = [\theta_1(X), \theta_2(X), \dots, \theta_p(X)] \in \mathbb{R}^{m \times p}$ form the dictionary of basis functions, and $\Xi \in \mathbb{R}^{p \times n}$ is the matrix of coefficients, where each column corresponds to an equation with p terms.

1.2.3 Nonlinear Filtering

The final component of this work involves nonlinear filtering, which aims to predict model components, with a specific focus on sea surface temperature (SST) anomalies. Nonlinear filtering, also known as the theory of stochastic filtering, is a branch of mathematics and engineering focused on estimating the state of a dynamical system from noisy observations, where the relationship between the state and the observations is nonlinear. The development of nonlinear filtering has a rich history that spans several decades and involves contributions from various fields, including control theory, probability theory, and signal processing.

The foundation of filtering theory began with Norbert Wiener, who developed the Wiener filter in the 1940s [165]. Although the Wiener filter is a linear filter and optimal for linear systems with Gaussian noise, it laid the groundwork for later developments in stochastic processes and filtering theory. In 1960, Rudolf Kálmán introduced the Kalman filter [69], which is optimal for linear systems with Gaussian noise. The Kalman filter was a significant advancement because it provided a recursive solution to the linear filtering problem. However, it was limited to linear systems, and its extension to nonlinear systems was not straightforward. The work of Ruslan L. Stratonovich in the 1960s was crucial for the development of nonlinear filtering [150]. He formulated the Stratonovich

filtering equation, which is a nonlinear extension of the Kalman filter. This equation governs the conditional distribution of the state of a system given noisy observations but is difficult to solve analytically. During the same period, contributions from Tyrone Duncan, Hans Mortensen, and Moshe Zakai led to the Duncan-Mortensen-Zakai equation, another form of the nonlinear filtering problem [170]. Zakai's equation is particularly notable for simplifying the mathematical treatment of the filtering problem by converting it into a linear stochastic differential equation (SDE), albeit for an unnormalized version of the probability density. In 1967, Harold Kushner further developed nonlinear filtering by introducing an alternative formulation of the filtering problem known as the Kushner equation [80]. This is an SPDE that describes the evolution of the conditional probability density of the state given the observations.

Due to the complexity of nonlinear filtering equations, exact solutions are often intractable. During the 1970s and 1980s, researchers focused on developing approximation methods. Techniques such as the extended Kalman filter (EKF), unscented Kalman filter (UKF), and particle filters (sequential Monte Carlo methods) were developed to handle nonlinear systems and non-Gaussian noise [45]. Recent developments in nonlinear filtering have been driven by advances in computational power and the rise of machine learning. Techniques like ensemble Kalman filters [37], Gaussian processes, and deep learning-based approaches have been integrated with traditional filtering methods to handle increasingly complex and high-dimensional systems, such as those found in climate modeling, robotics, and finance [15, 54, 130].

In the continuous-time framework, the evolution of the hidden (unobserved) state of the system is typically described by a stochastic differential equation (SDE). This equation captures the system's dynamics, incorporating random noise and uncertainty in both the state transitions and observations. The model equation is given as

$$d\mathbf{x}(t) = f(\mathbf{x}(t), t) dt + \sigma(\mathbf{x}(t), t) d\mathbf{w}_1(t) \quad (1.8)$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the state vector at time t , $f(\mathbf{x}(t), t)$ is a nonlinear function representing the system dynamics, $\sigma(\mathbf{x}(t), t)$ is the diffusion term (possibly state-dependent) and $\mathbf{w}_1(t) \in \mathbb{R}^n$ is a Wiener process, representing process noise, respectively. The observation equation describes how the noisy measurements relate to the system's state. It is also a stochastic differential equation:

$$d\mathbf{y}(t) = h(\mathbf{x}(t), t) dt + d\mathbf{w}_2(t) \quad (1.9)$$

where $\mathbf{y}(t) \in \mathbb{R}^m$ is the observation vector at time t . $h(\mathbf{x}(t), t)$ is a nonlinear function relating the state to the observations, $d\mathbf{w}_2(t)$ is the observation noise, typically modeled as a Wiener process. To formulate the nonlinear filter, we assume f is a given bounded function R^d and $\hat{f}(x(t))$ is the optimal filter. A fundamental result of filtering theory says that the optimal filter is given by the formula

$$\hat{f}(\mathbf{x}(t)) = \frac{\int_{R^d} f(\mathbf{x}) u(t, \mathbf{x}) d\mathbf{x}}{\int_{R^d} u(t, \mathbf{x}) d\mathbf{x}} \quad (1.10)$$

where $u(t, \mathbf{x})$ is the so-called unnormalized filtering density (UFD). A standard way to study the UFD is to treat it as a solution of the Zakai equation

$$du(t, \mathbf{x}) = \mathcal{L}^* u(t, \mathbf{x}) dt + h(\mathbf{x}) u(t, \mathbf{x}) dy(t) \quad (1.11)$$

where $\mathcal{L}^*$ is the formally adjoint operator to the generator of the Markov process $\mathbf{x}(t)$ and is given

by:

$$\mathcal{L}^* u(\mathbf{x}, t) = -\nabla_{\mathbf{x}} \cdot (f(\mathbf{x}, t)u(\mathbf{x}, t)) + \frac{1}{2} \nabla_{\mathbf{x}} \cdot (\sigma(\mathbf{x}, t)\sigma^\top(\mathbf{x}, t)\nabla_{\mathbf{x}} u(\mathbf{x}, t))$$

The first term, $-\nabla_{\mathbf{x}} \cdot (f(\mathbf{x}, t)u(\mathbf{x}, t))$, corresponds to the drift of the system (deterministic dynamics). The second term, $\frac{1}{2} \nabla_{\mathbf{x}} \cdot (\sigma(\mathbf{x}, t)\sigma^\top(\mathbf{x}, t)\nabla_{\mathbf{x}} u(\mathbf{x}, t))$, represents the diffusion term, which accounts for the stochastic noise in the system. The solution can be written as [17, 94]:

$$u(t, \mathbf{x}) = \sum_{\alpha} \frac{1}{\sqrt{\alpha!}} \varphi_{\alpha}(t, \mathbf{x}) V_{\alpha}(y)$$

where $V_{\alpha}(y)$ are Wick polynomials and $\varphi_{\alpha}(t, x)$ are deterministic Hermite-Fourier coefficients in the Cameron-Martin orthogonal decomposition of $u(t, \mathbf{x})$. This method is useful for exploring the evolution of probability density. In this thesis, we present some results for the two-dimensional model. However, the model encounters several challenges when its dimensionality exceeds two.

On the other hand, we employ a particle filter to make predictions of SST anomalies. The particle filter approximates the posterior distribution $p(\mathbf{x}(t) \mid \mathbf{y}(t))$ using a set of weighted particles $\{\mathbf{x}_i^t, w_i^t\}_{i=1}^N$ [1]. At each time step, the particles are propagated forward using the state dynamics:

$$\mathbf{x}_i^t \sim p(\mathbf{x}^t \mid \mathbf{x}_i^{t-1}),$$

where $p(\mathbf{x}^t \mid \mathbf{x}_i^{t-1})$ is derived from the stochastic differential equation (1.8). The weights are updated based on the observation likelihood:

$$w_i^t \propto w_i^{t-1} \cdot p(\mathbf{y}^t \mid \mathbf{x}_i^t),$$

where the observation model is given by Equation (1.8)-(1.9). By using the particle filter, we are able to predict SST anomalies with time series. Such calculations form the core results of this thesis, with detailed computations provided within.

1.3 Thesis Layout

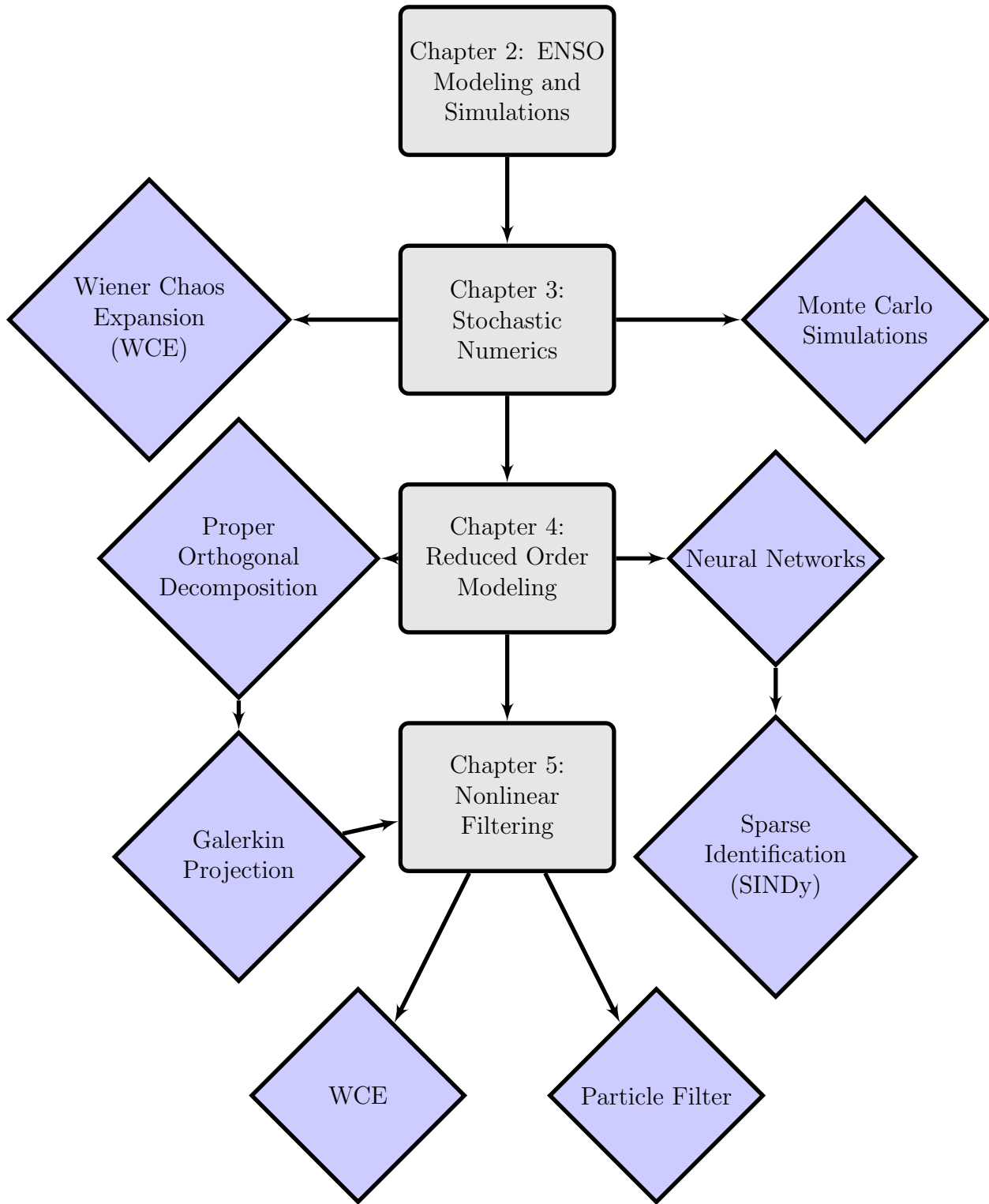


Figure 1.1: Thesis flowchart.

Chapter 2

Modeling and Simulations of the El Niño-Southern Oscillation (ENSO)

2.1 Climate Modeling

Climate can be defined as the average condition of the atmosphere, ocean, land surfaces and the ecosystems. It includes the average wind direction and strength, average cloud cover, the temperature of the sea surface nearby, which affects the previous quantities, and the ocean currents that affect the sea surface temperature. As an important area of climate science, climate modeling uses mathematical models. These models can be represented by mathematical equations, which describe the climate variables such as temperature, winds, and ocean currents. These models are solved numerically using computational methods, which are essential tools for understanding how natural factors and human activities influence the climate across different time scales, from seasons to centuries [24, 119].

Climate modeling requires different approximations depending on the location on Earth. For example, in mid-latitudes, the large scale atmospheric and ocean flows with their intense boundary currents and mesoscale eddies are very robust features and may be studied as the direct response to the curl of the wind stress since there is a fairly clear separation in the time and space scales. Mid-latitudes cover a significant part of Earth's regions and give rise to weather fronts and extratropical cyclones as well as occasional tropical cyclones. However, at low-latitudes (tropics), where the Coriolis number becomes very small and Rossby number may be order of unity, the atmosphere and ocean take on rather different characters than in mid-latitudes. Moreover, the tropics are the only regions where the surface boundary conditions such as sea-surface temperature (SST) have a significant impact on the weather and climate structure.

Earth's climate dynamics is mainly driven by the coupling of several components such as atmosphere, ocean, sea-ice, lands together with the energy sources including sun and SST. Although their separated dynamics are worthwhile to analyze at the local regions, uncoupled models do not provide sufficient information to understand the evolution of climate. Thus, coupled climate models are essential for the prediction of weather and climate. In this thesis, we study one of the most important phenomena at low-latitudes called El Niño-Southern Oscillations (ENSO), which is the dominant mode of variability in the tropical Pacific at inter-annual timescales. ENSO is modeled by

coupled atmosphere-ocean model together with SST at the equatorial Pacific confined in a region between Peru and Australia.

ENSO is observed every 3-7 years and has huge impacts on the global climate. Especially strong El Niño events (observed in 1982-83 and 1997-98) have been connected to low agricultural yields, availability of water resources, animal movements, forest fires, the economic well-being of nations, and many others. On the other hand, climate change also alters the ENSO dynamics. According to a recent study [159], climate change is increasing the frequency of extreme El Niño events, leading to intensifying droughts, worsening floods, and shifting hurricane patterns. It is projected that strong El Niño events may be observed in frequency from one every 20 years to one every 10 years by the end of the 21st century and the strongest events may be much stronger than today's strongest events [109]. For example, ocean stratification in the equatorial Pacific is expected to increase sharply in climate change. Therefore, it is worthwhile to develop novel models and techniques and analyze the ENSO phenomenon in a changing climate.

This chapter is organized as follows. The relation between El Niño and Southern Oscillation is provided in Section (2.2), where we present the climate dynamics at low latitudes. The ENSO modeling framework, which consists of equatorial Kelvin and Rossby waves coupled with sea surface temperature, is introduced in Section (2.3). We present the ENSO simulations in Section (2.4), which forms the core of this chapter and thesis. Finally, we discuss the conclusions and future direction in Section (2.5).

2.2 Equatorial Circulation and the El Niño

2.2.1 El Niño and Southern Oscillation

El Niño is one of the most prominent phenomena in climate science, with substantial impacts on human societies and the environment. It refers to an anomalous warming of surface waters in the eastern equatorial Pacific Ocean, and represents a powerful climate event with significant global consequences. El Niño is associated with heavy rainfall in regions such as California and northern Argentina, while simultaneously causing anomalously dry conditions in Southeast Asia and northern and eastern Australia. Additionally, it raises the global average surface temperature by approximately 0.5°C. When coupled with associated atmospheric changes, the phenomenon is collectively referred to as the El Niño–Southern Oscillation (ENSO). ENSO is the largest and most significant source of global climate variability on interannual timescales.

Every few years, the surface temperatures of the eastern tropical Pacific Ocean undergo significant warming. The most pronounced warming occurs within the region extending approximately from 5° S to 5° N in latitude, and from the west coast of Peru, at a longitude of approximately 80° W, to near the International Dateline at 180° W. This phenomenon, as depicted in Figure (2.1), is characterized by substantial temperature anomalies, with differences reaching up to 6°C between El Niño years and non-El Niño years.

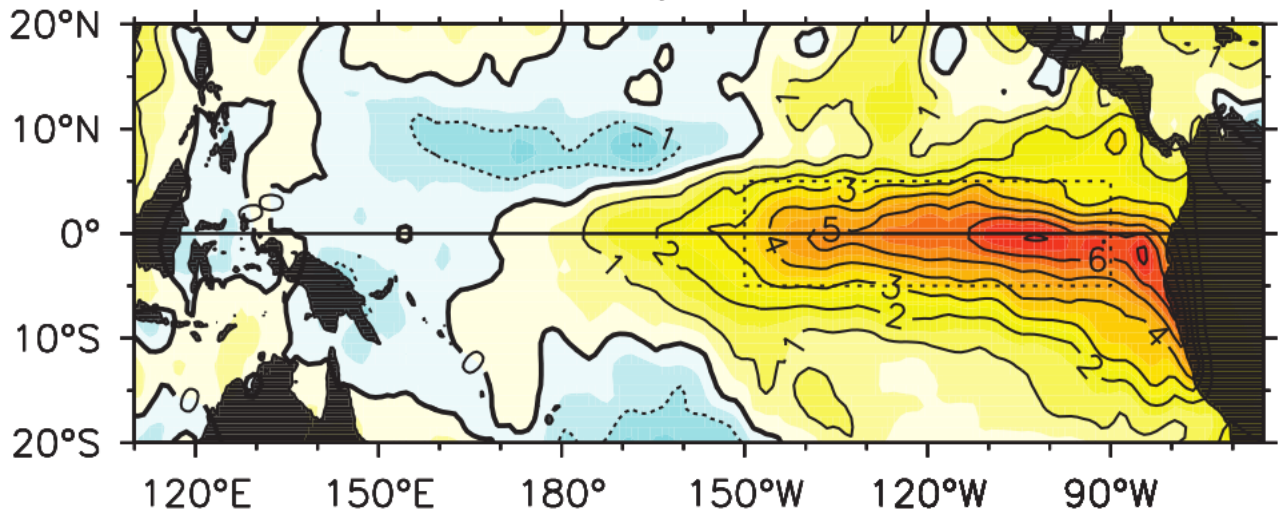


Figure 2.1: The sea-surface temperature anomalies during a strong El Niño year (December, 1997) [156].

The occurrence of these warm events is irregular, with intervals typically ranging from 3 to 7 years. Notably, particularly intense events have been documented in the years 1887–1888, 1982–1983, 1997–1998, and 2015–2016. The progression of the most recent major event, from 2015 to 2016, is illustrated in Figure (2.2).

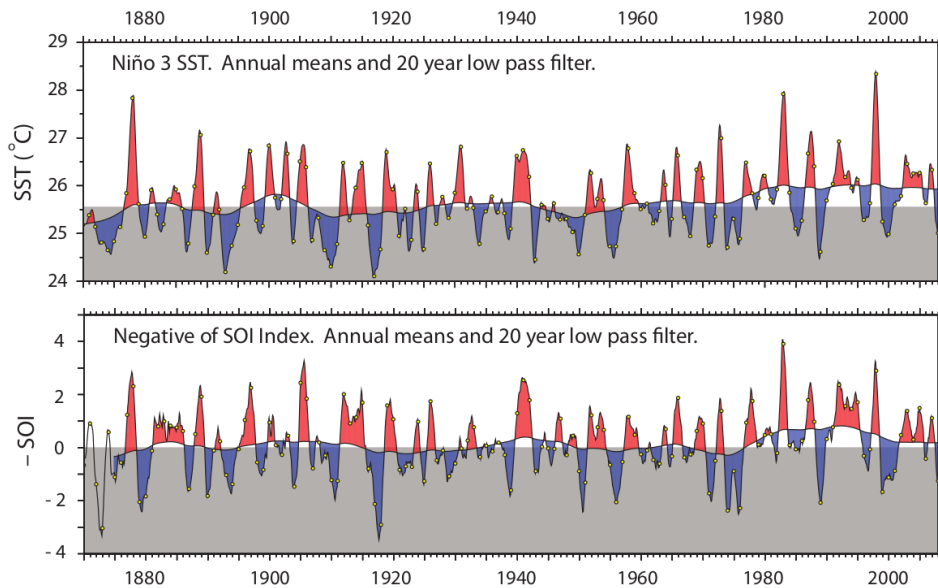


Figure 2.2: Upper: SST time series in the Eastern Equatorial Pacific (Niño 3) region. Lower: Negative of the Southern Oscillation Index (SOI), the anomalous pressure difference between Tahiti and Darwin [156].

2.2.2 Dynamics at Low Latitudes

For ENSO studies, the linearized shallow water equations (SWEs) on an equatorial beta-plane (which approximates the Coriolis force near the equator) are typically used. The SWEs are a simplified set of equations derived from the Navier-Stokes equations under the assumption that the horizontal scale is much larger than the vertical scale. The linearized equations are :

$$\begin{aligned}\frac{\partial u}{\partial t} - \beta y v &= -g \frac{\partial \eta}{\partial x} \\ \frac{\partial v}{\partial t} + \beta y u &= -g \frac{\partial \eta}{\partial y} \\ \frac{\partial \eta}{\partial t} + H \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) &= 0\end{aligned}\tag{2.1}$$

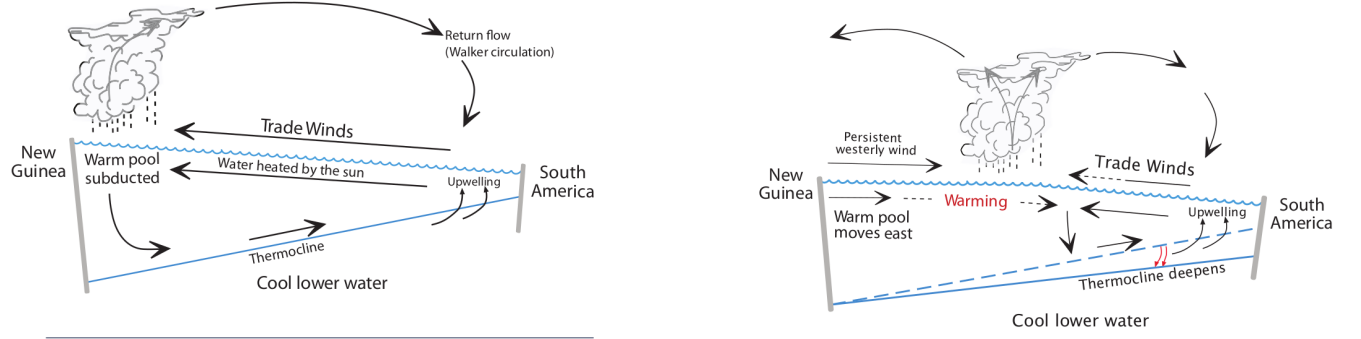
where u, v are zonal and meridional velocity components, η is free surface elevation (or thermocline displacement), g is gravitational acceleration, H is mean depth of the ocean layer, and $\beta = \frac{df}{dy}$ is variation of Coriolis parameter with latitude. These equations model equatorially trapped waves (especially Kelvin waves and Rossby waves) that alter the ocean's thermocline and transmit signals across the Pacific. Kelvin waves (eastward) are triggered by westerly wind anomalies and deepen the thermocline in the east while Rossby waves (westward) are reflected off the eastern boundary and help reestablish conditions. These waves are explained in detail in the next section.

At low-latitudes around the equator, the flow is not completely in a geostrophic balance and the Coriolis parameter due to the rotation of earth is very small and almost zero at the equatorial region. Geostrophy refers to a state of fluid flow in which the Coriolis force is balanced by the horizontal pressure gradient force. Mathematically, this balance can be shown as

$$f\vec{v} = \frac{1}{\rho}\nabla p\tag{2.2}$$

where f , $\vec{v}$, ∇p and ρ represent the Coriolis parameter, geostrophic velocity, pressure gradient and density, respectively. Geostrophic balance occurs when the Coriolis and pressure gradient forces are in equilibrium while it breaks down very close to the equator. In addition to the Coriolis effect, there are many other reasons why such an extraordinary mechanism is observed at low-latitudes rather than mid-latitudes [156]. The first reason is that the equatorial SSTs are usually high, around 27° and convection is easily generated as the temperature further increases. Secondly, thermocline is quite shallow at the equator and varies from 200m in the west to 50m in the east. The shallow thermocline causing upwelling leads to the close relation between SST and thermocline depth. Finally, equatorial Kelvin and Rossby waves are quite fast and that allows ENSO events to occur in a short time.

A cross-section of the equator during non-El Niño and El Niño phases



(a) A cross-section along the equator during non-El Niño.

(b) A cross-section along the equator at the peak of El Niño.

Figure 2.3: Sketch of the key members of the ENSO cycle [156].

Usual mechanism at equatorial Pacific is driven by the westerly trade winds (Walker circulation), which transfers the warm surface water from eastern part (South America) to western part (New Guinea) as demonstrated in Figure (2.3a). The upwelling of cold water coming from the below of thermocline reduces SST in the eastern part where typical SST in the eastern Pacific is low, around 21° , and is high in the western Pacific, around 30° . The ocean water near Peru and contiguous countries is normally colder and contains more nutrients than usual that makes fishing and related aquacultural activities as the main income source. However, during El Niño, upwelling weakens or even stops and westerly trade winds weaken and even reverse as shown in Figure (2.3b). Reversed trade winds bring the warm water to eastern regions. This atmosphere-ocean coupled mechanism depicts the seasonal cycle of ENSO. And, without the nutrients from the deep, there are fewer phytoplankton off the coast. This affects the food chain in the ocean such that fish eat phytoplankton and, in turn, affects everything that eats fish. That changes business factors in South America and extremely drought weather is observed in Australia and nearby countries during El Niño events.

2.2.3 Equatorial Waves

Equatorial Kelvin and Rossby waves are critical to the dynamics of the El Niño–Southern Oscillation (ENSO) as they facilitate the redistribution of heat and momentum across the equatorial Pacific. Kelvin waves are eastward-propagating waves characterized by their rapid speed and confinement to the equatorial region due to the Coriolis force. These waves are triggered by wind stress anomalies, such as the weakening of trade winds during El Niño events. When a Kelvin wave propagates eastward, it deepens the thermocline and raises sea surface temperatures (SSTs) in the eastern Pacific, suppressing the usual upwelling of cold, nutrient-rich water. This process is a key factor in initiating and amplifying the warming phase of El Niño. Kelvin waves are particularly efficient in transmitting the effects of atmospheric forcing across the vast Pacific basin.

Rossby waves, on the other hand, are westward-propagating and slower than Kelvin waves. They are generated when Kelvin waves reflect off the eastern boundary of the Pacific basin or as a direct response to atmospheric forcing. Rossby waves contribute to the redistribution of thermocline anomalies back toward the western Pacific. During the transition from El Niño to

La Niña, Rossby waves play a crucial role in restoring the thermocline structure to its normal state, allowing cooler waters to dominate the eastern Pacific. The interplay between these wave types creates a feedback mechanism that drives the evolution and termination of ENSO events. Together, Kelvin and Rossby waves act as key carriers of ocean-atmosphere coupling, regulating the timescales and spatial patterns of ENSO. The mechanism of oceanic Kelvin and Rossby waves along the equator is depicted in Figure (2.4). A detailed discussion of the reflection of equatorial Kelvin and Rossby waves is provided in [148].

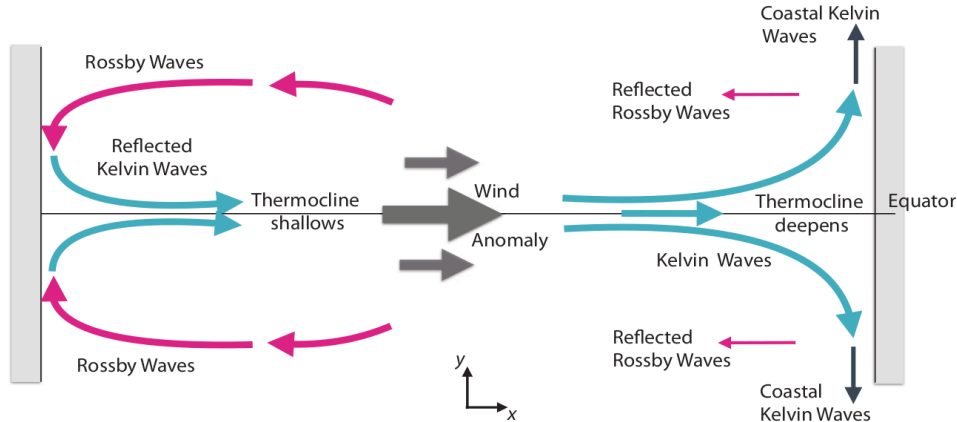


Figure 2.4: Oceanic waves induced by a wind anomaly on the equator [156].

2.3 ENSO Modeling

The modern analysis of climate models dates back to the 18th century with the works of Edmond Halley. For the first time, he went beyond the Ptolemaic view of climate and explained the trade winds as a function of latitude [49]. Since then, various perspectives of climate modeling have been developed [35]. First remarks of climatic anomalies at low-latitudes are based on the observations of Bjerknes in the late 1960s [12]. He recognized the correlation between SST and equatorial atmospheric circulations which is also named as the Walker Circulation. Such observations are considered as the links between El Niño and the Southern Oscillations. Along with the empirical approach, Matsuno (1966) and Gill (1980) utilized simplest and effective one-layer shallow-water models with damping and provided analytic solutions [42]. This model is known as the Matsuno-Gill model and the atmospheric response is represented in terms of equatorial Kelvin and Rossby waves. In addition to analytic solutions, the experimental forecasts of ENSO have been made since the 1970s. However, the atmosphere-ocean coupled model of ENSO has not been developed since the mid-1980s. Cane and Zebiak at Columbia University proposed a deterministic numerical model that predicted the ENSO event in 1986 [171]. This model is known as the Cane-Zebiak model and accurately captures the key elements of atmosphere-ocean coupled dynamics. Neelin, Dijkstra and their co-workers introduced a comprehensive ENSO theory and analyzed system components [29, 110, 120]. Moreover, works of Boyd [14] and Majda [100] on the equatorial wave dynamics are useful to understand linear and nonlinear structure of ENSO dynamics.

Simulating the ENSO remains a significant challenge in climate science due to several persistent issues. Despite advancements in modeling techniques, several persistent issues hinder accurate representation and prediction of ENSO events. A major difficulty lies in correcting mean state biases,

particularly the cold tongue bias in the equatorial Pacific, which undermines accurate representation of ENSO dynamics [88]. Moreover, uncertainty persists in how climate change will affect future ENSO behavior, with models offering divergent projections of its intensity and frequency [47]. Model resolution is another critical factor in ENSO simulation [162]. Higher-resolution ocean models better resolve narrow equatorial currents, coastal upwelling, and ocean eddies [91]. However, some studies suggest increasing atmospheric resolution alone can sometimes worsen ENSO asymmetry or shift the location of SST anomalies unrealistically [90].

Most of the climate models, which have been developed so far, are based on deterministic approaches. Although deterministic models explained the essence of dynamics, climate models are mostly chaotic and include random effects. Thual and Chen recently developed a simple stochastic model that captures the main dynamics of ENSO [23, 155]. Their approaches are based on the stochastic modeling of westerly wind bursts (WWB) that change the ocean phase. Depending on the strength of the stochastic forcing, the state is modeled as a multi-state Markov chain that switches between El Niño, non-El Niño and strong-El Niño phases. Although developed models provide information about ENSO, there are still several limitations for the prediction of ENSO phenomenon [75] due to the sensitivities to the stochastic forcing. Therefore, it is important to develop appropriate stochastic models, numerical schemes and effective techniques for the accurate prediction of ENSO dynamics.

Coupled ENSO models primarily focus on the upper layer of the tropical Pacific Ocean. This layer drives energy and momentum exchange with the tropopause. The deep ocean, in contrast, is considered dynamically inactive. The upper ocean of interest is a very long strip of water extending from 124E to 80W and from 29S to 29N with: i) a meridional (north-south) width of a few hundred kilometers, ii) a zonal (east-west) extension of about 10,000 kilometers and iii) an average thermocline depth (downwards from the sea surface) of about 120 meters, which is a relatively thin ocean layer above the deep cold waters.

2.3.1 A Meridionally Truncated ENSO Model

In this chapter and thesis, we present the nonlinear ENSO model developed by Majda and colleagues [23, 155] (Majda’s model), where the ocean and atmosphere are both modeled as non-dissipative linear shallow water equations, with a dissipative nonlinear thermodynamics budget. Similar to the Zebiak-Cane model, it is an anomaly model, where the anomalies are calculated relative to a mean annual cycle specified from observations, and thus all state variables in subsequent sections are anomalies from an equilibrium state and are nondimensional. Building on the foundation of the Cane-Zebiak model, Majda’s model addresses the limitations of linearity and determinism, making it a more versatile tool for studying the complexities of ENSO phenomena. This model provides a more realistic framework for studying ENSO variability, extremes, and predictability in the presence of noise, although we use deterministic version in this paper. A complete ENSO model described by zonal and meridional components is provided in Appendix (A). Such models can then be further simplified by meridional truncation that reduces the dimensionality of the problem by approximating the solution in the meridional (north-south) direction using a predefined set of modes.

The nonlinear coupled PDEs in five variables which model the interaction of equatorial atmo-

spheric and oceanic Kelvin and Rossby waves with SST anomalies, defined as

$$u(t, x) \stackrel{\text{def}}{=} [K^A(t, x), R^A(t, x), K^O(t, x), R^O(t, x), T(t, x)]^T$$

and can be written in differential equation form as

$$\mathcal{M} \frac{\partial u}{\partial t}(t, x) = \mathcal{A}(\kappa)u(t, x) + \mathcal{N}(u, u, \mu)(t, x) \quad (2.3)$$

where $\mathcal{M}$ is a 5×5 constant matrix with diagonals $[0, 0, 1, 1, 1]$, the linear operator $\mathcal{A}(\kappa)$ defined as

$$\mathcal{A}(\kappa) = \begin{bmatrix} -\frac{\partial}{\partial x} - \gamma & 0 & 0 & 0 & \chi_A \alpha_q (2 - 2\bar{Q})^{-1} \\ 0 & \frac{1}{3} \frac{\partial}{\partial x} - \gamma & 0 & 0 & \chi_A \alpha_q (3 - 3\bar{Q})^{-1} \\ \chi_O \frac{c\kappa}{2} & -\chi_O \frac{c\kappa}{2} & -c \frac{\partial}{\partial x} - \delta & 0 & 0 \\ -\chi_O \frac{c\kappa}{3} & \chi_O \frac{c\kappa}{3} & 0 & \frac{c}{3} \frac{\partial}{\partial x} - \delta & 0 \\ 0 & 0 & c\eta(x) & c\eta(x) & -c\xi\alpha_q \end{bmatrix} \quad (2.4)$$

depends on a parameter κ , the wind stress coefficient, and nonlinearity $\mathcal{N}(u, u, \mu)$ with advection coefficient μ , is defined by

$$\mathcal{N}(u, u, \mu) \stackrel{\text{def}}{=} [0, 0, 0, 0, -\mu \frac{\partial}{\partial x} [(K^O(t, x) - R^O(t, x))T(t, x)]]^T.$$

In equation (2.3), small damping terms γ, δ are introduced to guarantee a unique solution, E_q is latent heat, ξ is the latent heating exchange coefficient, $\bar{Q}$ is a constant representing mean vertical moisture gradient, $\eta(x)$ is the profile of thermocline feedback and the projection coefficients are defined as

$$\chi_A = \int_{-\infty}^{\infty} \phi_0(y) \psi_0(Y) dy \quad \text{and} \quad \chi_O = \int_{-\infty}^{\infty} \psi_0(Y) \phi_0(y) dY,$$

where ϕ_0 and ψ_0 denote the atmospheric and oceanic parabolic cylinder functions as a function in meridional position y and $Y = y/\sqrt{c}$. The initial conditions for the oceanic variables of (2.3) are given by

$$u_3(x, 0) = f_3(x), \quad u_4(x, 0) = f_4(x) \quad \text{and} \quad u_5(x, 0) = f_5(x). \quad (2.5)$$

Periodic boundary conditions that were adopted for the atmosphere model in [23] are modified to reflect the inclusion of the small damping term γ . Boundary conditions for the atmosphere variables that are confined in the ocean domain:

$$K^A(t, 0) = e^{-\gamma(L_A - L_O)} K^A(t, L_O) \quad \text{and} \quad R^A(t, 0) = e^{3\gamma(L_A - L_O)} R^A(t, L_O) \quad (2.6)$$

where L_A and L_O represent the length of the equatorial belt and Pacific ocean, respectively. The derivation of boundary conditions (2.6) is provided at the end of this section.

Having defined the ocean domain, the next step is to specify the boundary conditions that govern wave interactions and energy exchange as the following:

$$K^O(t, 0) = r_W R^O(t, 0) \quad \text{and} \quad R^O(t, L_O) = r_E K^O(t, L_O) \quad (2.7)$$

where $x = 0$ and $x = L_O$ represent the western and eastern parts of the ocean, respectively. The unidirectional Kelvin and Rossby oceanic equatorial waves are coupled through the boundary condition (2.7), where the oceanic waves are reflected with coupling coefficients r_W that represents

partial loss of energy in the western Pacific boundary and r_E that represents partial loss of energy due to the north-south propagation of the coastal Kelvin waves along the eastern Pacific boundary.

Equation (2.3) consists of a simple nonlinear advection $-\mu \frac{\partial}{\partial x} [(K^O(t, x) - R^O(t, x))T(t, x)]$ in the SST equation. We assume no heat flux at the eastern boundary, as specified by the following boundary conditions for $T(t, x)$:

$$\frac{\partial T}{\partial x} \Big|_{x=L_O} = 0 \quad (2.8)$$

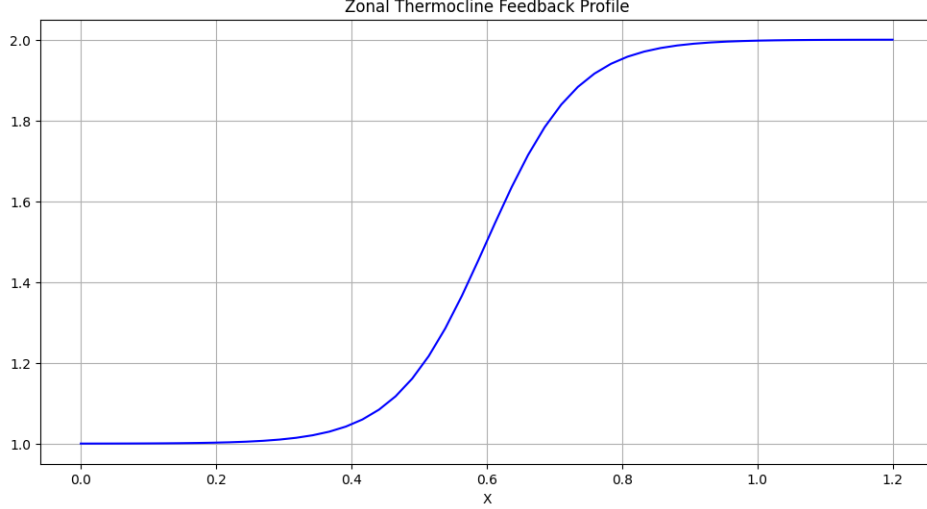


Figure 2.5: Zonal thermocline feedback profile $\eta(x)$.

The thermocline depth, influenced by the oceanic Kelvin wave $K^O(t, x)$ and Rossby wave $R^O(t, x)$, is incorporated into the SST component through the thermocline feedback variable $\eta(x) = 1.5 + 0.5 \times \tanh(7.5x - L_O/2)$ as shown in Figure (2.5). It is a feature of the coupled ENSO model (2.3) that there are no time derivatives of the first two components, $K^A(t, x)$ and $R^A(t, x)$. Periodic boundary conditions that were adopted for the atmosphere model in [23] are modified to reflect the inclusion of the small damping term γ . Hence, the truncated atmospheric model can be explicitly written as,

$$\begin{aligned} \partial_x K^A(t, x) &= -\gamma K^A(t, x) + \chi_A \alpha_q (2 - 2\bar{Q})^{-1} T(t, x) & 0 \leq x \leq L_O \\ \partial_x K^A(t, x) &= -\gamma K^A(t, x) & L_O \leq x \leq L_A \end{aligned} \quad (2.9)$$

$$\begin{aligned} \partial_x R^A(t, x) &= 3\gamma R^A(t, x) - 3\chi_A \alpha_q (3 - 3\bar{Q})^{-1} T(t, x) & 0 \leq x \leq L_O \\ \partial_x R^A(t, x) &= 3\gamma R^A(t, x) & L_O \leq x \leq L_A \end{aligned} \quad (2.10)$$

with the periodic boundary conditions $K^A(t, 0) = K^A(t, L_A)$ and $R^A(t, 0) = R^A(t, L_A)$. The solutions of (2.9) and (2.10) can be written as

$$\begin{aligned} K^A(t, x) &= C_1 e^{-\gamma x} \quad \Rightarrow \quad C_1 = K^A(t, L_O) e^{\gamma L_O} & L_O \leq x \leq L_A \\ R^A(t, x) &= C_2 e^{3\gamma x} \quad \Rightarrow \quad C_2 = R^A(t, L_O) e^{-3\gamma L_O} & L_O \leq x \leq L_A \end{aligned}$$

That provides the solutions for (2.9) and (2.10) as

$$K^A(t, x) = e^{-\gamma(x-L_O)} K^A(t, L_O) \quad \text{and} \quad R^A(t, x) = e^{3\gamma(x-L_O)} R^A(t, L_O)$$

Using the original boundary conditions $K^A(t, 0) = K^A(t, L_A) = e^{-\gamma(L_A-L_O)} K^A(t, L_O)$ and $R^A(t, 0) = R^A(t, L_A) = e^{3\gamma(L_A-L_O)} R^A(t, L_O)$ yields to the modified boundary condition confined in the ocean domain:

$$K^A(t, 0) = e^{-\gamma(L_A-L_O)} K^A(t, L_O) \quad \text{and} \quad R^A(t, 0) = e^{3\gamma(L_A-L_O)} R^A(t, L_O)$$

By making this modification, we confine the model domain to the ocean region rather than the entire equatorial belt, which is beneficial for the modeling and simulation framework. As discussed in Zebiak [171] the parameter γ is intended to represent the effect of a wide range of physical processes and as such is not easy to determine a priori. Gill [42] uses a value of 0.1, while Zebiak [171] uses a value of 0.3 arguing that that value gave, in his case, good agreement between model results and observations. A wide range of values of γ may be used without affecting qualitative conclusions. In the first order boundary value problems (2.9) and (2.10), t is a parameter and $T(t, x)$ is given. The general solutions for $K^A(t, x)$ and $R^A(t, x)$ are given explicitly in terms of $T(t, x)$ as

$$\begin{aligned} K^A(t, x) = & \frac{e^{-\gamma x} \chi_A (2 - 2\bar{Q})^{-1} \alpha_q e^{-\gamma L_A}}{1 - e^{-\gamma L_A}} \int_0^{L_O} e^{\gamma \xi} T(t, \xi) d\xi \\ & + \chi_A (2 - 2\bar{Q})^{-1} \alpha_q e^{-\gamma x} \int_0^x e^{\gamma \xi} T(t, \xi) d\xi, \quad 0 \leq x \leq L_O, \end{aligned} \quad (2.11)$$

and

$$\begin{aligned} R^A(t, x) = & \frac{-e^{3\gamma(x-L_O)} 3\chi_A (3 - 3\bar{Q})^{-1} \alpha_q e^{-3\gamma L_A}}{1 - e^{-3\gamma L_A}} \int_0^{L_O} e^{-3\gamma(\xi-L_O)} T(t, \xi) d\xi \\ & - 3\chi_A (3 - 3\bar{Q})^{-1} \alpha_q e^{3\gamma(x-L_O)} \int_x^{L_O} e^{-3\gamma(\xi-L_O)} T(t, \xi) d\xi, \quad 0 \leq x \leq L_O. \end{aligned} \quad (2.12)$$

2.4 ENSO Simulations

In this section, we provide the numerical solutions of the nonlinear ENSO model (2.3). Numerical integration is used to compute $K^A(t, x)$ and $R^A(t, x)$ as given by Equations (2.11) and (2.12). To compute $K^O(t, x)$, $R^O(t, x)$ and $T(t, x)$ in (2.3), we use finite difference method. Numerical scheme is summarized in Table (2.1) as the following :

Numerical Algorithm for the El Niño-Southern Oscillation

1. Choose initial conditions $u_3(x, 0) = f_3$, $u_4(x, 0) = f_4(x)$ and $u_5(x, 0) = f_5(x)$ in (2.5) such that they satisfy the boundary conditions (2.6) and (2.8). It should be noted that there is no initial condition for $K^A(t, x)$ and $R^A(t, x)$.
 2. Define grid points in the $(t, x) \in [0, T] \times U$ and Δt and Δx such that they satisfy the CFL condition for the numerical stability.
 3. Use numerical integration for $K^A(t, x)$ and $R^A(t, x)$ as given by Equations (2.11) and (2.12).
 4. Employ backward space finite difference for $K^O(t, x)$, forward space finite difference for $K^O(t, x)$. Use either backward or forward finite difference for $T(t, x)$ depending on the derivative term.
 5. Apply the boundary conditions (2.6), (2.7) and (2.8) at the boundaries $x = 0$ and $x = L_O$ for each iteration.
 6. Simulate the model for a sufficiently long duration to allow it to reach the steady state after the transient phase.
 7. Use the snapshots just after the transient phase as a new initial conditions to re-simulate the model, where the affect of the initial conditions disappears. This allows us to obtain simulations without transients for further analysis.
-

Table 2.1: Numerical scheme for the the El Niño-Southern Oscillation (ENSO).

Figure (2.6) illustrates the simulation of the ENSO model (2.3). In the figures, while $x = 0$ represents western boundary (Australia), $x = 1.2$ represent the eastern boundary (South America). We observe that atmospheric Kelvin $K^A(t, x)$ and atmospheric Rossby $R^A(t, x)$ has steady state dynamics as there is no time derivative in the equations. Oceanic Kelvin $K^O(t, x)$ and oceanic Rossby $R^O(t, x)$ move eastward with a speed c and westward with a speed $c/3$, respectively. Our simulation also investigates the evolution of SST anomalies $T(t, x)$. Especially, the SST dynamics is crucial at $x = L_0$ as it generates an El Niño condition for high values (shown by red). The model parameters, which are based on the values given in [23], are shown in Table (2.2).

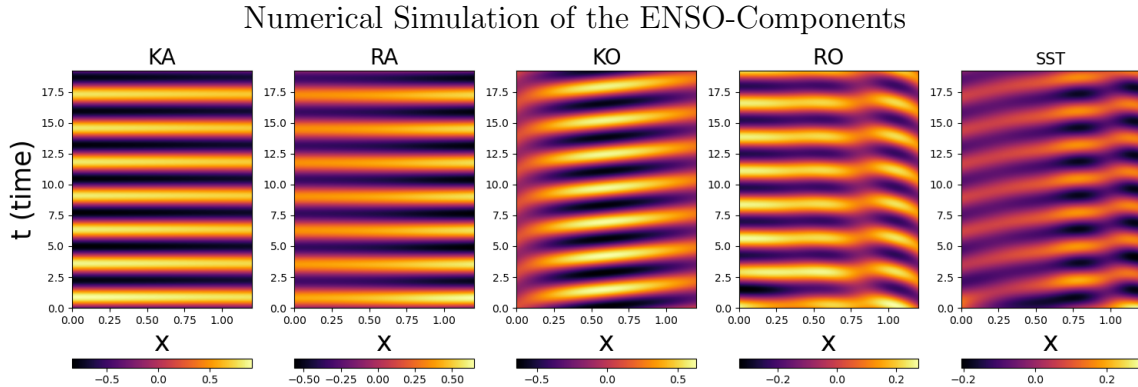


Figure 2.6: Simulation of the equatorial atmospheric Kelvin and Rossby waves, oceanic Kelvin and Rossby waves and sea surface temperature.

The Courant–Friedrichs–Lewy (CFL) condition is a necessary condition for the stability of many explicit time-stepping schemes when solving partial differential equations (especially hyperbolic PDEs) [26, 48]. For the PDE,

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0$$

a typical explicit finite difference scheme requires the CFL condition:

$$\frac{a \Delta t}{\Delta x} \leq 1,$$

In our simulations of the model (2.3), we choose $\Delta t = 0.006$ and $\Delta x = 0.04$ so that the CFL condition is satisfied for the numerical stability.

Parameter	Description	Nondimensional Values
κ	Wind stress coefficient	5.6
γ	Atmospheric damping coefficient	0.1
χ_A	Meridional projection coefficient from ocean to atmosphere	0.31
χ_O	Meridional projection coefficient from atmosphere to ocean	1.38
α_q	Latent heating factor	0.2
$\bar{Q}$	Mean vertical moisture gradient	0.9
c	Wave propagation speed	0.5
δ	Oceanic damping coefficient	0.5
ξ	Coupling strength between ocean and atmosphere	8.5
μ	Nonlinear zonal advection coefficient	0.04
r_W	Western boundary reflection coefficient in ocean	0.5
r_E	Eastern boundary reflection coefficient in ocean	0.5
L_A	Equatorial belt length	2.6
L_O	Equatorial Pacific length	1.2

Table 2.2: Descriptions of parameters in the operator $\mathcal{A}(\kappa)$ and nonlinearity $\mathcal{N}(u, u, \mu)$.

2.5 Conclusion and Future Work

In this chapter, we introduced the El Niño-Southern Oscillation (ENSO), which forms the modeling core of this research. We first discussed the climate modeling and simulations. Then, the climate dynamics at low latitudes around the equator was presented. Moreover, we emphasized the importance of ENSO on the global climate. This was followed by a discussion on atmosphere-ocean-SST feedback mechanism of ENSO, and geophysical flows, including equatorial Kelvin and Rossby waves.

In the next phase of this chapter, we presented the nonlinear ENSO model introduced by Chen and Majda [23]. By modifying the atmospheric boundary conditions, we proposed a slightly novel model that is confined to the Pacific ocean rather than the entire equatorial belt. Finally, we illustrated simulation results of the ENSO model at the end of this chapter.

Future work will focus on the stochastic ENSO modeling and simulations based on Wiener Chaos Expansion and Monte Carlo methods, which are useful methodologies to calculate the statistical

properties, i.e., mean and variance. Moreover, we will explore reduced-order modeling frameworks based on the POD-Galerkin and machine learning. These methods will be useful to reduce the computational complexity of the problems. Finally, we will apply data assimilation techniques to estimate the SST states of the ENSO model.

Chapter 3

A Numerical Study of Stochastic El Niño-Southern Oscillation (ENSO) Using Wiener Chaos Expansion

3.1 Stochastic Numerical Methods

Stochastic numerical methods have evolved significantly since the early 20th century to model systems influenced by randomness. Unlike the simulation of deterministic systems, which produce a single outcome for a given initial condition, simulations of stochastic systems incorporate the random elements to capture the inherent uncertainty and variability of real-world problems. The foundation of stochastic modeling was laid by Andrey Kolmogorov's work on probability theory [77] and Norbert Wiener's studies on Brownian motion [163]. These developments culminated in the 1940s with Itô stochastic calculus, which provided a rigorous mathematical framework for analyzing systems with randomness [63]. Around the same period, the Monte Carlo method—introduced during the Manhattan Project by Ulam, von Neumann, and Metropolis—became a pivotal tool for simulating random processes [112]. The evolution of stochastic numerical methods also includes the development of numerical solutions for stochastic differential equations (SDEs). The Euler-Maruyama method (1960s-1970s) provided a basic framework [76], followed by Milstein's scheme, which included strong convergence properties [114]. These methods addressed stability and efficiency challenges in stochastic simulations.

In addition to ensemble approaches in stochastic numerical methods like in Monte Carlo, the solution of stochastic differential equations using Wiener Chaos Expansion (WCE) is another important application of spectral methods. The WCE was originally introduced by Norbert Wiener in 1938 [164, 166]. Using Hermite polynomials, Wiener constructed an orthonormal random basis for expanding homogeneous chaos depending on white noise. Based on Wiener's idea, Cameron and Martin developed a more explicit formulation for the WCE [17]. This development is based on an explicit discretization of the white noise process through its Fourier expansion. This approach gained prominence in the late 20th century as a systematic framework to represent and solve SPDEs, particularly in contexts where randomness interacts with spatial and temporal dynamics.

The application of Wiener Chaos Expansion (WCE) to stochastic finite elements gained promi-

nence in the 1990s, particularly through the work of Ghanem and Spanos [41]. Their approach transformed SPDEs into a hierarchy of deterministic equations, making uncertainty quantification computationally efficient. Lototsky and Rozovsky further advanced the use of WCE by formulating SPDEs within a Hilbert space framework. By expressing solutions as an infinite series of orthogonal Wiener chaos terms, they derived a hierarchy of coupled deterministic equations, effectively handling systems driven by white noise or Gaussian processes. These equations can then be solved using numerical or analytical methods. Their approach proved particularly effective for SPDEs driven by white noise or other Gaussian processes, as WCE naturally accommodates these types of randomness [94–96].

In this chapter, we examine the effects of stochastic perturbations on ENSO dynamics and introduce novel modeling and numerical schemes based on the Wiener Chaos Expansion (WCE). We also compare these methods with Monte Carlo (MC) simulations. The key idea behind WCE is the explicit discretization of white noise through Fourier expansion. It is closely related to techniques like proper orthogonal decomposition (POD), Karhunen-Loève (KL) decomposition, and other spectral methods [168]. Theoretical foundations of WCE, along with its applications to the 1D wave equation, Burgers equations, and Navier-Stokes equations, can be found in [71, 95, 98]. The WCE framework offers several advantages, including the ability to separate random and deterministic components of the dynamics and the feasibility of approximating probability densities in random physical systems. However, it also has limitations, particularly in terms of high dimensionality, non-linearity, and approximation accuracy, as discussed in [101]. Therefore, combining WCE with the MC approach provides a useful strategy for addressing random processes with uncertain statistical properties.

This chapter is organized as follows. In Section (3.2), we provide a general framework highlighting the importance and application of stochastic methods for ENSO modeling and simulations. The theoretical foundation of WCE and its fundamental theorems are presented in Section (3.3). As an application of WCE, we study a minimal ocean problem in Section (3.4) and demonstrate the accuracy and efficiency of the proposed WCE framework. The main application of this chapter using WCE is introduced in Section (3.5), where we examine a linear coupled stochastic ENSO model forced by an Ornstein-Uhlenbeck process. To validate the convergence and accuracy of the WCE results, we compare them with those obtained from the same model using the Monte Carlo approach in Section (3.6).

3.2 The El Niño-Southern Oscillation and Stochastic Simulations

Stochastic climate models are mathematical frameworks that describe the evolution of climate systems by incorporating random processes to account for the inherent uncertainty and variability in climate dynamics. In contrast, traditional climate models are typically deterministic, relying on systems of differential equations to simulate physical processes like temperature, humidity, and atmospheric circulation. However, many climate phenomena—such as turbulence, cloud formation, and ocean currents—are uncertain or unresolved, making it necessary to incorporate stochastic methods to capture these effects more accurately [40, 50, 102].

Stochastic simulations are critical for understanding and predicting the El Niño-Southern Oscillation (ENSO), a key driver of climate variability. ENSO consists of irregular oscillations between warm (El Niño) and cold (La Niña) phases in the Pacific Ocean, significantly impacting global weather patterns. Due to ENSO’s inherent complexity and variability, it is particularly suited for stochastic modeling. Deterministic models alone often fail to capture the full spectrum of observed behaviors and the impact of random fluctuations. ENSO is a chaotic phenomenon characterized by complex interactions between the atmosphere and ocean on multiple scales, from sea surface temperatures in the equatorial Pacific to the broader global atmospheric circulation. Stochastic processes are incorporated into simulations to account for unresolved small-scale processes, such as turbulence, ocean-atmosphere coupling, and random atmospheric perturbations. By modeling these uncertainties, stochastic simulations offer a more accurate representation of ENSO’s variability, including its irregular periods and amplitudes.

Climate scientists often use stochastic differential equations (SDEs) to represent the coupled ocean-atmosphere dynamics that govern ENSO. These equations include deterministic terms describing the main physical processes (e.g., heat transport, wind stress), and stochastic terms that model random fluctuations such as atmospheric noise or small-scale turbulence. This randomness can influence the timing, intensity, and frequency of El Niño and La Niña events. A simplified stochastic model for ENSO could take the form:

$$dX_t = f(X_t)dt + g(X_t)dW_t \quad (3.1)$$

where X_t represents a key climate variable (e.g., sea surface temperature anomalies in the Pacific), $f(X_t)$ models deterministic ENSO dynamics and $g(X_t)dW_t$ introduces stochastic forcing driven by a Wiener process W_t . This type of equation captures the interplay between predictable dynamics and random perturbations in ENSO [103, 154]. In the next section, we provide the theoretical background for solving stochastic differential and partial differential equations like Equation (3.1) using the Wiener Chaos Expansion, which forms the core of this chapter.

3.3 Wiener Chaos Expansion

The Wiener Chaos Expansion (WCE) is a powerful mathematical tool used to represent random variables as infinite series of multiple Wiener-Itô integrals, organized in terms of orthogonal polynomials of Gaussian processes. By decomposing random variables into chaos levels, WCE provides a systematic framework for analyzing the effects of randomness in complex stochastic systems. The 0th level of the expansion corresponds to deterministic terms, the 1st level includes linear terms associated with the Wiener process, and higher levels capture increasingly complex nonlinear interactions. This hierarchical structure makes WCE particularly effective in studying stochastic differential equations, uncertainty quantification, and financial models, where random fluctuations significantly influence system behavior. Its ability to systematically manage and approximate the contributions of noise to such systems highlights its utility in simplifying the analysis and improving the understanding of stochastic dynamics [124].

WCE relies on Hermite polynomials because they form a natural basis for Gaussian random processes, ensuring orthogonality under the Gaussian probability measure. Each term in the WCE

corresponds to a multiple Wiener-Itô integral weighted by Hermite polynomials, where the polynomial degree reflects the chaos level. This relationship enables WCE to efficiently capture the structure of randomness in systems governed by Gaussian noise, providing a mathematically rigorous and computationally effective approach to stochastic analysis [65]. In Appendix (B), we provide some properties of Hermite polynomials and their relationships with Gaussian random variables.

3.3.1 Wiener Chaos Expansions of Functionals of Brownian Motions

In this part of this thesis, we are interested in a class of stochastic partial differential equations (SPDEs)

$$u_t(t, x) = \mathcal{L}(u) + \sigma \dot{W}(t) \quad (3.2)$$

where $\mathcal{L}(u)$ is a general linear or nonlinear differential operator in spatial variables, and $W(t)$ is a Brownian motion. The solution of (3.2) not only depends on the spatial-time variable (t, x) but also the Brownian motion path. For any fixed time $t > 0$, assume $m_i(s)$, $i = 1, 2, \dots$ are a set of complete orthonormal bases in the Hilbert space $L^2([0, t])$. Define the Ito integral

$$\xi_i = \int_0^t m_i(s) dW(s) \quad i = 1, 2, \dots \quad (3.3)$$

The equation (3.3) can be interpreted as the projection of the white noise $\dot{W}(s)$ onto L^2 basis function $m_i(s)$ because $\xi_i = (m_i, \dot{W}_s)$. From the isometry property of the Ito integrals, we have

$$\mathbb{E}[\xi_i \xi_j] = \int_0^T m_i(s) m_j(s) ds = \delta_{i,j} \quad (3.4)$$

where ξ_i , $i = 1, 2, \dots$ are independent standard Gaussian random variables. As a useful fact, the Brownian motion path $\{W(s), 0 \leq s \leq t\}$ can be decomposed as a linear combination of the Gaussian random variables ξ_i . In this work, we employ following the orthonormal bases $m_i(s)$ in $L^2([0, t])$, which are chosen as the trigonometric functions :

$$m_1(s) = \frac{1}{\sqrt{t}}, \quad m_i(s) = \sqrt{\frac{2}{t}} \cos\left(\frac{\pi(i-1)s}{t}\right) \quad k = 2, 3, \dots, 0 \leq s \leq t \quad (3.5)$$

The following theorem provides a Fourier expansion representation of Brownian motion, describing how it can be expressed as an infinite sum of deterministic basis functions weighted by random coefficients.

Theorem 1. *The Brownian motion $\{W_s; 0 \leq s \leq t\}$ has the Fourier expansion*

$$W(s) = \sum_{i=1}^{\infty} \xi_i \int_0^s m_i(\tau) d\tau \quad 0 \leq s \leq t \quad (3.6)$$

and the expansion converges in the mean square sense for all $s \leq t$

$$\mathbb{E}\left[W(s) - \sum_{i=1}^N \xi_i \int_0^s m_i(\tau) d\tau\right]^2 \leq \frac{t}{\pi N} \quad (3.7)$$

Proof. Denote the characteristic function of the interval $[0, s]$ by $\mathcal{X}_{[0,s]}(\tau) \in L_2([0, t])$ and has the

expansion

$$\mathcal{X}_{[0,s]}(\tau) = \sum_{i=1}^{\infty} (\mathcal{X}_{[0,s]}, m_i) m_i(\tau) = \sum_{i=1}^{\infty} m_i(\tau) \int_0^s m_i(\tau) d\tau$$

It follows that

$$W(s) = \int_0^s dW_\tau = \int_0^t \mathcal{X}_{[0,s]}(\tau) dW_\tau = \int_0^t \left(\sum_{i=1}^{\infty} m_i(\tau) \int_0^s m_i(\tau) d\tau \right) dW_\tau = \sum_{i=1}^{\infty} \xi_i \int_0^s m_i(\tau) d\tau$$

Denote the mean square error of the finite-term truncation

$$R(N) = \mathbb{E} \left[W(s) - \sum_{k=1}^N \xi_k \int_0^s m_k(\tau) d\tau \right]^2$$

Since ξ_i are independent unit Gaussian variables, we have

$$R(N) = \sum_{k=N+1}^{\infty} \left(\int_0^s m_k(\tau) d\tau \right)^2 = \sum_{k=N}^{\infty} \frac{2t}{k^2 \pi^2} \sin^2 \left(\frac{k\pi s}{t} \right) \leq \sum_{k=N}^{\infty} \frac{2t}{k^2 \pi^2} \leq \frac{2t}{\pi^2} \left(\frac{1}{N^2} + \int_N^{\infty} \frac{dx}{x^2} \right) \leq \frac{t}{\pi N}$$

So the expansion (3.6) converges in the mean square sense uniformly for all $s \leq t$ [98]. $\square$

To obtain recursive propagator equations, we denote the set of multi-indices as

$$\mathcal{J} = \{ \alpha = (\alpha_i; i \geq 1) | \alpha_i \in \{0, 1, 2, \dots\}, |\alpha| = \sum_{i=1}^{\infty} \alpha_i < \infty \} \quad (3.8)$$

where the set $\mathcal{J}$ is countable and for every $\alpha \in \mathcal{J}$, only finitely many of α_i are not equal to zero. For instance, $\alpha = (0, 0, 0, \dots, 0)$ is order $|\alpha| = 0$ index while $\alpha = (1, 0, 0, \dots, 0)$ or $\alpha = (0, 0, 1, \dots, 0)$ are the order $|\alpha| = 1$ indexes. Higher order indexes $\alpha = (1, 0, 1, 0, \dots, 0)$ or $\alpha = (2, 0, 0, \dots, 0)$ are order $|\alpha| = 2$ and are computed for the nonlinear problems. For $\alpha \in \mathcal{J}$, define

$$|\alpha| = \sum_i \alpha_i \text{ and } \alpha! = \prod_i \alpha_i!$$

and

$$V_\alpha(\xi) = \frac{1}{\sqrt{\alpha!}} \prod_i V_{\alpha_i}(\xi_i) \quad (3.9)$$

where V_n is the n^{th} degree Hermite polynomial. The following theorem describes how the solution $u(s, x)$ of an SPDE can be represented and analyzed using a Fourier-Hermite expansion and the properties of Wick polynomials.

Theorem 2. (Cameron-Martin) Assume that for any x and $s \leq t$, the solution $u(s, x)$ of (3.2) is a functional of the Brownian motion $\{W_s : 0 \leq s \leq t\}$ with $\mathbb{E}|u(s, x)|^2 < \infty$. Then $u(s, x)$ has the following Fourier-Hermite expansion:

$$u(s, x) = \sum_{\alpha \in \mathcal{J}} u_\alpha(s, x) V_\alpha(\xi) \quad u_\alpha(s, x) = \mathbb{E}[u(s, x) V_\alpha(\xi)] \quad (3.10)$$

where $V_\alpha(\xi)$ are the random Wick polynomials. Furthermore, the first two statistical moment of

$u(s, x)$ are given by the mean:

$$\mathbb{E}[u(s, x)] = u_0(s, x) \quad (3.11)$$

and the variance

$$\mathbb{E}[u^2(s, x)] = \sum_{\alpha \in \mathcal{J}, \alpha \neq 0} |u_\alpha(s, x)|^2 \quad (3.12)$$

The expansion (3.10) can be separated into two parts:

$$u(s, x) = \sum_{|\alpha| \leq 1} u_\alpha(s, x) V_\alpha(s, x) + \sum_{|\alpha| \geq 2} u_\alpha(s, x) V_\alpha(s, x) \quad (3.13)$$

The first part of expansion (3.13) is the Gaussian approximation of $u(s, x)$, while the second part accounts for higher-order non-Gaussian terms. It should be noted that due to the linear nature of the model in this section, we are only interested in the first part of the expansion. In the next section, we introduce a simple ocean model driven by white noise and present the first applications of WCE, comparing it with the method of characteristics.

On the other hand, we utilize Monte Carlo simulations to validate the convergence and accuracy of the results obtained from WCE for the linear stochastic ENSO model given in Section (3.5). The idea of MC simulations is to sample the random forcing and solve the model equations sample by sample. The Brownian motion path is sampled by generating Gaussian random variables $W(t_{n+1}) - W(t_n) = \sqrt{\Delta t} \mathcal{N}(0, 1)$ repeatedly. The statistical moments of the random solution are computed by averaging many realizations in the MC simulation. For example, the mean of the random solution is estimated by the MC ensemble average:

$$u_M^{MC}(t, x) = \frac{1}{M} \sum_{k=1}^M u(t, x, \omega_k) \quad (3.14)$$

where $u(t, x, \omega_k)$ is a realization computed by the MC simulations, ω_k is the sample path of Brownian motion and M is the total number of realizations [135, 138]. Moreover, the variance $u_{\text{Var}}^{MC}(t, x)$ can be calculated as

$$u_{\text{Var}}^{MC}(t, x) = \frac{1}{M} \sum_{k=1}^M [u(t, x, \omega_k) - u_M^{MC}(t, x)]^2$$

In Sections (3.5) and (3.6), we demonstrate the application of both methods using different numbers of recursive propagator equations (WCE modes) and MC ensembles.

3.4 Simple Ocean Model Driven by White Noise

An essential component of ENSO modeling is the propagation of ocean waves. To analyze and test the efficiency of the proposed WCE method, we first consider a simple ocean model governed by oceanic Kelvin and Rossby waves driven by white noise. The problem $u(t, x) = [K(t, x), R(t, x)]^T$ can be formulated as follows:

$$\mathcal{M} \frac{\partial u}{\partial t}(t, x) = \mathcal{L}(u) + \Xi \dot{W}(t) \quad \text{where} \quad \mathcal{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \mathcal{L} = \begin{bmatrix} -c \frac{\partial}{\partial x} & 0 \\ 0 & \frac{c}{3} \frac{\partial}{\partial x} \end{bmatrix} \quad \text{and} \quad \Xi = \begin{bmatrix} \sigma_K \\ \sigma_R \end{bmatrix} \quad (3.15)$$

where $K(t, x)$, $R(t, x)$ and $W(t)$ represent the oceanic Kelvin and Rossby waves and Brownian motion with a constant noise strength $\Xi = [\sigma_K, \sigma_R]^\top$, respectively. The model has the following reflective boundary conditions

$$K(t, 0) = r_W R(t, 0) \quad \text{and} \quad R(t, L) = r_E K(t, L) \quad (3.16)$$

where $r_W = 0.5$ and $r_E = 0.5$ are coupling coefficients at the western and eastern ocean boundaries, respectively. The ocean domain is chosen as $L = 1.2$. Specifically, at $x = 0$ the Rossby wave is partially reflected as a Kelvin wave with a coupling coefficient $r_W = 0.5$, and at $x = L$ the Kelvin wave is partially reflected as a Rossby wave with a coupling coefficient $r_E = 0.5$, ensuring wave interaction at the boundaries. Initial conditions that are shown explicitly in the next section are defined as:

$$K(0, x) = f(x) \quad \text{and} \quad R(0, x) = g(x) \quad (3.17)$$

Kelvin and Rossby waves play significant roles in the dynamics of the ENSO in the tropical Pacific Ocean. Kelvin waves travel eastward along the equator, constrained by the Coriolis effect, which is zero at the equator. On the other hand, Rossby waves, which travel westward and are much slower than Kelvin waves, are large-scale waves generated by the Earth's rotation and the variation of the Coriolis effect with latitude. In climate models, accurately representing these waves is essential for simulating the timing, intensity, and duration of ENSO events. These waves are also the first meridional basis projection to the originally two-dimensional PDEs- (x, y) and help modulate the sea surface temperature anomalies that are central to ENSO dynamics, as discussed in more detail in Section (3.5). In the next section, we provide some properties of the ocean model and compare solutions using the method of characteristics with those obtained numerically.

3.4.1 Discontinuities of the Homogeneous Model

To analyze the model features, we consider the homogeneous (unforced) case of equations (3.15)-(3.17) where $\sigma_K = 0$ and $\sigma_R = 0$. Due to the reflective boundary condition at the ocean boundaries, the model has discontinuities. We use a simple linear initial condition as shown in Figure (3.1). Although this initial condition may not be physical, it is one of the simplest choices that satisfies the boundary condition of the model.

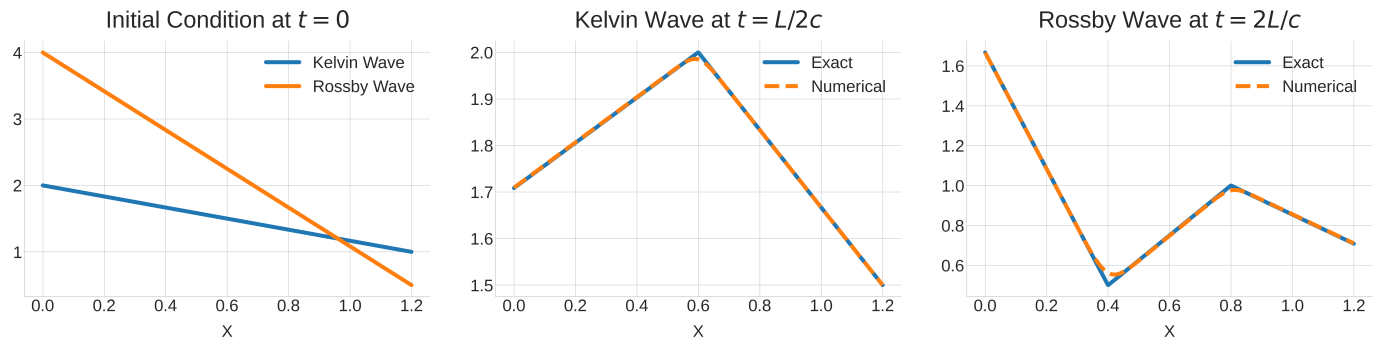


Figure 3.1: Initial conditions for Kelvin and Rossby waves at time $t = 0$ that satisfy the boundary conditions (left), Kelvin wave at time $t = L/2c$ (middle) and Rossby wave at time $t = 2L/c$ (right).

By using the method of characteristics, we observe that $K(t, x)$ has a jump discontinuity at all

times $0 < t < L/c$ and $3L/c < t < 4L/c$. At time $t = L/2c$, the exact solution can be written as:

$$K(L/2c, x) = \begin{cases} \frac{1}{2}R((-x + L/2)/3) & \text{if } 0 \leq x \leq L/2 \\ K(x - L/2) & \text{if } L/2 \leq x \leq L \end{cases} \quad (3.18)$$

Similarly, $R(t, x)$ has one jump discontinuity at times $0 < t \leq L/c$ and $3L/c < t < 4L/c$ and two jump discontinuities at times $L/c < t < 3L/c$. We have the solution for $R(t, x)$ at time $t = 2L/c$ as follows:

$$R(2L/c, x) = \begin{cases} R(x + \frac{2L}{3}) & \text{if } 0 \leq x \leq L/3 \\ \frac{1}{2}K(2L - 3x) & \text{if } L/3 \leq x \leq 2L/3 \\ \frac{1}{4}R(x - \frac{2L}{3}) & \text{if } 2L/3 \leq x \leq L \end{cases} \quad (3.19)$$

Figure (3.1) depicts the comparison of numerical solution and exact solution obtained from the characteristics of Kelvin wave at time $t = L/2c$ and Rossby wave at time $t = 2L/c$. The figures illustrate that the error between the exact and numerical solutions is remarkably small. In the next section, we consider the same problem with $\sigma_K \neq 0$ and $\sigma_R \neq 0$ and provide statistical properties.

3.4.2 Mean and Variance Through Characteristic Lines

In this section, we study the exact solution of the statistical moments of the ocean problem (3.15)-(3.17). We obtain mean and variance solutions at the terminal time and then compare these with the numerical solutions obtained from the WCE.

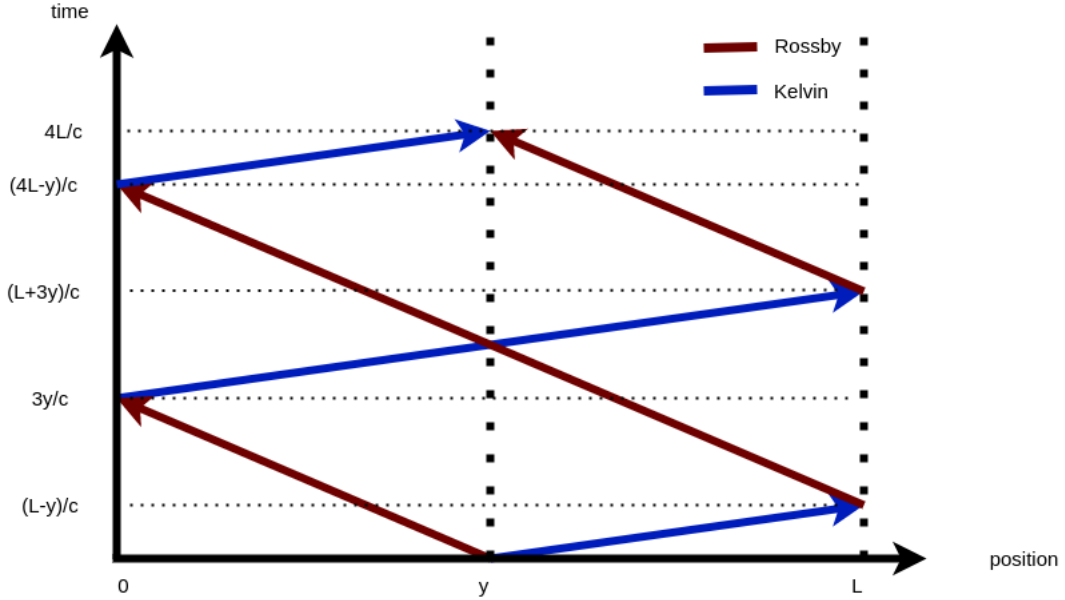


Figure 3.2: Kelvin and Rossby waves through characteristics.

Figure (3.2) illustrates the propagation of oceanic Kelvin and Rossby waves through characteristic lines. The horizontal axis represents the position of the waves in the ocean domain, while the vertical axis indicates the reflection times at the boundaries. Waves starting their motion at the point $y \in [0, L]$ return to the same point y within one-period time, calculated as $T = 4L/c$ for this

problem. At the boundaries, the waves transform into each other due to the reflective boundary conditions. The mean value of $K(4L/c, y)$ is $r_W r_E f(y)$ while the variance is given by

$$r_E^2 r_W^2 \sigma_K^2 \left(\frac{L-y}{c} \right) + r_W^2 \sigma_R^2 \left(\frac{3L}{c} \right) + \sigma_K^2 \left(\frac{y}{c} \right) \quad (3.20)$$

Similarly, the mean value of $R(4L/c, y)$ is $r_E r_W g(y)$ while the variance is given by

$$r_E^2 r_W^2 \sigma_R^2 \left(\frac{3y}{c} \right) + r_E^2 \sigma_K^2 \left(\frac{L}{c} \right) + \sigma_R^2 \left(\frac{3L-3y}{c} \right) \quad (3.21)$$

In the next section, we compare the exact solutions of the mean and variance equations (3.20)-(3.21) at the terminal time with the numerical solutions obtained from the WCE-based method.

3.4.3 Recursive Propagator Equations

In this section, we provide a detailed derivation of propagator equations of model (3.15)-(3.17), which will then be solved numerically. The solutions $K(t, x)$ and $R(t, x)$ of (3.15)-(3.17) have the following Fourier-Hermite expansion:

$$K(t, x) = \sum_{\alpha \in \mathcal{J}} K_\alpha(t, x) V_\alpha(\xi) \quad \text{and} \quad R(t, x) = \sum_{\alpha \in \mathcal{J}} R_\alpha(t, x) V_\alpha(\xi).$$

Write the SPDE's (3.15)-(3.17) in the integral form

$$\begin{aligned} K(t, x) &= f(x) - c \int_0^t \partial_x K(s, x) ds + \sigma_K W(t) \\ R(t, x) &= g(x) + \frac{c}{3} \int_0^t \partial_x R(s, x) ds + \sigma_R W(t). \end{aligned}$$

If we plug this expansion into the Equations (3.15), multiply both sides by V_β and take the expectation, we obtain

$$\begin{aligned} \sum_{\alpha \in \mathcal{J}} K_\alpha(t, x) \mathbb{E}[V_\alpha V_\beta] &= f(x) \mathbb{E}[V_\beta] - c \sum_{\alpha \in \mathcal{J}} \int_0^t \partial_x K_\alpha(s, x) \mathbb{E}[V_\alpha V_\beta] ds + \sigma_K \mathbb{E}[W(t) V_\beta] \\ \sum_{\alpha \in \mathcal{J}} R_\alpha(t, x) \mathbb{E}[V_\alpha V_\beta] &= g(x) \mathbb{E}[V_\beta] + \frac{c}{3} \sum_{\alpha \in \mathcal{J}} \int_0^t \partial_x R_\alpha(s, x) \mathbb{E}[V_\alpha V_\beta] ds + \sigma_R \mathbb{E}[W(t) V_\beta] \end{aligned}$$

Since $\mathbb{E}[V_\alpha(\xi) V_\beta(\xi)] = \delta_{\alpha\beta}$ and $\sum_\alpha K_\alpha \delta_{\alpha\beta} = K_\beta$, we get a set of infinite-dimensional equations

$$\begin{aligned} K_\beta(t, x) &= f(x) \mathbb{E}[V_\beta] - c \int_0^t \partial_x K_\beta(s, x) \mathbb{E}[V_\alpha V_\beta] ds + \sigma_K \mathbb{E}[W(t) V_\beta] \\ R_\beta(t, x) &= g(x) \mathbb{E}[V_\beta] + \frac{c}{3} \int_0^t \partial_x R_\beta(s, x) \mathbb{E}[V_\alpha V_\beta] ds + \sigma_R \mathbb{E}[W(t) V_\beta] \end{aligned}$$

Using the properties of Wick polynomials, we obtain the following recursive propagator equations for the Kelvin wave:

$$\begin{aligned} K_{|\beta|=0}(t, x) &= f(x) - c \int_0^t \partial_x K_{|\beta|=0}(s, x) ds \\ K_{|\beta|=1}(t, x) &= -c \int_0^t \partial_x K_{|\beta|=1}(s, x) ds + \sigma_K \int_0^t m_i(s) ds \\ K_{|\beta|\geq 2}(t, x) &= -c \int_0^t \partial_x K_{|\beta|\geq 2}(s, x) ds \end{aligned}$$

Similarly, the propagator equations for the Rossby wave are:

$$\begin{aligned} R_{|\beta|=0}(t, x) &= g(x) + \frac{c}{3} \int_0^t \partial_x R_{|\beta|=0}(s, x) ds \\ R_{|\beta|=1}(t, x) &= \frac{c}{3} \int_0^t \partial_x R_{|\beta|=1}(s, x) ds + \sigma_K \int_0^t m_i(s) ds \\ R_{|\beta|\geq 2}(t, x) &= \frac{c}{3} \int_0^t \partial_x R_{|\beta|\geq 2}(s, x) ds \end{aligned}$$

The order $|\beta| = 0$ propagator solution is obtained by using initial conditions with zero-forcing, while the order $|\beta| = 1$ propagator solutions are solved with forcing terms m_k and zero initial condition. Order $|\beta| \geq 2$ provide zero solution since they are solved with zero initial condition and without forcing terms. To obtain corresponding boundary conditions, we expand them using Fourier-Hermite expansion. At $x = 0$, we have the following relations:

$$K(t, 0) = \sum_{\alpha \in \mathcal{J}} K_\alpha(t, 0) V_\alpha(\xi) = r_W R(t, 0) = r_W \sum_{\alpha \in \mathcal{J}} R_\alpha(t, 0) V_\alpha(\xi).$$

Multiplying both sides by V_β and taking the expectation, one obtains

$$\sum_{\alpha \in \mathcal{J}} K_\alpha(t, 0) \mathbb{E}[V_\alpha V_\beta] = r_W \sum_{\alpha \in \mathcal{J}} R_\alpha(t, 0) \mathbb{E}[V_\alpha V_\beta].$$

Since $\mathbb{E}[V_\alpha(\xi) V_\beta(\xi)] = \delta_{\alpha\beta}$ and $\sum_\alpha K_\alpha \delta_{\alpha\beta} = K_\beta$, we get a set of infinite-dimensional equations

$$K_\beta(t, 0) = r_W R_\beta(t, 0)$$

Similarly, at $x = L$, we have the following relations:

$$R(t, L) = \sum_{\alpha \in \mathcal{J}} R_\alpha(t, L) V_\alpha(\xi) = r_E K(t, L) = \sum_{\alpha \in \mathcal{J}} K_\alpha(t, L) V_\alpha(\xi).$$

Multiplying both sides by V_β and taking the expectation, one obtains

$$\sum_{\alpha \in \mathcal{J}} R_\alpha(t, L) \mathbb{E}[V_\alpha V_\beta] = r_E \sum_{\alpha \in \mathcal{J}} K_\alpha(t, L) \mathbb{E}[V_\alpha V_\beta]$$

Since $\mathbb{E}[V_\alpha(\xi)V_\beta(\xi)] = \delta_{\alpha\beta}$ and $\sum_\alpha K_\alpha \delta_{\alpha\beta} = K_\beta$, we get a set of infinite-dimensional equations:

$$R_\beta(t, L) = r_E K_\beta(t, L)$$

To formulate the recursive deterministic propagator equations and numerical schemes, we can use differential forms and index notation as shown in (3.8). We show that the first WCE coefficients $K_{|0|}(t, x)$ and $R_{|0|}(t, x)$, which are the mean of the stochastic solutions, satisfy

$$(K_{(0,0,0,0,...)})_t + c(K_{(0,0,0,0,...)})_x = 0 \quad \text{with} \quad K_{(0,0,0,0,...)}(x, 0) = f(x) \quad (3.22)$$

$$(R_{(0,0,0,0,...)})_t - (c/3)(R_{(0,0,0,0,...)})_x = 0 \quad \text{with} \quad R_{(0,0,0,0,...)}(x, 0) = g(x) \quad (3.23)$$

with the boundary conditions

$$K_{(0,0,0,0,...)}(t, 0) = r_W R_{(0,0,0,0,...)}(t, 0)$$

$$R_{(0,0,0,0,...)}(t, L) = r_E K_{(0,0,0,0,...)}(t, L)$$

The mean $K_{|0|}$ and $R_{|0|}$ are not driven by the random forcing directly, since the forcing has zero mean. Order $|\alpha| = 1$ coefficients can be written for $\alpha = (0, 0, 0, \dots, 1, \dots)$ where the k th element is 1 as:

$$(K_{(0,0,0,...,1,...)})_t + c(K_{(0,0,0,...,1,...)})_x = m_k(t) \quad K_{(0,0,0,...,1,...)}(x, 0) = 0 \quad (3.24)$$

$$(R_{(0,0,0,...,1,...)})_t - (c/3)(R_{(0,0,0,...,1,...)})_x = m_k(t) \quad R_{(0,0,0,...,1,...)}(x, 0) = 0 \quad (3.25)$$

with the boundary conditions

$$K_{(0,0,0,...,1,...)}(t, 0) = r_W R_{(0,0,0,...,1,...)}(t, 0)$$

$$R_{(0,0,0,...,1,...)}(t, L) = r_E K_{(0,0,0,...,1,...)}(t, L)$$

The coefficients $\{ K_\alpha, R_\alpha ; |\alpha| = 1 \}$ of the Gaussian part are forced by the input $m_k(t)$ directly. We solve Equations (3.22)-(3.25) using the finite difference method and compare them with the exact solutions at the terminal time.

3.4.4 Comparison of Theoretical and Numerical Results

We simulate the model until $T = 4L/c$, which corresponds to one-period time where $L = 1.2$ and $c = 0.5$. Figure (3.3) illustrates the evolution of Kelvin and Rossby wave mean and variance, calculated using Equations (3.11)-(3.12), over one-period. We compare the mean and variance results of the exact solution (as obtained in the previous section using characteristics) with those from the WCE at the terminal time $T = 9.6$. We find reasonable agreement using 1-WCE mode for the mean and 9-WCE modes for the variance. As the mean solution decays due to dissipation at the boundaries, the variance increases over time due to stochastic forcing. Due to discontinuities in the boundary conditions, some errors are observed in the numerical solution at the boundaries for the mean calculation.

Mean and Variance of Kelvin and Rossby Waves

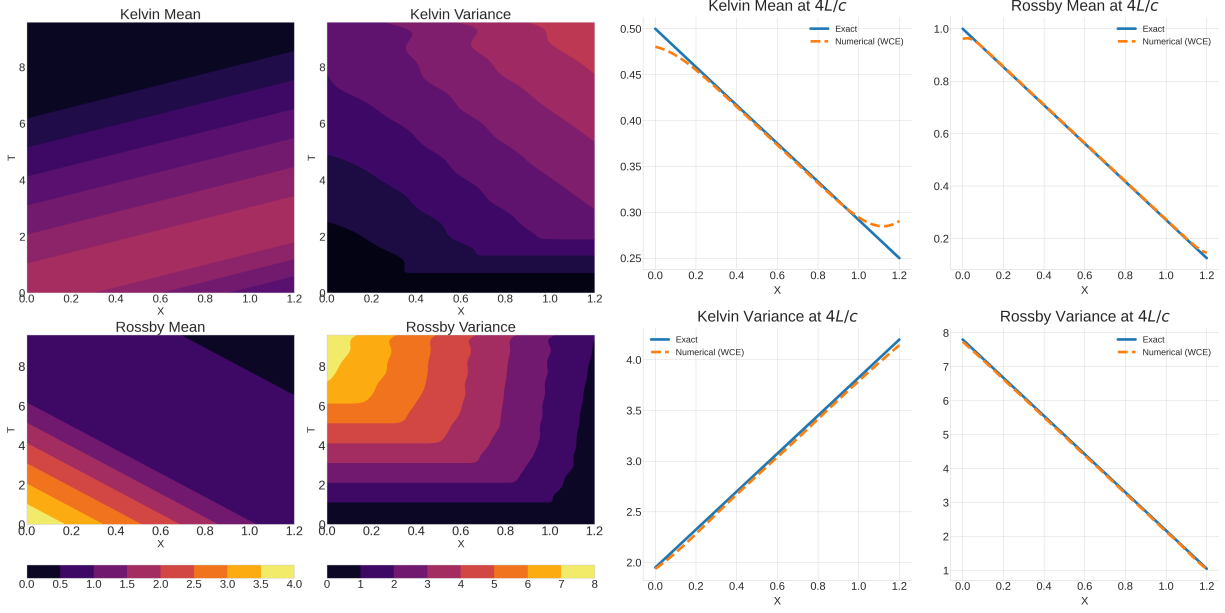


Figure 3.3: Mean and variance of Kelvin and Rossby waves over time (left) and at the terminal time $T = 4L/c$ (right).

To obtain the random solution for $u(t, x) = [K(t, x), R(t, x)]^T$, we combine the deterministic propagators $u_\alpha(t, x) = [K_\alpha(t, x), R_\alpha(t, x)]^T$ with the corresponding Hermite polynomials. Random numbers are generated by using the Python package *numpy.random.normal* with a normal distribution $\mathcal{N}(0, 1)$. Table (3.1) shows one set of random variables to obtain a complete random solution of model (3.15). After deriving the expressions for the propagators, we combine them with the corresponding random Hermite bases. Figure (3.4) displays five ensemble solutions of Kelvin and Rossby waves. Due to the stochastic nature of the solutions, each ensemble exhibits different wave evolution.

ξ_{P_0}	ξ_{P_1}	ξ_{P_2}	ξ_{P_3}	ξ_{P_4}	ξ_{P_5}	ξ_{P_6}	ξ_{P_7}	ξ_{P_8}	ξ_{P_9}
1.0000	-1.3102	-0.0109	1.7551	1.3406	0.7429	-1.2924	1.0150	0.2958	2.0311
1.0000	0.3678	-0.5975	-0.0969	0.0215	-0.3817	-0.7041	0.3456	-0.6991	-0.1101
1.0000	0.7719	-0.6630	1.3424	0.7646	-1.2890	0.1414	-1.5045	-1.1763	-1.2049
1.0000	0.0194	-0.2582	0.6723	0.8448	-1.0832	2.0494	0.4994	1.1276	-2.1716
1.0000	-0.1005	-0.1200	-0.0088	0.8219	0.3810	0.1785	1.5316	-0.3415	-0.6532

Table 3.1: Gaussian random variables multiplied by the corresponding propagator coefficients for different sets.

Stochastic Representations of Oceanic Kelvin and Rossby Waves

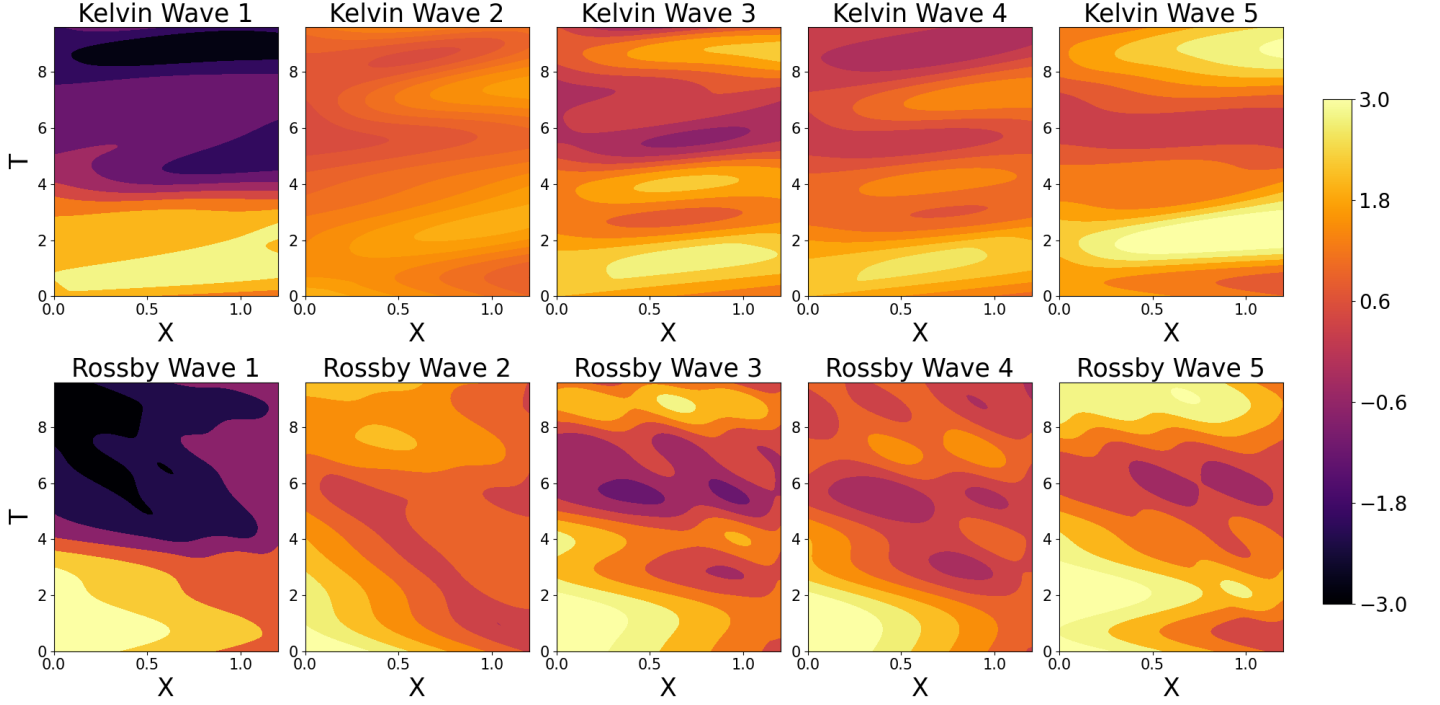


Figure 3.4: Random field solutions of oceanic Kelvin and Rossby waves generated using different stochastic bases.

In the next section, we apply the Wiener Chaos Expansion (WCE) to the linear coupled stochastic ENSO model and validate the results by comparing them with those obtained from Monte Carlo simulations.

3.5 Linear Coupled Stochastic ENSO Model

In this section, we utilize the linear stochastic ENSO model introduced by Majda and co-workers [23, 155] to show the efficiency and accuracy of the WCE method for high-dimensional coupled problems. A dissipative steady atmosphere model is driven by atmospheric Kelvin $K^A(t, x)$ and Rossby $R^A(t, x)$ waves respectively. Additionally, a dissipative shallow water ocean model forced by atmospheric waves and stochastic wind burst $\tau(t)$ is composed of oceanic Kelvin $K^O(t, x)$ and Rossby $R^O(t, x)$ waves. Final component of the coupled model is the sea surface temperature (SST) model $T(t, x)$ that characterizes the temperature anomalies. Linear coupled PDEs in six variables $u(t, x) = [K^A(t, x), R^A(t, x), K^O(t, x), R^O(t, x), T(t, x), \tau(t)]^T$ can be written in differential equation form as

$$\mathcal{M} \frac{\partial u}{\partial t}(t, x) = \mathcal{A}(\kappa) u(t, x) \quad (3.26)$$

where the linear operator $A(\kappa)$ in terms of bifurcation parameter κ , the wind stress coefficient is

$$\mathcal{A}(\kappa) = \begin{bmatrix} -\frac{\partial}{\partial x} - \gamma & 0 & 0 & 0 & \chi_A \alpha_q (2 - 2\bar{Q})^{-1} & 0 \\ 0 & \frac{1}{3} \frac{\partial}{\partial x} - \gamma & 0 & 0 & \chi_A \alpha_q (3 - 3\bar{Q})^{-1} & 0 \\ \chi_O \frac{c\kappa}{2} & -\chi_O \frac{c\kappa}{2} & -c \frac{\partial}{\partial x} - \delta & 0 & 0 & s_p(x) \\ -\chi_O \frac{c\kappa}{3} & \chi_O \frac{c\kappa}{3} & 0 & \frac{c}{3} \frac{\partial}{\partial x} - \delta & 0 & s_p(x) \\ 0 & 0 & c\eta(x) & c\eta(x) & -c\xi\alpha_q & 0 \\ 0 & 0 & 0 & 0 & 0 & -d_p \end{bmatrix}$$

and $\mathcal{M}$ is a 6×6 constant matrix with diagonals $[0, 0, 1, 1, 1, 1]$. Randomness enter the model through an Ornstein-Uhlenbeck process $\tau(t)$ that has the form:

$$\frac{\partial \tau(t)}{\partial t} = -d_p(\tau(t) - \hat{\tau}) + \sigma \dot{W}(t) \quad (3.27)$$

where d_p , σ and $\hat{\tau}$ represent dissipation, noise strength and mean of the process, respectively. The coupled model $u(t, x)$ is driven by the Brownian motion $W(t)$, making it random. We have the following boundary conditions for the atmosphere variables that are confined in the ocean domain:

$$K^A(t, 0) = e^{-\gamma(L_A - L_O)} K^A(t, L_O) \quad \text{and} \quad R^A(t, 0) = e^{3\gamma(L_A - L_O)} R^A(t, L_O) \quad (3.28)$$

where L_A and L_O represent the length of the equatorial belt and Pacific ocean, respectively. We have the following reflective boundary conditions for the ocean variables:

$$K^O(t, 0) = r_W R^O(t, 0) \quad \text{and} \quad R^O(t, L_O) = r_E K^O(t, L_O) \quad (3.29)$$

where $x = 0$ and $x = L_O$ represent the western and eastern parts of the ocean, respectively. Initial conditions for the model can also be written as:

$$K^O(0, x) = f_O(x) \quad ; \quad R^O(x, 0) = g_O(x) \quad ; \quad T(0, x) = T_O(x) \quad \text{and} \quad \tau(0) = 0 \quad (3.30)$$

Atmospheric and oceanic Kelvin and Rossby waves play crucial roles in ENSO dynamics by influencing oceanic heat distribution and large-scale circulation patterns. During an El Niño event, equatorial Kelvin waves propagate eastward along the equator, deepening the thermocline in the central and eastern Pacific. This deepening reduces upwelling of cooler subsurface waters, allowing warmer surface waters to accumulate, thereby increasing sea surface temperature (SST) in these regions. This warming further reinforces El Niño conditions by altering atmospheric convection patterns and weakening the trade winds.

In contrast, Rossby waves propagate westward and act as a feedback mechanism that modulates the influence of Kelvin waves. These waves contribute to large-scale ocean circulation changes by redistributing heat and affecting thermocline depth in the western Pacific. When Rossby waves reflect off the western boundary of the Pacific basin, they generate eastward-propagating Kelvin waves, which can influence the transition between El Niño and La Niña phases. Additionally, Rossby waves contribute to SST variability by modulating oceanic heat transport and mixing processes, ultimately playing a crucial role in the evolution and termination of ENSO events.

In our numerical simulations, we observe that the initial conditions of the model persist in the

system until time $T = 9.6$. Therefore, we simulate the model until 3-period (which is $T_f = 28.8$) to analyze both transient and steady states. The ocean is forced by the atmosphere in two ways : (i) deterministic forcing caused by the wind flow $K^A(t, x)$ and $R^A(t, x)$ and (ii) stochastic forcing created by the wind burst activities $\tau(t)$. Figure (3.5) depicts the time evolution of atmospheric forcing ($K^A(t, x) - R^A(t, x)$) at the mid-ocean with coupling coefficient $\kappa = 5.75$, dissipation $d_p = 1.0$, noise strength $\sigma = 1.2$ and mean $\hat{\tau} = -0.25$.

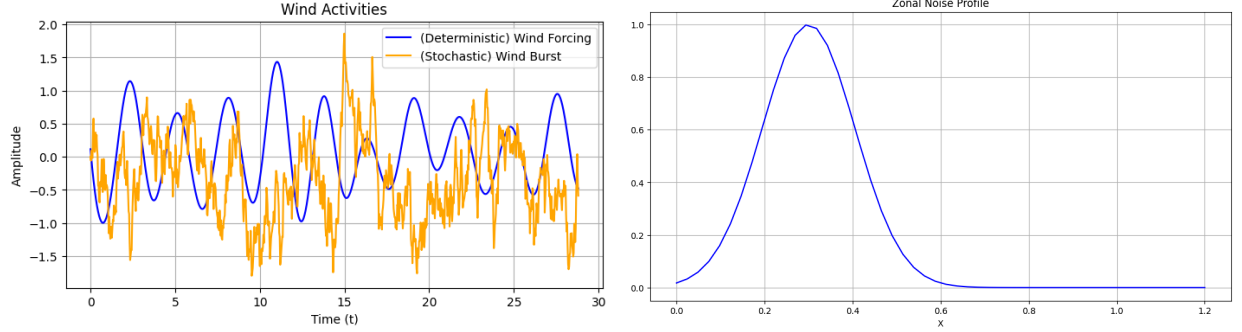


Figure 3.5: Deterministic and stochastic wind activities (left) and the noise profile confined to the western Pacific (right).

While the atmospheric winds $K^A(t, x)$ and $R^A(t, x)$ are observed all over the Pacific ocean, the stochastic component representing the wind burst activities are confined to the western Pacific as shown in the noise profile in Figure (3.5). These bursts, known as Westerly Wind Bursts (WWBs), play a crucial role in the onset and amplification of El Niño events by generating equatorial Kelvin waves that deepen the thermocline and enhance eastward transport of warm water in the Pacific. The stochastic nature of WWBs, often linked to the Madden-Julian Oscillation (MJO), introduces variability in ENSO evolution, influencing its timing and intensity [87].

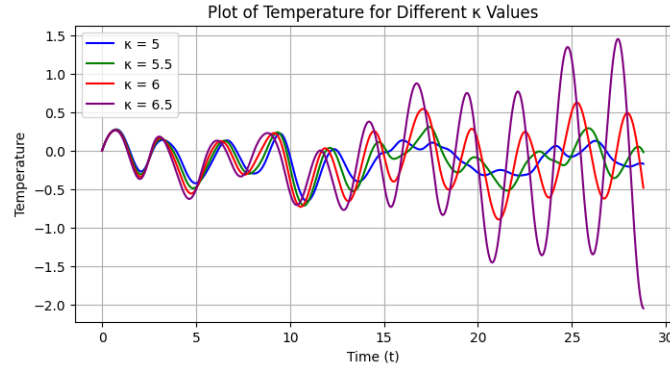


Figure 3.6: Response of the SST to different atmosphere-ocean coupling κ at the mid-ocean.

Figure (3.6) depicts the response of SST at the mid-ocean to the coupling coefficient κ between atmosphere and ocean. As the coupling strength increases, the ocean is forced by the atmosphere with higher strength. This response grows significantly after surpassing a certain threshold value of κ which acts as a bifurcation parameter.

3.6 Results

We first apply WCE to the Ornstein-Uhlenbeck (OU) process, which models stochastic wind burst activities affecting the ocean. Figure (3.7) shows the variance of the OU process for different numbers of WCE modes, along with the corresponding errors relative to the exact solution. We observe that a small number of modes, such as 2, 5, and 10, fail to capture more than 50% of the variance. In contrast, after using 40 modes, the WCE variance results converge closely to the exact values. Since 50 modes capture over 90% of the variance, we can effectively use 50 modes to represent the OU variance and coupled model, achieving a satisfactory convergence to the random field solution of the problem.

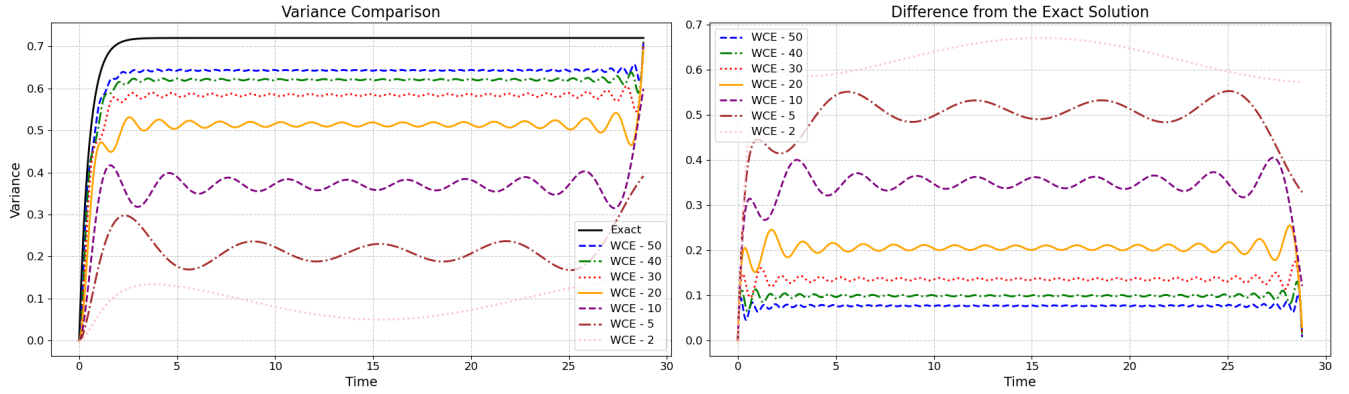


Figure 3.7: Variance of OU-process obtained from WCE with different number of modes (with $\sigma = 1.2$, $d_p = 1.0$ and $\hat{\tau} = -0.25$) (left) and $L1$ errors relative to the exact (analytical) solution (right).

To apply WCE for the coupled high-dimensional problem, we first truncate the propagator equations that can be solved by well-known numerical schemes. We use upwind finite difference schemes for the ocean and SST components and numerical integration for the atmosphere. The solutions of propagators are used to calculate the mean and variance of the model. Finally, one can construct the stochastic bases by using Gaussian random numbers and can combine them with the deterministic propagators to calculate random field solutions. The numerical scheme and spectral decomposition of the problem (3.26) are summarized in Table (3.2).

Numerical Algorithm for Wiener Chaos Expansion

1. Choose a two-way of truncation for the (deterministic) propagator equations and random basis functions.
2. Truncate the number of propagator equations $K \geq 1$.
3. Define grid points in the $(t, x) \in [0, T] \times U$ for the appropriate finite difference scheme
4. In this problem, we use $\Delta t \approx 3.0 \times 10^{-2}$ and $\Delta x \approx 2.4 \times 10^{-2}$ and simulate the model with $t \in [0, 28.8]$ and $x \in [0, 1.2]$.
5. For each time $t_i = i\Delta t$ and $x_j = j(\Delta x_1, \dots, \Delta x_n)$ solve the deterministic propagator system of equations.
6. From the propagator solutions, construct the statistical moments, i.e. mean and variance.
7. (optional) Generate random variable ξ_k , $k = 1, 2, \dots$ and compute the random field solution

$$\begin{aligned}
 K^A(t, x) &= \sum_{\alpha \in \mathcal{J}} K_\alpha^A(t, x) V_\alpha(\xi) \quad ; \quad R^A(t, x) = \sum_{\alpha \in \mathcal{J}} R_\alpha^A(t, x) V_\alpha(\xi) \quad ; \quad T(t, x) = \sum_{\alpha \in \mathcal{J}} T_\alpha(t, x) V_\alpha(\xi) \\
 K^O(t, x) &= \sum_{\alpha \in \mathcal{J}} K_\alpha^O(t, x) V_\alpha(\xi) \quad ; \quad R^O(t, x) = \sum_{\alpha \in \mathcal{J}} R_\alpha^O(t, x) V_\alpha(\xi) \quad ; \quad \tau(t) = \sum_{\alpha \in \mathcal{J}} \tau_\alpha(t) V_\alpha(\xi)
 \end{aligned}$$

Table 3.2: Numerical scheme for the solution of deterministic propagators combined with the appropriate random basis to obtain the WCE solution of SPDEs.

For example, the randomness enters through the Orsntein-Uhlenbeck process. If we plug in the expansion using integral form and take the inner product with Wick polynomial V_β , we obtain the following expression:

$$\sum_{\alpha \in \mathcal{J}} \tau_\alpha(t) \mathbb{E}[V_\alpha V_\beta] = \tau(0) \mathbb{E}[V_\beta] - d_p \int_0^t \sum_{\alpha \in \mathcal{J}} \tau_\alpha(s) \mathbb{E}[V_\alpha V_\beta] ds + d_p \int_0^t \hat{\tau} ds + \sigma \mathbb{E}[W(t) V_\beta] \quad (3.31)$$

Using the properties of the Wick polynomial (or Hermite polynomials), we obtain $|\alpha| = 0$ and $|\alpha| = 1$ order expressions as follows:

$$\tau_{|\beta|=0}(t) = \tau(0) - d_p \int_0^t \tau_{|\beta|=0}(s) ds + d_p \int_0^t \hat{\tau} ds \quad (3.32)$$

$$\tau_{|\beta|=1}(t) = -d_p \int_0^t \tau_{|\beta|=1}(s) ds + \sigma \int_0^t m_i(s) ds \quad (3.33)$$

The order $|\beta| = 0$ equation (3.32) consists of the initial condition and deterministic forcing while the order $|\beta| = 1$ equation (3.33) involves a combination of L^2 functions $m_i(s)$ as given in (3.5).

$$\tau_{|\beta| \geq 2}(t) = -d_p \int_0^t \tau_{|\beta| \geq 2}(s) ds \quad (3.34)$$

The order $|\beta| \geq 2$ Equations (3.34) provide zero solutions, as they involve no initial condition or forcing terms. Equations (3.32)-(3.34) contribute to corresponding components of $u(t, x) = [K^A(t, x), R^A(t, x), K^O(t, x), R^O(t, x), T(t, x)]$ by forming a system of recursive deterministic propagator equations.

3.6.1 Comparison of Wiener Chaos Expansion and Monte Carlo Simulations

WCE can provide significant advantages in terms of convergence rate, efficiency, and analytical insights for many problems involving uncertainty quantification. However, for extremely high-dimensional, nonlinear and discontinuous problems, WCE can become infeasible. Moreover, to analyze models where analytical or exact solutions are not easily available, such as stochastic ENSO models, one needs to utilize other numerical techniques to confirm convergence and efficiency.

For this purpose, the coupled model (3.26) can be numerically solved using the Monte Carlo method. Monte Carlo methods are straightforward to implement and require minimal modifications to the underlying deterministic solver. This simplicity makes them accessible and widely applicable over WCE-based methods. MC methods provide direct estimates of statistical quantities such as mean, variance, and higher-order moments through sampling. Although each method has its own advantages and drawbacks, a combination of them provides numerical proof of the convergence and accuracy of simulations.

To analyze the convergence behavior of the MC simulation, we compute different numbers of ensembles as $M \in [10, 50, 100, 300]$. The relative error of the mean is computed as

$$\epsilon_u(M) = \frac{\|u_M^{MC}(t, x) - \mathbb{E}[u(t, x)]\|}{\|\mathbb{E}[u(t, x)]\|} \quad (3.35)$$

where $\mathbb{E}[u(t, x)]$ is the benchmark mean computed by WCE simulation. We also calculate the variance error using WCE and MC calculations. The variance of Monte Carlo ensembles is calculated as

$$\sigma_u^{MC}(M) = \frac{(u_M^{MC}(t, x) - \mathbb{E}[u(t, x)])^2}{M} \quad (3.36)$$

and the relative variance error is calculated as

$$\epsilon_\sigma(M) = \frac{\|\sigma_u^{MC}(M) - \sigma_u^{WCE}\|}{\|\sigma_u^{WCE}\|} \quad (3.37)$$

Table (3.3) shows the relative error of WCE modes and MC ensembles for sea surface temperature $T(t, x)$ computation times for simulating the ensembles of the coupled model. The benchmark is computed with WCE mean and variance solutions. We observe that the mean and variance errors decrease as the number of Monte Carlo ensembles increases. However, the computation time dramatically increases depending on the ensembles. On the other hand, the computation time of the WCE solution takes remarkably shorter time with small mean and variance error.

Relative Error Comparison: MC Ensembles vs. WCE			
Method	Mean Error	Computation Time	Variance Error
10 MC Ensembles	0.3792	168 sec	0.4150
50 MC Ensembles	0.0853	798 sec	0.1241
100 MC Ensembles	0.0645	1627 sec	0.1066
300 MC Ensembles	0.0635	4880 sec	0.0616
50 WCE Modes	-	801 sec	-

Table 3.3: Comparison of relative mean and variance errors for MC ensembles and WCE method for sea surface temperature $T(t, x)$, including computation times on a common laptop.

Figure (3.8) shows the comparison of convergence of MC results of $T(t, x)$ to the WCE at a fixed space point $x = 0.6$ where $x \in [0, 1.2]$. As the number of ensembles increases, the $L1$ error between the two methods decreases. However, the computational cost of the MC method is considerably higher compared to the WCE method.

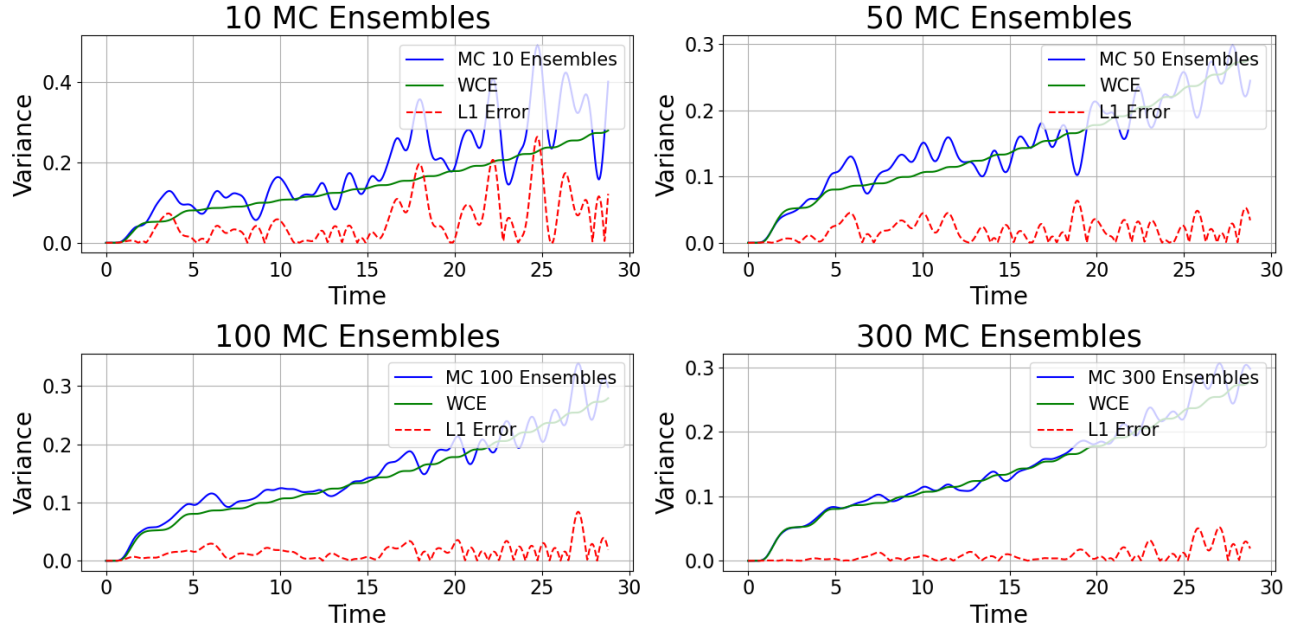


Figure 3.8: Comparison of WCE and MC variance results of $T(t, x)$ with different number of ensembles at a fixed space point $x = 0.6$.

Figures (3.9), (3.10) and (3.11) depict the mean and variance results of oceanic Kelvin wave $K^O(t, x)$, $R^O(t, x)$ and $T(t, x)$ using both WCE and MC methods. We only need the first (unforced) propagator solution for the WCE mean results, which takes 15 seconds. In contrast, the mean results for the MC simulations are obtained using 300 ensembles, requiring 4880 seconds. Although the $L1$ error between these two methods is relatively small, the WCE method requires significantly less computation time to obtain the mean and variance solutions. As we make longer simulations, the error increases due to numerical inaccuracies. The variance of oceanic Kelvin and Rossby waves is confined in the first half of the ocean due to noise profile $s_p(x)$. The SST variance shifts towards the eastern part of the ocean, resulting in warmer water initiating El Niño events. These results have been published in *Tellus A: Dynamic Meteorology and Oceanography* [5], with earlier findings presented at the European Geosciences Union (EGU) in Vienna in 2023 [4].

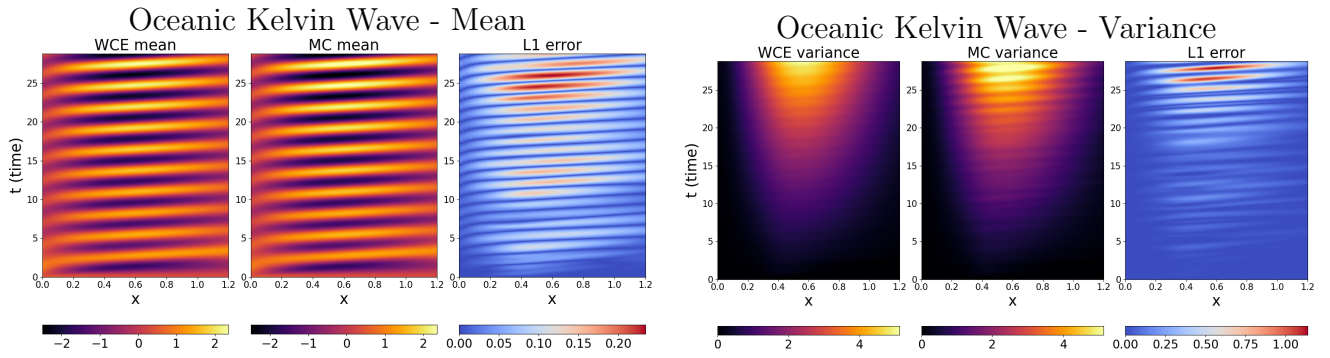


Figure 3.9: (a) Mean of oceanic Kelvin wave $K^O(t, x)$ obtained from WCE (1st propagator), MC (300 ensembles), and the corresponding $L1$ error. (b) Variance of oceanic Kelvin wave obtained from WCE (50 propagators), MC (300 ensembles).

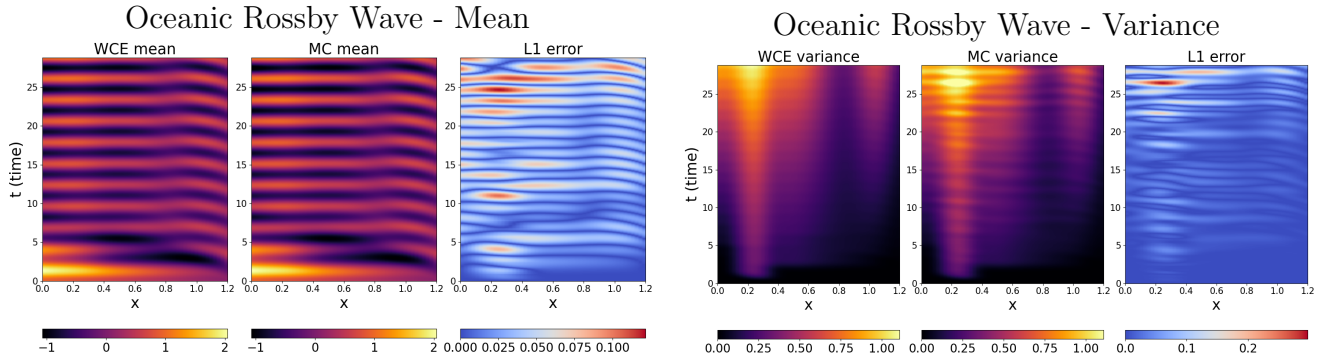


Figure 3.10: (a) Mean of oceanic Rossby wave $R^O(t, x)$ obtained from WCE (1st propagator), MC (300 ensembles), and the corresponding $L1$ error. (b) Variance of oceanic Rossby wave obtained from WCE (50 propagators), MC (300 ensembles).

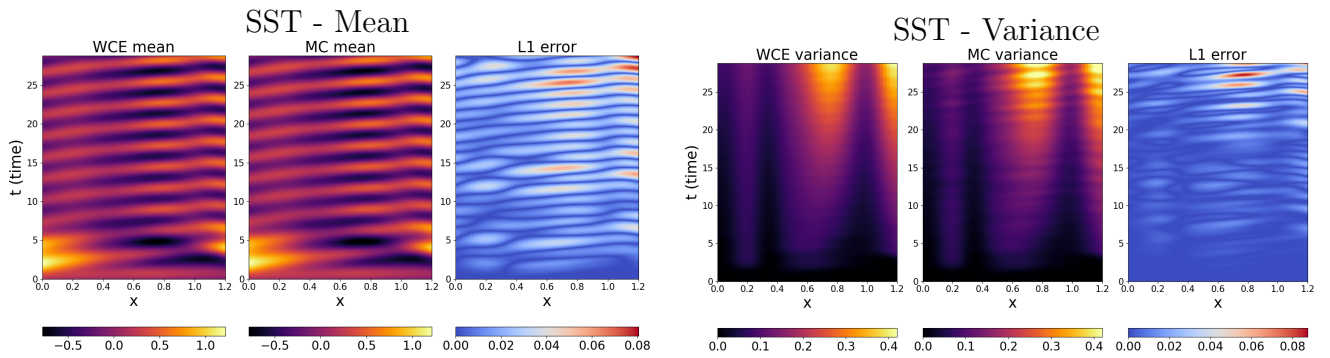


Figure 3.11: (a) Mean of sea surface temperature (SST) $T(t, x)$ obtained from WCE (1st propagator), MC (300 ensembles), and the corresponding $L1$ error. (b) Variance of sea surface temperature (SST) obtained from WCE (50 propagators), MC (300 ensembles).

3.7 Conclusion

In this chapter, we explored a stochastic ENSO model by the numerical simulations based on Wiener chaos expansion and Monte Carlo methods. The utilization of WCE provided a comprehensive study based on a spectral approach while MC simulations enhanced our ability to analyze stochastic models through ensemble methods. Although WCE and MC-based methods have been applied to many problems, this chapter investigated a relatively uncommon application on high-dimensional coupled SPDE models.

Initially, we applied the WCE-based scheme on the simple ocean model formed by oceanic Kelvin and Rossby waves forced with white noise. The WCE approach with 10 propagators converged exact mean and variance values obtained from the method of characteristics with high accuracy. These results demonstrated that the WCE method would be an appropriate method for high-dimensional and more complicated models. Subsequently, we utilized a linear stochastic ENSO model forced by the Ornstein-Uhlenbeck process to show the accuracy of the method. We compared the WCE results formed by 50 modes, which is assumed to be the benchmark for the solution, with 300 MC ensembles. We showed that the mean and variance results of the stochastic model based on the MC method converges to the WCE results. Notably, the model simulation with WCE required for less computational resources as opposed to MC ensembles. Along with the noise construction with specific model parameters, this chapter provides an alternative approach for simulations of stochastic climate models with less computational resources that can be considered as a first step of more complex and longer simulations.

However, the fact that WCE provides several advantages both numerical and theoretical, it has its drawbacks. As WCE is applied to particularly high-dimensional nonlinear problems, it becomes more computationally expensive due to the increasing number terms in the chaos expansions. The simulation of stochastic models involving non-Gaussian processes through WCE may not be the most efficient or effective approach. Depending on the type of problem, MC or other numerical methods may be more suitable.

Chapter 4

Reduced-Order Modeling (ROM) of the El Niño-Southern Oscillation (ENSO)

4.1 Reduced Order Modeling (ROM)

Simulations of complex fluid dynamics problems or climate models take weeks to complete even when run parallel in state-of-the-art supercomputers. Given computational resource constraints and the need for adaptable simulation settings, cost-efficient and accurate algorithms are essential. To overcome these difficulties, reduced order modeling (ROM) aims to reduce the computational complexity and present high-dimensional problems (usually PDEs) with reduced order equations (ODEs). Among the prominent techniques, one effective data-driven approach is dynamic mode decomposition (DMD), introduced by Peter Schmid in 2010. DMD decomposes spatiotemporal data into modes with distinct temporal dynamics, providing a linear approximation of the system's behavior [68, 81, 137, 142]. A closely related concept is the Koopman operator, which provides a linear representation of nonlinear dynamical systems by projecting them onto an infinite-dimensional function space. This operator has gained attention for its ability to extract global dynamic features, making it valuable for prediction and control of nonlinear systems [113, 129, 167]. While DMD and the Koopman operator provide linear approximations, manifold learning techniques, such as Isomap developed by Tenenbaum and co-workers and diffusion maps proposed by Coifman and Lafon uncover the low-dimensional latent space, providing a nonlinear perspective on ROM by capturing intrinsic geometric structures within high-dimensional data [25, 32, 153]. Additionally, autoencoders, a class of deep learning architectures, learn compact, nonlinear latent representations by encoding data into a reduced space and reconstructing it, offering flexibility for complex systems [10, 53, 158].

Building on these foundational methods, modern advancements in ROM tackle challenges in strongly nonlinear and high-dimensional systems. For instance, DMD and Koopman operator approaches excel in analyzing fluid flows or oscillatory systems, where global modes dominate. Manifold learning, on the other hand, is particularly effective for systems exhibiting nonlinear structures, as it provides insights that linear methods may overlook. Autoencoders introduce deep learning to ROM, enabling the discovery of intricate patterns and reconstructions even in chaotic systems. Hybrid approaches, such as integrating Koopman theory with autoencoders, enhance ROM capabilities by leveraging the strengths of both linear and nonlinear techniques, making them effective

for high-dimensional and nonlinear systems.

Proper Orthogonal Decomposition (POD) is another widely used ROM technique that identifies the most energetic modes of a system. These modes are optimal in capturing the system's dominant features, making POD particularly effective for reducing the dimensionality of systems governed by partial differential equations, such as fluid flows or heat transfer. Within the POD framework, the governing equations are projected onto a reduced basis using Galerkin projection, yielding a computationally efficient approximation of the original system. However, since POD prioritizes energy-dominant modes and is inherently linear, it may struggle to capture complex nonlinear dynamics. Despite its limitations in handling strong nonlinearities, POD remains a cornerstone of ROM due to its simplicity and strong theoretical foundation.

POD has its origins in early 20th-century mathematical and statistical developments. It is closely linked to Principal Component Analysis (PCA), introduced by Pearson (1901) [126] and later formalized by Hotelling (1933) [56]. PCA identifies dominant modes in data by analyzing covariance matrices, a concept that later influenced POD. A few decades later, the Karhunen–Loève (KL) [72, 92] theorem emerged as a significant mathematical framework in stochastic processes. Developed by Karhunen and Loève in the 1940s, the KL theorem describes a way to represent random processes through orthogonal components, forming a foundation for what would later become POD. The development of PCA and the KL theorem laid the groundwork for applying orthogonal decomposition techniques to both deterministic and random systems.

In the analysis of functional data of random processes, each observation is a curve on a functional space assumed to be sampled from a stochastic process $X \in L^2[0, T]$ - the Hilbert space of square integrable functions on $[0, T]$ with the usual inner product $\langle \varphi, \psi \rangle \stackrel{\text{def}}{=} \int_0^T \varphi(s)\psi(s)ds$ for any two functions $\varphi, \psi \in L^2[0, T]$. As the primary theoretical foundation of POD, the Karhunen–Loève expansion of X , under certain technical conditions, is assumed to exist such that

$$X(t) = \mu_X(t) + \sum_{j=1}^{\infty} \xi_j \varphi_j(t),$$

where the mean function $\mu_X(t) \stackrel{\text{def}}{=} [X(t)]$ and the basis functions $\varphi_j(t)$ are the orthonormal eigenfunctions of the covariance kernel $\Gamma_X(s, t) = \text{Cov}[X(s), X(t)]$ [72, 92]. The eigenvalues corresponding to $\varphi_j(t) \in L^2[0, T]$ satisfy

$$\int_0^T \Gamma_X(s, t) \varphi_j(s) ds = \lambda_j \varphi_j(t) \quad \text{and} \quad \Gamma_X(s, t) = \sum_{j=1}^{\infty} \lambda_j \varphi_j(s) \varphi_j(t).$$

The random coefficients ξ_j are given by the projection of $X(t) - \mu_X(t)$ in the direction of the j^{th} eigenfunction φ_j , that is,

$$\xi_j \stackrel{\text{def}}{=} \langle X(t) - \mu_X(t), \varphi_j \rangle.$$

Furthermore, $\{\xi_j\}$ are uncorrelated sequences of random variables with zero mean and variance λ_j . Since the process $X \in L^2[0, T]$, we have $\sum_{j=1}^{\infty} \lambda_j < \infty$.

In the 1960s and 1970s, POD began to gain recognition in fluid dynamics and turbulence research. Lumley [97] introduced an essential extension to POD by incorporating spatial velocity correlations, significantly impacting fluid dynamics research. The fluid flows are assumed to be turbulent and stationary in time. Under these conditions, researchers computed ensemble averages to derive spatial velocity correlations. The orthonormal eigenfunctions of the velocity correlations were then used as a rational and quantitative basis of identifying coherent structures. Lumley’s work highlighted POD’s ability to identify dominant energy-carrying modes in fluid flows, offering insights into complex phenomena like turbulence. This application of POD to physical systems marked its transition from a purely mathematical tool to a practical method in engineering and applied sciences. By the 1980s, POD was extended to computational mechanics. Sirovich [147] introduced the “snapshot method”, which made POD computationally feasible for high-dimensional systems by utilizing time-series data from simulations or experiments. Recent advancements further highlight POD’s role in fluid dynamics and reduced-order modeling. For example, [43] applied POD-Galerkin methods for large eddy simulation (LES) filtering, while [141] and [160] explored closure models for turbulent flows, demonstrating the continued evolution of POD techniques.

Although solutions are defined on the interval $[0, T]$ they are seldom observed there; instead they are observed on a discrete subset of points. The POD method of snapshots, where the data are taken at different times as a way to efficiently determine the $n \ll \infty$ basis functions, was first suggested by Sirovich [147] to study fluid dynamic applications for extracting coherent or dominant structures. The only requirement is that the snapshots are linearly independent (or data generated by ergodic dynamical systems). Given a function $w(t, x)$, the POD seeks a decomposition of the form

$$w(t, x) \approx \sum_{\alpha=1}^n a_{\alpha}(t) \psi_{\alpha}(x)$$

with spatial basis functions $\psi_{\alpha}(x)$ (also called modes), orthogonal in the spatial domain, and temporal coefficients $a_{\alpha}(t)$, capturing the time evolution of the function. n is the number of retained modes, chosen based on the desired level of approximation.

An open question is how a given functional data set can effectively be spanned by finite basis functions *without the assumption of ergodicity*. There is little in the current literature that provides a sufficient answer to this question with a notable exception of Djouadi and Sahyoun [31]. Their work explicitly investigates the theoretical properties of dimension reduction techniques within a functional analytic framework. We extend that conceptual framework to explore dimension reduction techniques in this chapter. The techniques presented in [31] circumvent one of the confounding issues that the snapshots are linearly independent. In addition, we develop the theoretical properties of the techniques within a functional analytic framework to identify an optimal set of basis functions for vector valued coupled nonlinear PDEs in one spacial dimension. These results provide a systematic approach to constructing the basis functions for model reduction.

Along with POD, Galerkin approximations are used to build weak solutions of the PDEs by constructing solutions of certain finite-dimensional (n) approximations to the PDEs. Many ROM techniques in fluid mechanics arise from the POD-Galerkin projection approach [3, 147]. Suppose

we are given the PDE in the following form

$$\frac{\partial w}{\partial t}(t, x) = \mathcal{A}w(t, x) + \mathcal{N}(w, w)(t, x) \quad (4.1)$$

The Galerkin projection reduces the PDE model to a set of reduced-order equations (ODEs) by projecting the governing equations onto a subspace spanned by the reduced basis functions, resulting in the following form:

$$\dot{a}_\gamma(t) = \sum_{\alpha=1}^n a_\alpha(t) \langle \mathcal{A}\psi_\alpha, \psi_\gamma \rangle + \sum_{\alpha, \beta=1}^n a_\alpha(t) a_\beta(t) \langle \mathcal{N}(\psi_\alpha, \psi_\beta), \psi_\gamma \rangle \quad \gamma = 1, \dots, n$$

In this chapter, we explore dimensionality reduction and model discovery methods applied to the nonlinear ENSO model. The primary focus is on POD-Galerkin reduced-order modeling, where we also introduce the main theorem developed by [31] and later extended by [9]. We present a stable, efficient, and accurate reduced-order ENSO model using POD-Galerkin method. Additionally, we investigate model reduction using neural networks and model identification using sparse identification of nonlinear dynamics (SINDy) method as proposed by [16], to compare different reduction approaches.

This chapter is organized as follows. We describe the nonlinear deterministic ENSO model in Section (4.2). We introduce the POD method with the main theorem in Section (4.3). This is followed by the POD-Galerkin ROM in Section (4.4) where we discuss efficiency, accuracy and stability with the main results. In addition, we study neural networks in Section (4.5) and SINDy method in Section (4.6), where we provide additional results. Finally, the discussion of the conclusion of the proposed methods and future work occurs in Section (4.7).

4.2 Nonlinear El Niño-Southern Oscillation Model

Let $\Omega := \{x \in [0, L]\}$ denote the extent of the equatorial Pacific Ocean (the physical space between Australia and South America) and the time domain be defined with a terminal time T as $\mathbb{T} := \{t \in [0, T]\}$. Then the nonlinear coupled PDEs in five variables which model the interaction of Kelvin and Rossby waves with SST anomalies are defined as

$$u(t, x) \stackrel{\text{def}}{=} [K^A(t, x), R^A(t, x), K^O(t, x), R^O(t, x), T(t, x)]^T$$

and can be written in differential equation form as

$$\mathcal{M} \frac{\partial u}{\partial t}(t, x) = \mathcal{A}(\kappa)u(t, x) + \mathcal{N}(u, u, \mu)(t, x), \quad \text{in } \mathbb{T} \times \Omega \quad (4.2)$$

where $\mathcal{M}$ is a 5×5 constant matrix with diagonals $[0, 0, 1, 1, 1]$, the linear operator $\mathcal{A}(\kappa)$ defined as

$$\mathcal{A}(\kappa) = \begin{bmatrix} -\frac{\partial}{\partial x} - \gamma & 0 & 0 & 0 & \chi_A \alpha_q (2 - 2\overline{Q})^{-1} \\ 0 & \frac{1}{3} \frac{\partial}{\partial x} - \gamma & 0 & 0 & \chi_A \alpha_q (3 - 3\overline{Q})^{-1} \\ \chi_O \frac{c\kappa}{2} & -\chi_O \frac{c\kappa}{2} & -c \frac{\partial}{\partial x} - \delta & 0 & 0 \\ -\chi_O \frac{c\kappa}{3} & \chi_O \frac{c\kappa}{3} & 0 & \frac{c}{3} \frac{\partial}{\partial x} - \delta & 0 \\ 0 & 0 & c\eta(x) & c\eta(x) & -c\xi \alpha_q \end{bmatrix} \quad (4.3)$$

depends on a parameter κ , the wind stress coefficient, and nonlinearity $\mathcal{N}(u, u, \mu)$ with advection coefficient μ , is defined by

$$\mathcal{N}(u, u, \mu) \stackrel{\text{def}}{=} [0, 0, 0, 0, -\mu \frac{\partial}{\partial x} [(K^O(t, x) - R^O(t, x))T(t, x)]^T.$$

In equation (4.2), small damping terms γ, δ are introduced to guarantee a unique solution, E_q is latent heat, ξ is the latent heating exchange coefficient, $\bar{Q}$ is a constant representing mean vertical moisture gradient, $\eta(x)$ is the profile of thermocline feedback and the projection coefficients are defined as

$$\chi_A = \int_{-\infty}^{\infty} \phi_0(y) \psi_0(Y) dy \quad \text{and} \quad \chi_O = \int_{-\infty}^{\infty} \psi_0(Y) \phi_0(y) dY,$$

where ϕ_0 and ψ_0 denote the atmospheric and oceanic parabolic cylinder functions dependent on meridional position y and $Y = y/\sqrt{c}$.

The initial conditions for the oceanic variables of (4.2) are given by

$$u_3(x, 0) = f_3(x), \quad u_4(x, 0) = f_4(x) \quad \text{and} \quad u_5(x, 0) = f_5(x). \quad (4.4)$$

Periodic boundary conditions that were adopted for the atmosphere model in [23] are modified to reflect the inclusion of the small damping term γ . Boundary conditions for the atmosphere variables that are confined in the ocean domain are:

$$K^A(t, 0) = e^{-\gamma(L_A - L_O)} K^A(t, L_O) \quad \text{and} \quad R^A(t, 0) = e^{3\gamma(L_A - L_O)} R^A(t, L_O) \quad (4.5)$$

where L_A and L_O represent the length of the equatorial belt and Pacific ocean, respectively. Having defined the ocean domain, the next step is to specify the boundary conditions that govern wave interactions and energy exchange as follows:

$$K^O(t, 0) = r_W R^O(t, 0) \quad \text{and} \quad R^O(t, L_O) = r_E K^O(t, L_O) \quad (4.6)$$

where $x = 0$ and $x = L_O$ represent the western and eastern ends of the ocean, respectively. The unidirectional Kelvin and Rossby oceanic equatorial waves are coupled through the boundary conditions (4.6), where the oceanic waves are reflected with coupling coefficients r_W that represents partial loss of energy in the western Pacific boundary and r_E that represents partial loss of energy due to the north-south propagation of the coastal Kelvin waves along the eastern Pacific boundary.

Equation (4.2) consists of a simple nonlinear advection $-\mu \frac{\partial}{\partial x} [(K^O(t, x) - R^O(t, x))T(t, x)]$ in the SST equation. We assume no heat flux at the eastern boundary, as specified by the following boundary conditions for $T(t, x)$:

$$\frac{\partial T}{\partial x} \Big|_{x=L_O} = 0$$

Suppose $\{u(t, x) : 0 \leq t \leq T, x \in \Omega\}$ is the solution of the coupled PDE (4.2), defined on some domain Ω and taking values in some Euclidean space $\mathbb{R}^5$. Assume the solution $u(t, \cdot) \in L^2([0, T], H)$ where H is some Hilbert space of functions $\Omega \rightarrow \mathbb{R}^5$ such as an L^2 (square-integrable vector functions u on Ω) or $H^1(D) \subset L^2(D)$ which denotes the Sobolev space of u of order one that satisfies the boundary conditions (4.5) and (4.6).

4.3 Proper Orthogonal Decomposition (POD)

In this section, the theoretical properties of the POD techniques within a functional analytic framework are established. That framework was originally introduced by Djouadi and Sahyoun [31] and later extended by Baxendale [9]. That approach aims to identify an optimal set of basis functions for vector-valued, coupled nonlinear PDEs. We present below the results in [9], which provide an efficient approach for constructing basis functions to be used in model reduction.

Suppose $\{w(t, x) : 0 \leq t \leq T, x \in \Omega\}$ is the solution of some PDE, defined on some domain Ω and taking values in some Euclidean space $\mathbb{R}^d$. Assume the solution $w(t, \cdot) \in L^2([0, T], H)$, where H is some Hilbert space of functions $\Omega \rightarrow \mathbb{R}^d$ such as an L^2 or Sobolev space. In order to construct a reduced order model (ROM) we first look for a finite-dimensional subspace H_0 of H and a function $\tilde{w} : [0, T] \rightarrow H_0$ taking values in H_0 so that the L^2 difference

$$\int_0^T \|w(t, \cdot) - \tilde{w}(t)\|_H^2 dt$$

is small. The selection of H_0 , or equivalently the set of POD modes forming a basis for H_0 , is determined solely based on observations of the solution $w(t, x)$. The information from the PDE will later be used to derive an approximate ODE governing the finite-dimensional process. $\{\tilde{w}(t) : 0 \leq t \leq T\}$.

To formulate a computationally suitable framework, assume that we have snapshots $\{w(x, t_i); 1 \leq i \leq M\}$ of the function $w(t, x)$ taken at times t_i evenly distributed over the interval $[0, T]$. That is, instead of continuous observations, we have elements $w(t_i) = w(t_i, \cdot) \in H$ for $1 \leq i \leq M$, and we look for a finite-dimensional subspace H_0 of H and a sequence $\tilde{w}_1, \dots, \tilde{w}_M$ taking values in H_0 so that the L^2 difference

$$\sum_{i=1}^M \|w(t_i) - \tilde{w}_i\|_H^2$$

is small. Specifically, given $n \geq 1$ let $\mathcal{S}_n$ be the set of all functions $F : \{1, 2, \dots, M\} \rightarrow H$ of the form

$$F_i = \sum_{\alpha=1}^n a_i^\alpha \psi_\alpha$$

where all the vectors a^α are in $\mathbb{R}^M$ and all the ψ_α are in H . Thus F_i takes values in the linear span $H_0 = \text{span}\{\psi_1, \dots, \psi_n\}$ which has dimension at most n . Define

$$\mu_n = \inf_{F \in \mathcal{S}_n} \left(\sum_{i=1}^M \|w(t_i) - F_i\|_H^2 \right). \quad (4.7)$$

Theorem 3. *Given the snapshots $w(t_1), \dots, w(t_M)$ define the $M \times M$ matrix C by*

$$C_{ij} = \langle w(t_i), w(t_j) \rangle_H, \quad 1 \leq i, j \leq M. \quad (4.8)$$

The matrix C is symmetric and non-negative, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ (counted with multiplicities) and corresponding orthonormal eigenvectors $\gamma^1, \gamma^2, \dots$ in $\mathbb{R}^M$. Then the minimization problem (4.7) has solution $\mu_n = \sum_{\alpha > n} \lambda_\alpha$ and the optimal F is given by $F_i = \sum_{\alpha=1}^n a_i^\alpha \psi_\alpha$ where

$\{\psi_1, \dots, \psi_n\}$ is orthonormal and is given by

$$\psi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^M \gamma_i^\alpha w(t_i), \quad 1 \leq \alpha \leq n,$$

and

$$a_i^\alpha = \sqrt{\lambda_\alpha} \gamma_i^\alpha, \quad 1 \leq i \leq M, 1 \leq \alpha \leq n.$$

Theorem (3) plays a crucial role in our research, providing the foundation for computing the POD modes ψ_α and the corresponding time coefficients a^α using the eigenvectors γ^α eigenvalues λ^α of matrix C and snapshots $w(t_i)$. The proof is provided in the next section.

4.3.1 Proof of Main Theorem

This section presents the proof of Theorem (3), which serves as the theoretical foundation for POD. To derive the proof step by step following an analytic framework, operators Γ and Γ^* are defined as an initial stage. In Lemma (1), it is shown that $\Gamma^*\Gamma$ is a trace class of operator with a finite trace. Then, minimization problem (3) is expressed in terms of maximization problem as defined in Problem (4.11). Finally, eigenvalues and eigenvectors of $\Gamma^*\Gamma$ are introduced, revealing a close relationship with the correlation matrix C . To simplify notation, the snapshots $w(t_i)$ is written as w_i for $1 \leq i \leq M$ in the proof steps.

Operators Γ and Γ^*

To express the problem in functional analytic framework, the operators Γ and Γ^* are defined in this section. The usual Euclidean inner product on $\mathbb{R}^M$ is denoted by $\langle \cdot, \cdot \rangle_M$. Given the snapshots $\{w_i \in H : 1 \leq i \leq M\}$, the operator $\Gamma : H \rightarrow \mathbb{R}^M$ is defined by

$$(\Gamma\phi)_i = \langle \phi, w_i \rangle_H, \quad 1 \leq i \leq M. \quad (4.9)$$

If $\phi \in H$ and $g \in \mathbb{R}^M$ then

$$\langle \Gamma\phi, g \rangle_M = \sum_{i=1}^M \langle \phi, w_i \rangle_H g_i = \langle \phi, \Gamma^*g \rangle_H$$

where $\Gamma^* : \mathbb{R}^M \rightarrow H$ is given by

$$\Gamma^*g = \sum_{i=1}^M g_i w_i.$$

The following lemma introduces the trace-class operator $\Gamma^*\Gamma$ that maps H into itself and shows that its trace is finite.

Lemma 1. *Let $\{\phi^\alpha : \alpha \geq 1\}$ be any complete orthonormal basis for H . Then*

$$\sum_{\alpha \geq 1} \langle \phi^\alpha, \Gamma^*\Gamma\phi^\alpha \rangle_H = \sum_{i=1}^M \|w_i\|_H^2 < \infty. \quad (4.10)$$

Proof. We have

$$\sum_{\alpha \geq 1} \langle \phi^\alpha, \Gamma^* \Gamma \phi^\alpha \rangle_H = \sum_{\alpha \geq 1} \left(\sum_{i=1}^M \langle \phi^\alpha, w_i \rangle_H^2 \right) = \sum_{i=1}^M \left(\sum_{\alpha \geq 1} \langle \phi^\alpha, w_i \rangle_H^2 \right) = \sum_{i=1}^M \|w_i\|_H^2.$$

□

Changing the problem

To formulate the proof of Theorem (3), the problem is reformulated by introducing $\hat{\mu}$ as the supremum taken over all sets of vectors $\psi^1, \dots, \psi^n$ in H as follows:

$$\hat{\mu}_n = \sup \left\{ \sum_{\alpha=1}^n \langle \Gamma^* \Gamma \psi^\alpha, \psi^\alpha \rangle_H : \text{the set } \{\psi^1, \dots, \psi^n\} \text{ is orthonormal in } H \right\}. \quad (4.11)$$

The following result closely follows the main idea presented in Theorem 1 of [31].

Theorem 4. Suppose the supremum in the maximization problem (4.11) is attained by an orthonormal set $\{\hat{\psi}^1, \dots, \hat{\psi}^n\}$. Then the minimization problem (4.7) has $\mu_n = \sum_{i=1}^M \|w_i\|_H^2 - \hat{\mu}_n$ and the infimum is attained with $F_i = \sum_{\alpha=1}^n (\Gamma \hat{\psi}^\alpha)_i \hat{\psi}^\alpha$.

Proof. Write $F_i = \sum_{\alpha=1}^n a_i^\alpha \psi^\alpha$. Applying the Gram-Schmidt orthogonalization process yields a complete orthonormal system $\{\phi^\alpha : \alpha \geq 1\}$ such that $\psi^\alpha \in \text{span}\{\phi^1, \dots, \phi^n\}$ for all $\alpha = 1, \dots, n$. Thus, after some elementary rearranging, we may assume F is of the form $F_i = \sum_{\alpha=1}^n b_i^\alpha \phi^\alpha$ where the set $\{\phi^1, \dots, \phi^n\}$ gives the first n terms of a complete orthonormal system $\{\phi^\alpha : \alpha \geq 1\}$. Then

$$\begin{aligned} \sum_{i=1}^M \|w_i - F_i\|_H^2 &= \sum_{i=1}^M \left\| w_i - \sum_{\alpha=1}^n b_i^\alpha \phi^\alpha \right\|_H^2 \\ &= \sum_{i=1}^M \sum_{\beta \geq 1} \left\langle \phi^\beta, \left(w_i - \sum_{\alpha=1}^n b_i^\alpha \phi^\alpha \right) \right\rangle_H^2 \\ &= \sum_{\beta \geq 1} \sum_{i=1}^M \left((\Gamma \phi^\beta)_i - \sum_{\alpha=1}^n b_i^\alpha \langle \phi^\beta, \phi^\alpha \rangle_H \right)^2 \\ &= \sum_{\beta=1}^n \sum_{i=1}^M \left((\Gamma \phi^\beta)_i - b_i^\beta \right)^2 + \sum_{\beta > n} \sum_{i=1}^M \left((\Gamma \phi^\beta)_i \right)^2 \\ &= \sum_{\beta=1}^n \sum_{i=1}^M \left((\Gamma \phi^\beta)_i - b_i^\beta \right)^2 + \sum_{i=1}^M \|w_i\|_H^2 - \sum_{\beta=1}^n \langle \Gamma^* \Gamma \phi^\beta, \phi^\beta \rangle_H. \end{aligned}$$

where the last equality is obtained by applying (4.10). Minimizing the first term is achieved by selecting $b_i^\beta = (\Gamma \phi^\beta)_i$ for $1 \leq \beta \leq n$ and the minimization of the third term is accomplished by maximizing $\hat{\mu}_n$ as given in Problem (4.11). Thus $\mu_n = \sum_{i=1}^M \|w_i\|_H^2 - \hat{\mu}_n$ and the optimizing function is given by $\phi^\alpha(x) = \hat{\psi}^\alpha(x)$ and then $b^\alpha = (\Gamma \hat{\psi}^\alpha)$. □

Eigenvalues and eigenfunctions of $\Gamma^*\Gamma$

Since $\Gamma^*\Gamma$ is symmetric and non-negative and of rank at most M on a Hilbert space H , it can be diagonalized. Thus there is a complete orthonormal set of eigenfunctions $\{\nu^\alpha : \alpha \geq 1\}$ in H and corresponding eigenvalues $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3 \geq \dots \geq 0$ (counted with multiplicities) so that

$$(\Gamma^*\Gamma)\nu^\alpha = \hat{\lambda}^\alpha \nu^\alpha, \quad \alpha \geq 1.$$

(Given that H is a function space, the eigenvectors of $\Gamma^*\Gamma$ are referred to as eigenfunctions.)

Theorem 5. *For the maximization problem (4.11), we have $\hat{\mu}_n = \sum_{\alpha=1}^n \hat{\lambda}_\alpha$, attained by the orthonormal set $\{\nu^1, \dots, \nu^n\}$.*

Proof. Let $\{\psi^1, \dots, \psi^n\}$ be an orthonormal set in H , and define the $n \times n$ matrix A by $A_{\alpha\beta} = \langle \Gamma^*\Gamma\psi^\alpha, \psi^\beta \rangle$. Let $\kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_r > 0$ denote the strictly positive eigenvalues of A . The Rayleigh-Ritz procedure applied to $-\Gamma^*\Gamma$ and $-A$, see for example Theorem 10.12 in [118], implies $\hat{\lambda}_\alpha \geq \kappa_\alpha$ for $1 \leq \alpha \leq r$. Then

$$\sum_{\alpha=1}^n \langle \Gamma^*\Gamma\psi^\alpha, \psi^\alpha \rangle_H = \sum_{\alpha=1}^n A_{\alpha\alpha} = \text{trace}(A) = \sum_{\alpha=1}^n \kappa_\alpha \leq \sum_{\alpha=1}^n \hat{\lambda}_\alpha.$$

Therefore $\hat{\mu} \leq \sum_{\alpha=1}^n \hat{\lambda}_\alpha$, and equality is attained by choosing $\psi^\alpha = \nu^\alpha$. $\square$

Applying the result of Theorem (5) in Theorem (4) and noting

$$\sum_{\alpha \geq 1} \hat{\lambda}^\alpha = \sum_{\alpha \geq 1} \langle \nu^\alpha, (\Gamma^*\Gamma)\nu^\alpha \rangle_H = \sum_{i=1}^M \|w_i\|_H^2,$$

This leads to the following corollary.

Corollary 1. *Suppose $\Gamma^*\Gamma$ has eigenvalues $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq 0$ (counted with multiplicities) with orthonormal eigenfunctions $\nu^1, \nu^2, \dots$. Then for $n \geq 1$, the minimization problem (4.7) has solution $\mu_n = \sum_{\alpha > n} \hat{\lambda}_\alpha$ attained with $F_i = \sum_{\alpha=1}^n (\Gamma\nu^\alpha)_i \nu^\alpha$.*

The following section introduces the relationship between the correlation matrix C and $\Gamma^*\Gamma$, establishing a framework that is more suitable for computation.

Eigenvalues and eigenvectors of C

Corollary (1) provides the solution of the minimization problem (4.7) expressed in terms of the operator $\Gamma^*\Gamma$ on the Hilbert space H . Here it is reformulated in the computationally convenient form given in Theorem (3) in terms of the $M \times M$ matrix C . Notice that for $\gamma \in \mathbb{R}^M$ we have $(C\gamma)_i = \sum_{j=1}^M \langle w_i, w_j \rangle \gamma_j = \langle w_i, \Gamma^*\gamma \rangle = (\Gamma\Gamma^*\gamma)_i$, so that $C = \Gamma\Gamma^*$.

Proposition 1. *(i) If $\{\phi^1, \dots, \phi^n\}$ is any orthogonal set of non-zero eigenfunctions of $\Gamma^*\Gamma$ with non-zero eigenvalues $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$, then $\{\Gamma\phi^1, \dots, \Gamma\phi^n\}$ is an orthogonal set of non-zero eigenvectors of C with the same non-zero eigenvalues $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$.*

(ii) Conversely, if $\{\gamma^1, \dots, \gamma^n\}$ is any orthogonal set of non-zero eigenvectors of C with non-zero eigenvalues $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$, then $\{\Gamma^* \gamma^1, \dots, \Gamma^* \gamma^n\}$ is an orthogonal set of non-zero eigenfunctions of $\Gamma^* \Gamma$ with the same non-zero eigenvalues $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$.

Proof. (i) We have

$$C(\Gamma \phi^\alpha) = \Gamma \Gamma^* \Gamma \phi^\alpha = \Gamma(\Gamma^* \Gamma \phi^\alpha) = \Gamma(\tilde{\lambda}_\alpha \phi^\alpha) = \tilde{\lambda}_\alpha (\Gamma \phi^\alpha)$$

and

$$\langle \Gamma \phi^\alpha, \Gamma \phi^\beta \rangle_M = \langle \phi^\alpha, \Gamma^* \Gamma \phi^\beta \rangle_H = \tilde{\lambda}_\beta \langle \phi^\alpha, \phi^\beta \rangle_H.$$

Together, we see $\Gamma \phi^\alpha \neq 0$ and the $\Gamma \phi^1, \dots, \Gamma \phi^n$ are orthogonal and $\Gamma \phi^\alpha$ is an eigenvector of C with the same eigenvalue $\tilde{\lambda}_\alpha$ as the eigenfunction ϕ^α .

(ii) Similarly we have

$$\Gamma^* \Gamma(\Gamma^* \gamma^\alpha) = \Gamma^*(C \gamma^\alpha) = \Gamma^*(\tilde{\lambda}_\alpha \gamma^\alpha) = \tilde{\lambda}_\alpha (\Gamma^* \gamma^\alpha)$$

and

$$\langle \Gamma^* \gamma^\alpha, \Gamma^* \gamma^\beta \rangle_H = \langle \gamma^\alpha, \Gamma \Gamma^* \gamma^\beta \rangle_M = \langle \gamma^\alpha, C \gamma^\beta \rangle_M = \tilde{\lambda}_\beta \langle \gamma^\alpha, \gamma^\beta \rangle_M, \quad (4.12)$$

and the result follows as in (i). $\square$

Proof of Theorem (3). Let λ_α and γ^α be the eigenvalues and orthonormal eigenvectors of C , as in the statement of Theorem (3). Then in Corollary (1) we have $\hat{\lambda}_\alpha = \lambda_\alpha$ and orthogonal eigenfunction $\Gamma^* \gamma^\alpha$. Using (4.12) we have orthonormal eigenvectors

$$\psi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} (\Gamma^* \gamma^\alpha) = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^M \gamma_i^\alpha w_i$$

and then

$$a_i^\alpha = (\Gamma \psi_\alpha)_i = \frac{1}{\sqrt{\lambda_\alpha}} (\Gamma \Gamma^* \gamma^\alpha) = \sqrt{\lambda_\alpha} \gamma_i^\alpha.$$

$\square$

4.4 POD-Galerkin ROM Applied to Nonlinear ENSO Model

In this section, we apply POD techniques presented in Section (4.3) to the ENSO model described in Section (4.2). Suppose $\{u(t, x) : 0 \leq t \leq T, x \in \Omega\}$ is the solution of the coupled PDE (4.2), defined on some domain Ω and taking values in some Euclidean space $\mathbb{R}^5$. In (4.2), the equatorial atmosphere is modeled by a steady-state linear equation forced by heat from the equatorial Pacific Ocean surface, which is proportional to the SST $T(t, x)$. Hence, the first two components $K^A(t, x)$ and $R^A(t, x)$ of $u(t, x)$ are driven by $T(t, x)$. For each fixed time t the first two equations along with the boundary conditions on $K^A(t, x)$ and $R^A(t, x)$ the functions $x \mapsto K^A(t, x)$ and $x \mapsto R^A(t, x)$ are determined by the function $x \mapsto T(t, x)$. Thus there are linear transformations Φ_1 and Φ_2 sending functions of x to functions of x such that for each t we have

$$K^A(t, \cdot) = \Phi_1(T(t, \cdot)), \quad \text{and} \quad R^A(t, \cdot) = \Phi_2(T(t, \cdot)).$$

Therefore, every solution $u(t, x)$ of (4.2) satisfying the reflection boundary conditions that couple $K^O(t, x)$ and $R^O(t, x)$ is of the form

$$u(t, x) = [\Phi_1(T(\cdot, t))(x), \Phi_2(T(\cdot, t))(x), K^O(t, x), R^O(t, x), T(t, x)] \quad (4.13)$$

Assume the solution $u \in L^2([0, T], H)$ where H is some Hilbert space of functions $\Omega \rightarrow \mathbb{R}^5$ such as an L^2 or Sobolev space. Due to the functional dependance of $K^A(t, x)$ and $R^A(t, x)$ on $T(t, x)$ as shown in (4.13), for the ENSO system, we concentrate on the three functions

$$\bar{u}(t, x) = (K^O(t, x), R^O(t, x), T(t, x)) \quad (4.14)$$

and the space $\hat{H}$, which is again a Hilbert space of functions $\Omega \rightarrow \mathbb{R}^3$. For each fixed time t the mapping $x \mapsto (K^O(t, x), R^O(t, x), T(t, x))$ is in $\hat{H}$. Recall that the solution $u(t, x)$ satisfies $u(t, \cdot) \in H$ for all t , where H is a Hilbert space consisting of functions $[0, L] \rightarrow \mathbb{R}^5$. Define the new Hilbert space

$$\hat{H} = \{v : [0, L] \rightarrow \mathbb{R}^3 : (0, 0, v) \in H\}$$

with the inner product

$$\langle v, w \rangle_{\hat{H}} = \langle (0, 0, v), (0, 0, w) \rangle_H, \quad v, w \in \hat{H}.$$

Here if $v(x) = (K^O(t, \cdot), R^O(t, \cdot), T(t, \cdot))$ then $(0, 0, v) : [0, L] \rightarrow \mathbb{R}^5$ is the function

$$x \mapsto (0, 0, K^O(t, \cdot), R^O(t, \cdot), T(t, \cdot)).$$

Remark 4.1. : If the Hilbert space H is the product of Hilbert spaces $H_1 \times H_2 \times H_3 \times H_4 \times H_5$, so that $u \in H$ can be written

$$u(x) = (u_1(x), u_2(x), u_3(x), u_4(x), u_5(x))$$

with $u_i \in H_i$ for $i = 1, \dots, 5$ and

$$\langle u, v \rangle_H = \sum_{i=1}^5 \langle u_i, v_i \rangle_{H_i}, \quad u, v \in H$$

then $\hat{H} = H_3 \times H_4 \times H_5$. Therefore, for each fixed time t the mapping $x \mapsto (K^O(t, x), R^O(t, x), T(t, x))$ is in $\hat{H}$. It should be noted that we define $H = H_{KA} \times H_{RA} \times H_{KO} \times H_{RO} \times H_T$ and $\hat{H} = H_{KO} \times H_{RO} \times H_T$ in the PDE model (4.2), respectively. Considering this, we can define full snapshots

$$u_i(x) = (\Phi_1(T(t_i, \cdot))(x), \Phi_2(T(t_i, \cdot))(x), K^O(t_i, x), R^O(t_i, x), T(t_i, x)) \quad 1 \leq i \leq M$$

and reduced snapshots as

$$\bar{u}_i(x) = (K^O(t_i, x), R^O(t_i, x), T(t_i, x)) \quad 1 \leq i \leq M \quad (4.15)$$

where $\bar{u}_i(x) \in \hat{H}$. We can form an $M \times M$ matrix using the reduced snapshots such as

$$\mathcal{C}_{ij} = \frac{1}{M} \langle \bar{u}_i, \bar{u}_j \rangle_{\hat{H}} = \frac{1}{M} \left(\langle K_i^O(x), K_j^O(x) \rangle_{H_{KO}} + \langle R_i^O(x), R_j^O(x) \rangle_{H_{RO}} + \langle T_i(x), T_j(x) \rangle_{H_T} \right) \quad 1 \leq i, j \leq M \quad (4.16)$$

Using the reduced snapshots given in (4.15) and the eigenvectors γ^α and eigenvalues λ_α of matrix $\mathcal{C}$, we can define the orthonormal modes by (see Theorem (3))

$$\tilde{\Psi}_\alpha(x) = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^M \gamma_i^\alpha \bar{u}_i(x) \quad \text{and} \quad \hat{\Psi}_\alpha(x) = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^M \gamma_i^\alpha u_i(x) \quad (4.17)$$

where $\tilde{\Psi}(x)$ and $\hat{\Psi}(x)$ are constructed using reduced and full snapshots, respectively. It is worth pointing out that POD modes $\tilde{\Psi}(x)$ and $\hat{\Psi}(x)$ remain invariant regardless of initial conditions for the PDE model (4.2). This invariance of POD modes is presented in Appendix (C). Figure (4.1) illustrates the POD modes $\tilde{\Psi}_\alpha(x)$ and projected initial conditions $\langle (u_3(x, 0), u_4(x, 0), u_5(x, 0)), \tilde{\Psi}_\alpha \rangle$ where $\alpha = 1, \dots, 4$. As the full-order PDE model given in Equation (4.2) is discretized using 168 spatial nodes, each POD mode $\tilde{\Psi}_\alpha(x)$ consists of 504 elements, combining the corresponding modes for $K^O(t, x)$, $R^O(t, x)$ and $T(t, x)$, respectively. The reconstruction results are presented in the next section to show the accuracy of the method. After finding optimal POD bases $\{\tilde{\Psi}_1, \dots, \tilde{\Psi}_n\}$ and $\{\hat{\Psi}_1, \dots, \hat{\Psi}_n\}$, one can project PDE (4.2) onto these POD modes using Galerkin projection to find set of ODEs for the $\mathbb{R}^n$ -valued process $t \rightarrow (a_1(t), \dots, a_n(t))$ so that the solution given by $u_n(t, x)$ stays close to the solution $u(t, x)$. The steady state properties of first two components of model (4.2) suggest that the time evolution of the system is controlled by the evolution of the last three components $K^O(t, x)$, $R^O(t, x)$ and $T(t, x)$ satisfying

$$\begin{aligned} \frac{\partial}{\partial t} \bar{u}(t, x) &= \left[\frac{\partial K^O}{\partial t}(t, x) \quad \frac{\partial R^O}{\partial t}(t, x) \quad \frac{\partial T}{\partial t}(t, x) \right]^T \\ &= \mathcal{A}_0 u(t, x) + \mathcal{N}_0(u, u)(t, x) \end{aligned} \quad (4.18)$$

where $u(t, x)$ is given by (4.13) and $\mathcal{A}_0$ and $\mathcal{N}_0(u, u)$ denote the last three rows of $\mathcal{A}$ and $\mathcal{N}(u, u)$, respectively. Replacing $u(t, x)$ and $\bar{u}(t, x)$ with $u_n(t, x)$ and $\bar{u}_n(t, x)$ in (4.18) results in

$$\sum_{\alpha=1}^n \dot{a}_\alpha(t) \tilde{\Psi}_\alpha = \sum_{\alpha=1}^n a_\alpha(t) \mathcal{A}_0 \hat{\Psi}_\alpha + \sum_{\alpha, \beta=1}^n a_\alpha(t) a_\beta(t) \mathcal{N}_0(\hat{\Psi}_\alpha, \hat{\Psi}_\beta) \quad (4.19)$$

The PDEs for K^O and R^O are coupled through the boundary conditions (4.6), that is, $K^O(t, 0) = r_W R^O(t, 0)$ and $R^O(t, L_O) = r_E K^O(t, L_O)$. Hence, due to this coupling, the amplitudes $(a_1(t), a_2(t), \dots, a_n(t))$ in (4.19) are common for all the POD modes, as described in Remark 4.2 below. Taking the inner product of each side with the function $\tilde{\Psi}_\gamma$ such that

$$\sum_{\alpha=1}^n \dot{a}_\alpha(t) \langle \tilde{\Psi}_\alpha, \tilde{\Psi}_\gamma \rangle_{\hat{H}} = \sum_{\alpha=1}^n a_\alpha(t) \langle \mathcal{A}_0 \hat{\Psi}_\alpha, \tilde{\Psi}_\gamma \rangle_{\hat{H}} + \sum_{\alpha, \beta=1}^n a_\alpha(t) a_\beta(t) \langle \mathcal{N}_0(\hat{\Psi}_\alpha, \hat{\Psi}_\beta), \tilde{\Psi}_\gamma \rangle_{\hat{H}} \quad (4.20)$$

These assumptions suggest the following reduced order equations:

$$\dot{a}_\gamma(t) = \sum_{\alpha=1}^n a_\alpha(t) \langle \mathcal{A}_0 \hat{\Psi}_\alpha, \tilde{\Psi}_\gamma \rangle_{\hat{H}} + \sum_{\alpha, \beta=1}^n a_\alpha(t) a_\beta(t) \langle \mathcal{N}_0(\hat{\Psi}_\alpha, \hat{\Psi}_\beta), \tilde{\Psi}_\gamma \rangle_{\hat{H}} \quad \gamma = 1, \dots, n \quad (4.21)$$

Since

$$\langle \tilde{\Psi}_\alpha, \tilde{\Psi}_\gamma \rangle_{\hat{H}} = \begin{cases} 1 & \alpha = \gamma \\ 0 & \alpha \neq \gamma \end{cases} \quad (4.22)$$

Equation (4.21) represents the reduced order model (ODEs) of the full-order PDE model (4.2). Similar to POD analysis, we observe that four modes ($n = 4$) is sufficient to capture the main properties using the POD-Galerkin method. To simulate the ODEs, the initial conditions of the PDE are projected to determine the corresponding initial conditions given in (4.4) for the reduced-order model using $\langle (u_3(x, 0), u_4(x, 0), u_5(x, 0)), \tilde{\Psi}_\alpha \rangle$.

Remark 4.2 : Using this approach, we emphasize that each component of ENSO model (4.2) has distinct spatial modes written in vector form as $\{\tilde{\Psi}_\alpha(x)\}_{\alpha=1}^n$ and $\{\tilde{\Psi}_\alpha(x)\}_{\alpha=1}^n$, reflecting their unique variability patterns. However, the temporal coefficients $\{a_\alpha(t)\}_{\alpha=1}^n$ remain the same across all components, ensuring a unified low-dimensional representation of the system's dynamics as given in Equation (4.21). That approach makes it uniquely different from other applications of the POD-Galerkin method.

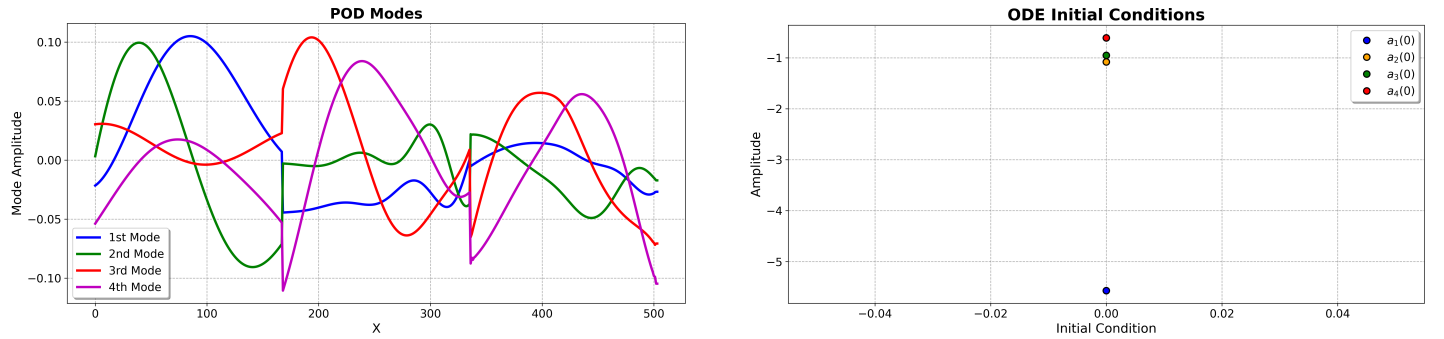


Figure 4.1: (a) POD modes $\tilde{\Psi}(x)$ and (b) projected initial conditions used to simulate the reduced-order equations.

Figure (4.2) illustrates initial conditions of the PDE model (4.2) $(u_3(x, 0), u_4(x, 0), u_5(x, 0))$ and reconstructed initial conditions using $\sum_{\alpha=1}^4 \langle (u_3(x, 0), u_4(x, 0), u_5(x, 0)), \tilde{\Psi}_\alpha \rangle \tilde{\Psi}_\alpha$. This result verifies that the ODEs (4.21) are simulated using correct initial conditions $\langle (u_3(x, 0), u_4(x, 0), u_5(x, 0)), \tilde{\Psi}_\alpha \rangle$.

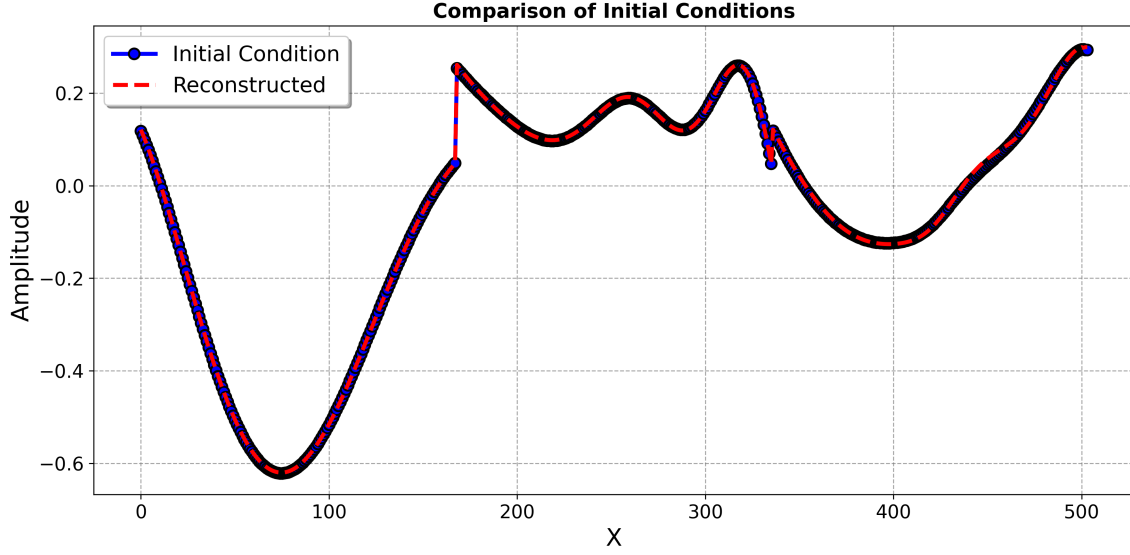


Figure 4.2: Comparison of initial conditions of PDE $\langle (u_3(x, 0), u_4(x, 0), u_5(x, 0)) \rangle$ and reconstructed initial conditions using four POD modes $\langle (\bar{u}_3(x, 0), \bar{u}_4(x, 0), \bar{u}_5(x, 0)) \rangle$.

Figure (4.3) depicts the comparison of POD time coefficients $q_\alpha(t)$ and POD-Galerkin ODEs $a_\alpha(t)$. $q_\alpha(t)$ is obtained using the formula $q_i^\alpha = \sqrt{\lambda_\alpha} \gamma_i^\alpha$ where λ_α and γ_α are the eigenvalues and eigenvectors of correlation matrix (4.4), respectively. Moreover, α and i represent the POD modes and elements of modes, respectively. On the other hand, $a_\alpha(t)$ is obtained using reduced order equations (4.21) with the initial conditions $\langle (u_3(x, 0), u_4(x, 0), u_5(x, 0)), \tilde{\Psi}_\alpha \rangle$. These results confirm the high stability and accuracy of the POD-Galerkin approach. We also observe that four modes are sufficient to capture the main features of the full-order model $u(t, x)$ while the first two modes are the most dominant, effectively capturing the primary dynamics of the full-order model as shown in Table (4.1). A detailed discussion of the results is provided in Section (4.4.1).

	$n = 1$	$n = 2$	$n = 3$	$n = 4$
KA	0.54387	0.04312	0.04293	0.03683
RA	0.31148	0.03050	0.02771	0.02116
KO	0.18504	0.02298	0.02268	0.02224
RO	0.04198	0.01884	0.01277	0.00971
T	0.06859	0.01562	0.01186	0.00838

Table 4.1: Reconstruction errors of model PDEs (root mean square) K^A, R^A, K^O, R^O, T , using different number of modes with $\alpha = 1, \dots, 4$

Comparison of Truth $q(t)$ and ROM $a(t)$ (Deterministic)

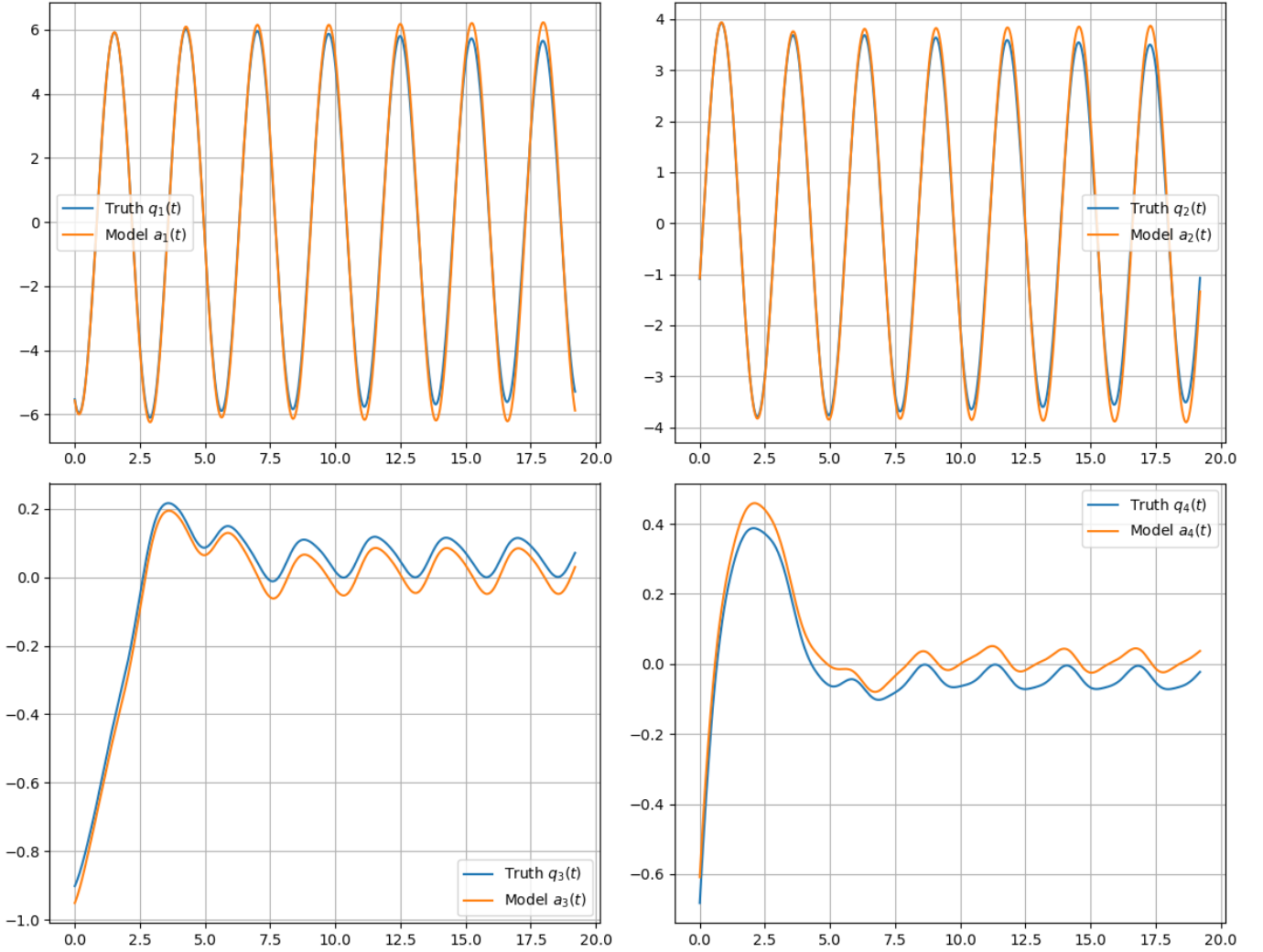


Figure 4.3: Comparison of POD time coefficients $q_\alpha(t)$ and simulations of reduced order equations $a_\alpha(t)$ for $\alpha = 1, \dots, 4$.

4.4.1 Reconstruction (Accuracy) Results Using POD-Galerkin Reduced Order Modeling

The reconstruction of model components given in Equation (4.14) can be performed in two ways: (i) using POD, where the approximation with four modes $\tilde{\Psi}_\alpha(x)$ and time coefficients $q_\alpha(t)$ is given by

$$\bar{u}_{\text{POD}}(t, x) \approx \sum_{\alpha=1}^4 q_\alpha(t) \tilde{\Psi}_\alpha(x) \quad (4.23)$$

and (ii) using POD-Galerkin with POD modes $\tilde{\Psi}_\alpha(x)$ and ODEs $a_\alpha(t)$, where the reconstruction takes the form

$$\bar{u}_{\text{POD-G}}(t, x) \approx \sum_{\alpha=1}^4 a_\alpha(t) \tilde{\Psi}_\alpha(x) \quad (4.24)$$

The first method involves POD modes $\tilde{\Psi}(x)$ and time coefficients $q_\alpha(t)$ *without* any reduced-order equations. On the other hand, the second method incorporates the simulation of ODEs, $a_\alpha(t)$ where the coefficients for the equations are obtained with the Galerkin projection using POD modes. Both $q_\alpha(t)$ and $a_\alpha(t)$ are shown in Figure (4.3). In both cases, while POD modes $\tilde{\Psi}_\alpha(x)$ are same, the time coefficients $a_\alpha(t)$ and $q_\alpha(t)$ are slightly different. However, both approximations are highly accurate using four time coefficients and four POD modes. Figures (4.4)-(4.8) illustrate the reconstruction results of PDE model (4.2). In each figure, first columns depict the simulation of the full-order model (4.2) using well-known numerical methods. In our problem, we employed numerical integration for the $K^A(t, x)$ and $R^A(t, x)$ due to their steady-state nature, and finite difference for $K^O(t, x), R^O(t, x)$ and $T(t, x)$. Moreover, while seconds columns represent the POD reconstruction using formula (4.23), third columns show the POD-Galerkin reconstruction using the formula (4.24). Finally, fourth and fifth columns are the $L1$ errors between FOM and POD and POD-Galerkin approximations, respectively. It should be noted that while FOM and reconstructed models are on the scale 10^{-1} , errors are on the scale 10^{-2} . This verifies that the accuracy of each method is always above 95% with the POD approach being slightly better.

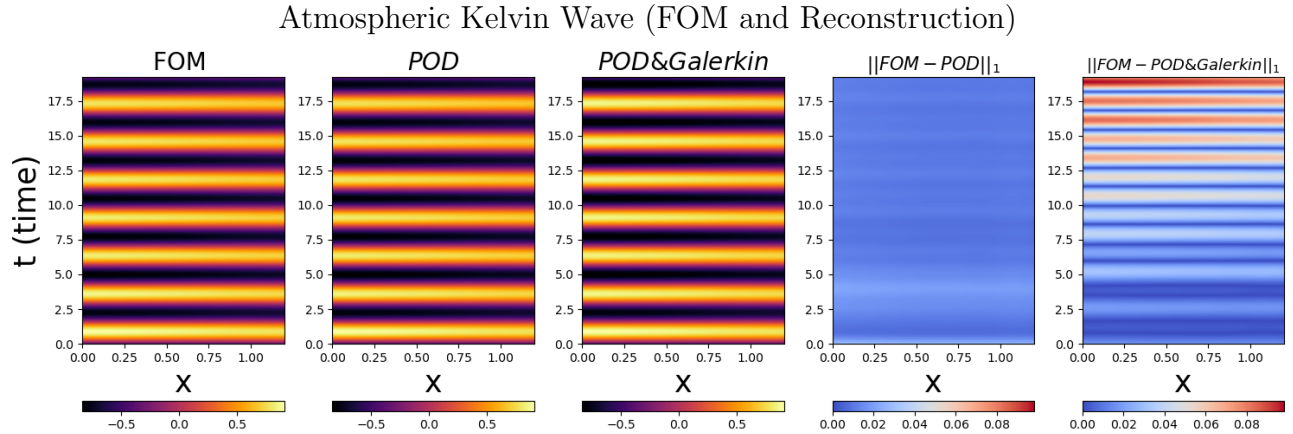


Figure 4.4: Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of atmospheric Kelvin $K^A(t, x)$ together with reconstruction (L1) errors.

Atmospheric Rossby Wave (FOM and Reconstruction)

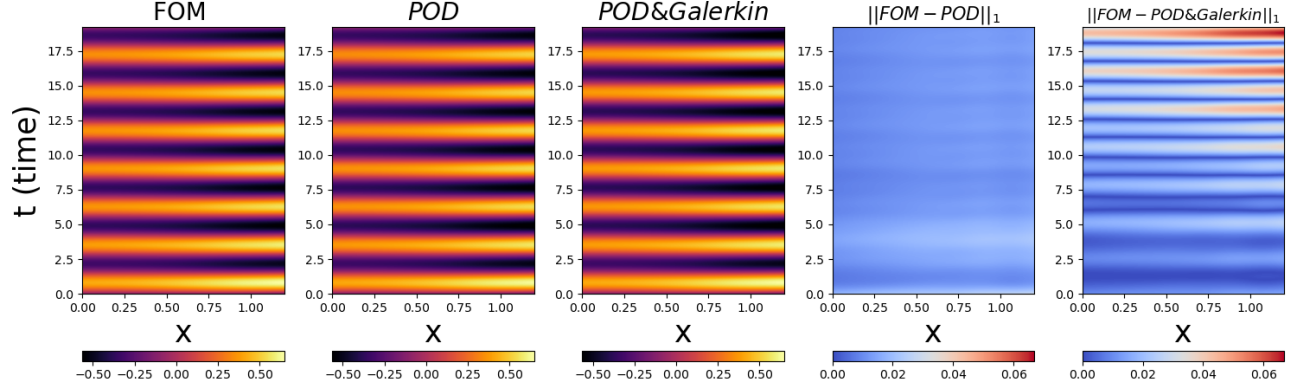


Figure 4.5: Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of atmospheric Rossby $R^A(t, x)$ together with reconstruction (L1) errors.

Oceanic Kelvin Wave (FOM and Reconstruction)

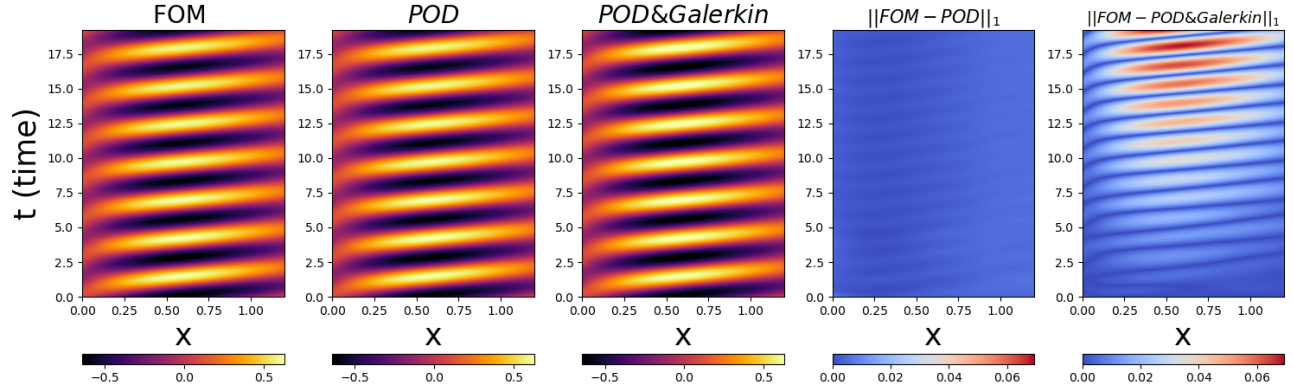


Figure 4.6: Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of oceanic Kelvin $K^O(t, x)$ together with reconstruction (L1) errors.

Oceanic Rossby Wave (FOM and Reconstruction)

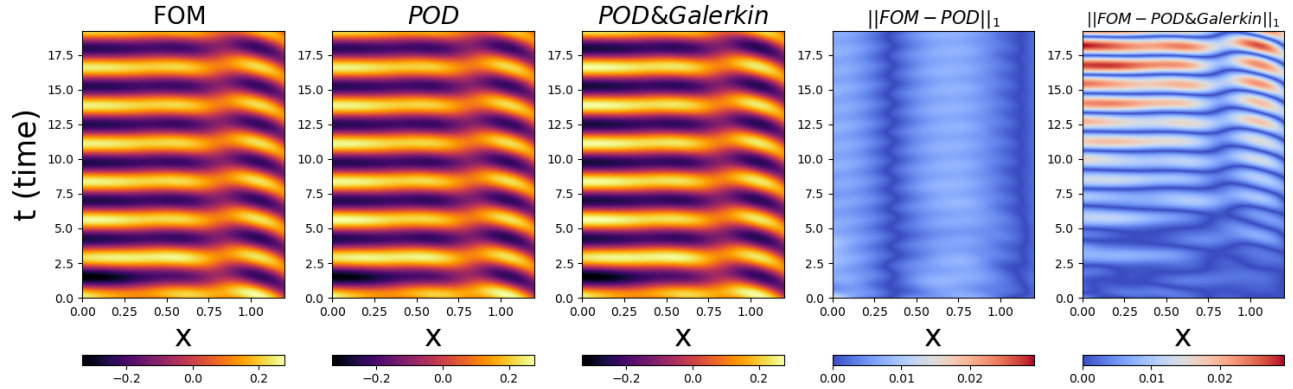


Figure 4.7: Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of oceanic Rossby $R^O(t, x)$ together with reconstruction (L1) errors.

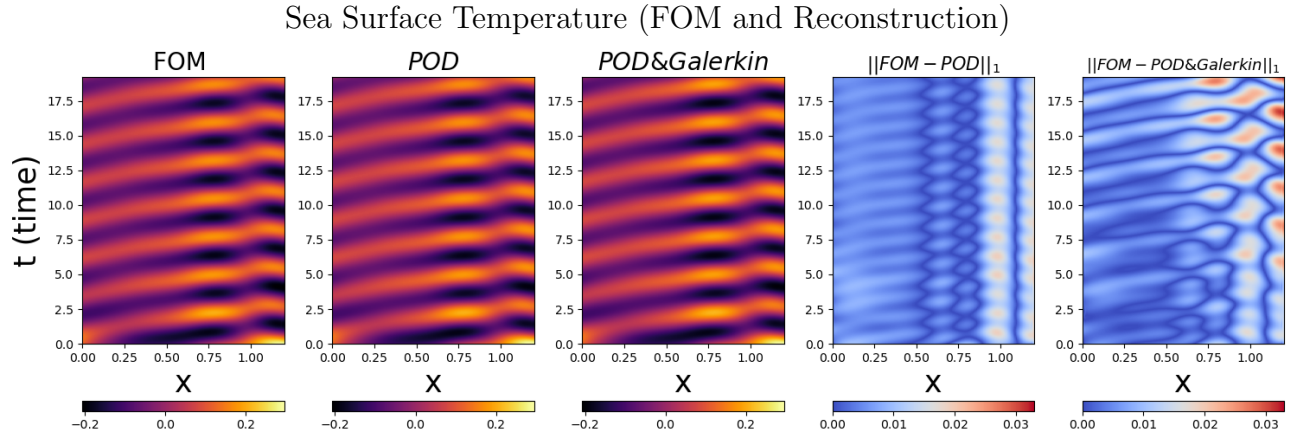


Figure 4.8: Full-order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of SST $T(t, x)$ together with reconstruction (L1) errors.

4.5 Neural Networks

In the previous section, we have employed POD and POD-Galerkin methods to reduced the dimension of high-dimensional model (4.2) and demonstrated accurate and stable results, which form the core of this chapter and thesis. On the other hand, in this section, we investigate the deep learning and neural networks for the dimensionality reduction. Deep learning, a subset of machine learning, has revolutionized artificial intelligence (AI). It enables models to automatically learn hierarchical feature representations from data. Inspired by the structure of biological neural networks, deep learning architectures typically consist of multiple layers of artificial neurons that process and transform input data through nonlinear activation functions. The resurgence of deep learning can be attributed to the availability of large-scale datasets, advances in computational power (e.g., GPUs), and algorithmic innovations such as backpropagation [140]. One of the pioneering breakthroughs was the development of convolutional neural networks (CNNs) by LeCun et al. [85], which significantly improved image recognition tasks. More recently, deep learning models such as ResNet [51] and Vision Transformers [33] have further enhanced performance across various domains, including computer vision and natural language processing (NLP).

Neural networks have been increasingly used for solving partial differential equations (PDEs), leveraging their ability to approximate complex functions and model high-dimensional systems. One prominent approach is Physics-Informed Neural Networks (PINNs) [133], which incorporate the governing PDEs into the loss function to enforce physical constraints during training. PINNs have been successfully applied to forward and inverse problems in fluid dynamics, heat transfer, and elasticity, demonstrating advantages over traditional numerical solvers, especially for high-dimensional or data-scarce scenarios. Extensions of PINNs, such as Variational PINNs (VPINNs) [73], utilize variational principles to improve accuracy and convergence.

Another significant neural network-based method for PDEs is the Deep Galerkin Method (DGM) [146], which employs deep neural networks to approximate solutions of high-dimensional PDEs, particularly useful for financial mathematics and stochastic differential equations. Additionally, Fourier

Neural Operators (FNOs) [89] provide an alternative approach by learning solution operators rather than explicit solutions, allowing for fast and generalizable predictions across different PDE instances. These methods have demonstrated the potential of deep learning in accelerating PDE simulations and expanding the scope of solvable problems beyond traditional computational methods. In this thesis, we employ autoencoders to reduce the dimensionality of the nonlinear ENSO model and compare with the POD-Galerkin approach.

4.5.1 Autoencoders

Autoencoders are a type of neural network designed to learn efficient representations of data, typically for the purpose of dimensionality reduction. They consist of two main parts: an encoder and a decoder as illustrated in Figure (4.9). The encoder compresses the input data into a lower-dimensional latent space, while the decoder reconstructs the original data from this compressed representation. The input-output relationship is given by

$$\hat{\mathbf{y}} = \mathbf{W}_2(\mathbf{W}_1\mathbf{y} + \mathbf{b}_1) + \mathbf{b}_2 \quad (4.25)$$

where $\mathbf{W}_1, \mathbf{W}_2^T \in \mathbb{R}^{m \times n}$ are the encoder and decoder weight matrices, and $\mathbf{b}_1 \in \mathbb{R}^m$, $\mathbf{b}_2 \in \mathbb{R}^n$ are the encoder and decoder bias vectors. Once the optimal $\{\mathbf{W}_1, \mathbf{W}_2, \mathbf{b}_1, \mathbf{b}_2\}$ are found, we can construct an encoder to reduce the input into a reduced order data $\mathbf{z} \in \mathbb{R}^m$ using $\{\mathbf{W}_1, \mathbf{b}_1\}$ and a decoder to convert the encoded data back to its original dimension using $\{\mathbf{W}_2, \mathbf{b}_2\}$. The goal of training an autoencoder is to minimize the reconstruction error between the input and the reconstructed output, ensuring that the latent space captures the most important features of the data. $\mathbf{y}$ and the reconstructed input $\hat{\mathbf{y}}$. A common choice for the reconstruction error is the mean squared error (MSE):

$$L(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \|\mathbf{y} - (\mathbf{W}_2(\mathbf{W}_1\mathbf{y} + \mathbf{b}_1) + \mathbf{b}_2)\|^2$$

The overall objective is to minimize the reconstruction error over the entire dataset. Let $\mathcal{D} = \{\mathbf{y}^{(i)}\}_{i=1}^N$ be the dataset with N examples. The minimization problem can be formulated as:

$$\min_{\{\mathbf{W}_1, \mathbf{W}_2, \mathbf{b}_1, \mathbf{b}_2\}} \frac{1}{N} \sum_{i=1}^N \|\mathbf{y}^{(i)} - (\mathbf{W}_2(\mathbf{W}_1\mathbf{y}^{(i)} + \mathbf{b}_1) + \mathbf{b}_2)\|^2$$

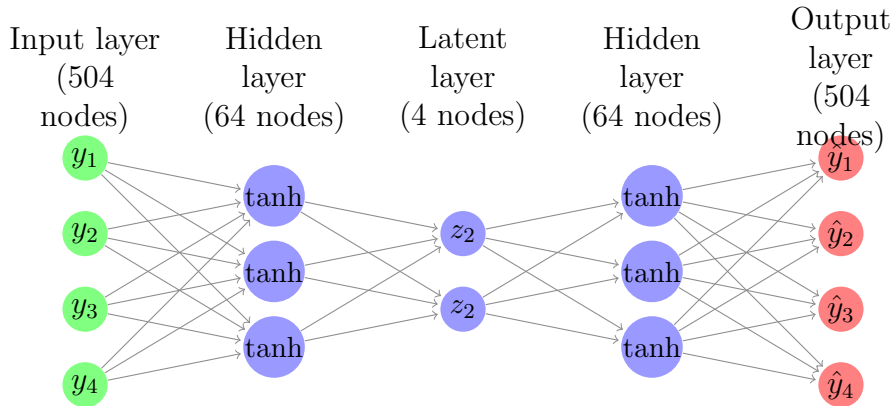


Figure 4.9: Autoencoder architecture for ENSO data with $\tanh()$ activations function

Autoencoders can capture intricate, non-linear patterns in data thanks to the use of non-linear activation functions in their neural network layers. This capability makes them more versatile and powerful than linear techniques like principal component analysis (PCA). Additionally, the architecture of autoencoders can be customized for specific tasks by adjusting the number of layers, neurons per layer, and types of activation functions, allowing for their application to a wide variety of problems. However, the latent space representations produced by autoencoders can be challenging to interpret, unlike the clear principal components in PCA. Furthermore, training autoencoders can be computationally demanding and time-consuming, especially for deep networks with numerous layers. They require careful hyperparameter tuning and a substantial amount of training data. In the next section, we discuss the system identification of reduced-order data obtained from POD and autoencoders using sparse identification of nonlinear dynamics (SINDy) method. With this approach, we gain more insights on the model discovery in addition to POD-Galerkin projection.

4.6 Sparse Identification of Nonlinear Dynamics (SINDy)

Sparse identification of nonlinear dynamics (SINDy) [16] is an algorithm for discovering the dynamical equations directly from the data. The problem of model discovery from data can be formulated as a feature selection problem in machine learning [74]. The SINDy algorithm takes m -time measurements of $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{X} = [\mathbf{x}(t_1), \dots, \mathbf{x}(t_m)]^T \in \mathbb{R}^{m \times n}$ and attempts to discover the structure of a nonlinear differential equation of the form

$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}(t)) \approx \mathbf{\Theta}(\mathbf{X})\mathbf{\Xi} \quad (4.26)$$

where $\mathbf{\Theta}(\mathbf{X}) = [\theta_1(\mathbf{X}), \theta_2(\mathbf{X}), \dots, \theta_p(\mathbf{X})] \in \mathbb{R}^{m \times p}$ form the dictionary of basis functions, and $\mathbf{\Xi} \in \mathbb{R}^{p \times n}$ is the matrix of coefficients, where each column corresponds to an equation with p terms. p is the maximal number of n -multivariate monomials of degree at most d . The majority of $\mathbf{\Xi}$ entries are zero while the remaining non-zero entries identify the active terms contributing to the sparse representation of the dynamics $\mathbf{f}(\mathbf{X})$.

To guarantee sparsity, SINDy is reformulated as a Least Absolute Shrinkage and Selection Operator (LASSO) problem

$$\min_{\mathbf{\Xi}} \frac{1}{m} \sum_{i=1}^m \|\dot{\mathbf{x}}(t_i) - \mathbf{\Theta}(\mathbf{x}(t_i))\mathbf{\Xi}\|_2^2 + \alpha \|\mathbf{\xi}\|_1. \quad (4.27)$$

where $\mathbf{\xi}(\mathbf{\Xi}) \in \mathbb{R}^{1 \times pn}$ is a vector of all entries inside $\mathbf{\Xi}$. LASSO is an optimization algorithm that finds a sparse solution for (4.27) by initializing $\mathbf{\Xi} = \mathbf{0}$ and at each iteration, it tries to find an update for $\mathbf{\Xi}$ one matrix entry at a time. The l_1 regularization coefficient α acts as a threshold such that if the an optimal condition involving α is not satisfied for a particular $\mathbf{\Xi}$ entry, the entry is chosen to be equal to zero. Increasing the value of α leads to more zero entries in $\mathbf{\Xi}$, resulting in a sparse model. The dictionary of basis functions for monomial sampling of a dynamical system is

$$\mathbf{\Theta}(\mathbf{X}) = \begin{bmatrix} | & | & | & | & | \\ \mathbf{1} & \mathbf{X} & \mathbf{X}^{\mathbf{P}_2} & \mathbf{X}^{\mathbf{P}_3} & \dots \\ | & | & | & | & | \end{bmatrix}. \quad (4.28)$$

The dictionary $\Theta(\mathbf{X})$ is constructed by appending candidate nonlinear functions of $\mathbf{X}$ column-wise. Here, higher order polynomials are denoted as $\mathbf{X}^{P_d}$ where d is the order of the polynomial considered. For example, element 1 is a column-vector of ones, element $\mathbf{X}$ is as defined above, element $\mathbf{X}^{P_2}$ is the matrix containing the set of all quadratic polynomial functions of the state vector $\mathbf{x}$, and is constructed as follows:

$$\mathbf{X}^{P_2} = \begin{bmatrix} x_1^2(t_1) & x_1(t_1)x_2(t_1) & \dots & x_2^2(t_1) & \dots & x_n^2(t_1) \\ x_1^2(t_2) & x_1(t_2)x_2(t_2) & \dots & x_2^2(t_2) & \dots & x_n^2(t_2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_1^2(t_m) & x_1(t_m)x_2(t_m) & \dots & x_2^2(t_m) & \dots & x_n^2(t_m) \end{bmatrix}. \quad (4.29)$$

In our research group, we applied neural networks and SINDy on the Moore-Greitzer compressor model where my contribution was on dimensionality reduction and SINDy components [123]. Moreover, we applied a compressed sensing, which forms the foundation of SINDy, to the Lorenz-96 atmospheric model where my contribution involved constructing a reduced-order ensemble Kalman filter [117]. These applications motivated me to apply similar methods to the ENSO model (4.2), which forms the core of this thesis.

4.6.1 Reconstruction (Accuracy) Results Using SINDy Reduced Order Modeling

In this section, the combined results using the proposed methods in this chapter are presented. As for the dimensionality reduction methods, we employ POD and autoencoders, while Galerkin projection and SINDy are used for model discovery. The reconstruction results of model components $K^O(t, x)$, $R^O(t, x)$, and $T(t, x)$ of the model (4.2) using four modes are presented in Table (4.2). Here, POD and AE represent the reduction methods while Galerkin and SINDy are the model discovery methods. The best reconstruction results are obtained using POD and AE, respectively. We observe more reconstruction errors when we use the model discovery results. Among them, POD-G provides the best reconstruction results that is followed by POD-SINDy and AE-SINDy, respectively. The corresponding reduced-order equations and their simulations are provided in Appendix (C).

	POD	POD-G	POD-SINDy	AE	AE-SINDy
KO	0.00350	0.02224	0.05892	0.01740	0.07206
RO	0.00496	0.00971	0.03354	0.01708	0.02697
T	0.00712	0.00838	0.02195	0.01352	0.02757

Table 4.2: Root mean square error between FOM and used methods for K^O , R^O , and T using four modes.

4.7 Conclusion and Future Work

This study has demonstrated the efficacy of the POD-Galerkin projection in modeling the El Niño-Southern Oscillation (ENSO) phenomena. By capturing the full-order model with only four modes

and four reduced order equations, we achieved a substantial reduction in computational complexity without significant loss of accuracy. Due to the special properties of the model, we introduced a novel approach using different POD bases, but the same time coefficients for all model components.

Our results illustrated the capability of the reduced-order model to replicate key oceanographic features associated with ENSO, including oceanic Kelvin waves, oceanic Rossby waves, and variations in sea surface temperature (SST). The reduced-order model was able to maintain the essential dynamics and spatiotemporal patterns of these features, which are critical for understanding and predicting ENSO events. Overall, the application of the POD-Galerkin method to ENSO modeling presents a promising approach for efficient and accurate simulations. This reduction in model complexity could facilitate more extensive and rapid assessments of climate scenarios, potentially enhancing predictive capabilities. Moreover, we applied neural networks and SINDy to further explore reduced-order modeling. Autoencoders were employed to reduce the model’s dimensionality, while SINDy was used to identify the underlying equations. We analyzed each method by comparing their reconstruction results.

Future work will concentrate on the reduced-order modeling of stochastic coupled ENSO models and assessing the efficiency of the proposed approach. This study lays the groundwork for developing reduced-order filtering techniques, a data assimilation technique that integrates observational data into reduced-order models to estimate and update the state of complex dynamical systems efficiently. By combining dimensionality reduction methods with state estimation techniques, like particle filters, reduced-order filtering retains essential dynamics while significantly reducing computational costs. This approach is vital for enhancing the accuracy and efficiency of climate models, enabling better predictions of ENSO.

Chapter 5

Data Assimilation of the El Niño-Southern Oscillation (ENSO)

5.1 Data Assimilation of the El Niño-Southern Oscillation (ENSO)

Data Assimilation (DA) is a mathematical technique used to combine observational data with numerical models to improve state estimation and forecasting of dynamical systems. It is widely used in meteorology, oceanography, climate science, and engineering applications. By integrating noisy, sparse, and incomplete observations, DA corrects model predictions in an optimal way using filtering algorithms [2, 84]. In numerical climate prediction (NCP), DA plays a crucial role by improving the initialization, accuracy, and reliability of climate models, where it combines real-time observations from satellites, radars, ocean buoys, and ground stations. Climate prediction involves long-term simulations of the Earth's atmosphere, oceans, and land systems, which require precise initial conditions to reduce forecast errors. DA alleviates biases and uncertainties in key climate variables such as sea surface temperature (SST), atmospheric circulation, and ocean currents [13, 18, 127].

DA is also important for enhancing sea surface temperature (SST) predictions for El Niño–Southern Oscillation (ENSO) by integrating observational data with numerical models [20–22]. ENSO-related SST anomalies in the equatorial Pacific strongly influence global climate, making accurate forecasts essential for early warning systems. Methods like 4D-Var and the Ensemble Kalman Filter (EnKF) are widely used by organizations like NOAA, ECMWF, and JMA to refine global weather forecasts and track extreme events like hurricanes [6, 131]. Building upon the developments in the field of data assimilation and the findings from the previous chapter, we present a reduced-order particle filter, which is a widely used data assimilation technique, for ENSO-related SST prediction as the final stage of this thesis.

This chapter is organized as follows. In Section (5.2), we present the nonlinear stochastic ENSO model. In Section (5.3), we introduce the POD-Galerkin ROM for this stochastic model and provide model validation. Section (5.4) covers nonlinear filtering theory and particle filters. In Section (5.5), we present the results of the particle filter applied to the reduced order ENSO model and prediction of ENSO-related SST. Finally, we discuss conclusions and future work in Section (5.6).

5.2 Stochastic Nonlinear ENSO Model

Stochastic nonlinear PDEs in six variables

$$u(t, x, \omega) = [K^A(t, x, \omega), R^A(t, x, \omega), K^O(t, x, \omega), R^O(t, x, \omega), T(t, x, \omega), \tau(t, \omega)]^T$$

can be written in differential equation form as

$$\mathcal{M} \frac{\partial u}{\partial t}(t, x, \omega) = \mathcal{A}(\kappa)u(t, x, \omega) + \mathcal{N}(u, u, \mu)(t, x, \omega) \quad (5.1)$$

where $\mathcal{M}$ is a 6×6 constant matrix with diagonals $[0, 0, 1, 1, 1, 1]$, the linear operator $\mathcal{A}(\kappa)$ in terms of bifurcation parameter κ , the wind stress coefficient, is

$$\mathcal{A}(\kappa) = \begin{bmatrix} -\frac{\partial}{\partial x} - \gamma & 0 & 0 & 0 & \chi_A \alpha_q (2 - 2\bar{Q})^{-1} & 0 \\ 0 & \frac{1}{3} \frac{\partial}{\partial x} - \gamma & 0 & 0 & \chi_A \alpha_q (3 - 3\bar{Q})^{-1} & 0 \\ \chi_O \frac{c\kappa}{2} & -\chi_O \frac{c\kappa}{2} & -c \frac{\partial}{\partial x} - \delta & 0 & 0 & \epsilon \chi_O \frac{c\kappa}{2} s_p(x) \\ -\chi_O \frac{c\kappa}{3} & \chi_O \frac{c\kappa}{3} & 0 & \frac{c}{3} \frac{\partial}{\partial x} - \delta & 0 & -\epsilon \chi_O \frac{c\kappa}{3} s_p(x) \\ 0 & 0 & c\eta(x) & c\eta(x) & -c\xi \alpha_q & 0 \\ 0 & 0 & 0 & 0 & 0 & d_p \end{bmatrix} \quad (5.2)$$

and $\mathcal{N}(u, u, \mu)$, nonlinearity with advection coefficient μ , is

$$\mathcal{N}(u, u, \mu) = [0 \quad 0 \quad 0 \quad 0 \quad -\mu \frac{\partial}{\partial x} [(K^O(t, x, \omega) - R^O(t, x, \omega))T(t, x, \omega)] \quad 0]^T$$

Randomness enter the model through the Ornstein-Uhlenbeck process $\tau(t)$ that has the form:

$$\frac{\partial \tau(t, \omega)}{\partial t} = -d_p(\tau(t, \omega) - \hat{\tau}) + \Xi_\tau \dot{W}(t, \omega) \quad (5.3)$$

where d_p , σ and $\hat{\tau}$ represent dissipation, noise strength and mean of the process, respectively. The boundary and initial conditions are same defined as in Chapter (4).

The coupled model $u(t, x, \omega)$ (5.1) is driven by the Ornstein-Uhlenbeck process $\tau(t)$, making it random. Here, $\epsilon = 0.1$ is included in the model to scale the noise intensity properly. Since we are working with stochastic PDEs, it is important to generate ensembles to validate the proposed method. For this purpose, we simulate ten ensembles of the Ornstein-Uhlenbeck process as shown in Figure (5.1). Each ensemble produces a different solution $u(t, x, \omega)$ depending on the choice of ω . In the next section, we show the projection of stochastic PDE (5.1) onto POD modes along with Galerkin projection.

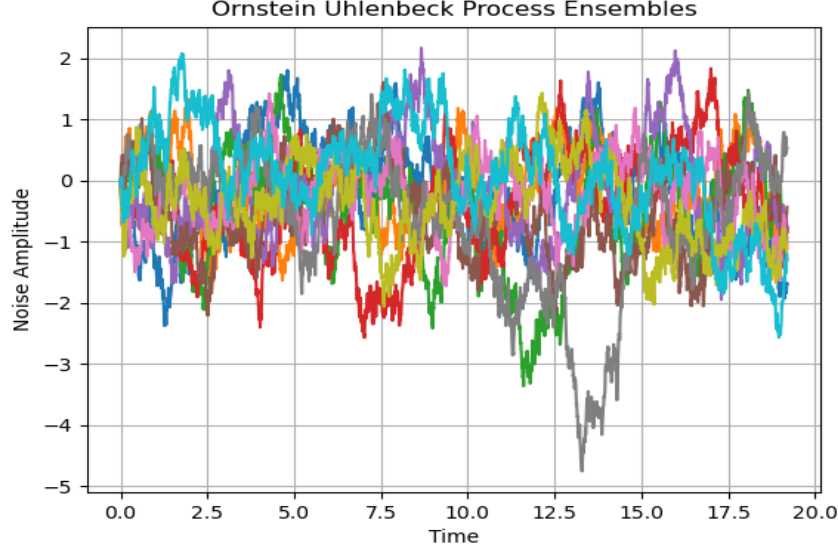


Figure 5.1: Ensembles of Ornstein-Uhlenbeck process $\tau(t, \omega)$ (ten ensembles).

5.3 Reduced Order Modeling of Stochastic Nonlinear ENSO Model

In Chapter (4), we have shown that the deterministic case of coupled model (5.1) (without the presence of $\tau(t)$) can be expanded in terms of POD modes $\Psi_\alpha^D(x)$ such that :

$$u_D(t, x) \approx \sum_{\alpha=1}^n q_\alpha^D(t) \Psi_\alpha^D(x) \quad (5.4)$$

where $u_D(t, x)$ is the solution of deterministic model, $q_\alpha^D(t)$ are the POD time coefficients and $\Psi_\alpha^D(x)$ are the POD modes obtained from deterministic problem. The solution can be projected onto modes $\Psi_\alpha^D(x)$ to obtain the time coefficients $q_\alpha^D(t)$ with the following inner product:

$$q_\alpha^D = \langle u_D(t, x), \Psi_\alpha^D(x) \rangle \quad (5.5)$$

On the other hand, we have shown that the PDE model can be projected onto those modes using Galerkin projection yielding the following reduced order equations:

$$\dot{a}_\gamma^D(t) = \sum_{\alpha=1}^n a_\alpha^D(t) \langle \mathcal{A}_0 \Psi_\alpha, \Psi_\gamma \rangle + \sum_{\alpha, \beta=1}^n a_\alpha^D(t) a_\beta^D(t) \langle \mathcal{N}_0(\Psi_\alpha, \Psi_\beta), \Psi_\gamma \rangle \quad (5.6)$$

where $\gamma = 1, \dots, n$. In this section, we apply the same procedure for the stochastic nonlinear model (5.1) to obtain POD time coefficients and *stochastic* reduced order equations. To distinguish deterministic solution $u_D(t, x)$, we shall use $u_S(t, x, \omega)$ in this section. As in the deterministic case,

the coupled model (5.1) can be expanded in terms of POD modes $\Psi_\alpha^D(x)$ such that

$$u_S(t, x, \omega) \approx \sum_{\alpha=1}^n q_\alpha^S(t, \omega) \Psi_\alpha^D(x) \quad (5.7)$$

where $u_S(t, x, \omega)$ is the solution of stochastic model, $q_\alpha^S(t, \omega)$ are the stochastic POD time coefficients. It is worth pointing out that we use the POD modes $\Psi_\alpha^D(x)$ obtained from the deterministic problem. The solution can be projected onto modes $\Psi_\alpha^D(x)$ to obtain the time coefficients $q_\alpha^S(t, \omega)$ with the following inner product:

$$q_\alpha^S(t, \omega) = \langle u_S(t, x, \omega), \Psi_\alpha^D(x) \rangle \quad (5.8)$$

On the other hand, the PDE model (5.1) can be projected onto those modes using Galerkin projection yielding the following *stochastic* reduced order equations:

$$\dot{a}_\gamma^S(t, \omega) = \sum_{\alpha=1}^n a_\alpha^S(t, \omega) \langle \mathcal{A}_0 \Psi_\alpha, \Psi_\gamma \rangle + \sum_{\alpha, \beta=1}^n a_\alpha^S(t, \omega) a_\beta^S(t, \omega) \langle \mathcal{N}_0(\Psi_\alpha, \Psi_\beta), \Psi_\gamma \rangle + \begin{bmatrix} \sigma_1 \\ \cdots \\ \sigma_n \end{bmatrix} \tau(t, \omega) \quad (5.9)$$

together with the Ornstein-Uhlenbeck process

$$\dot{\tau}(t, \omega) = -d_p(\tau(t, \omega) - \hat{\tau}) + \Xi_\tau \dot{W}(t, \omega) \quad (5.10)$$

The reduced order model equations (5.9) and (5.10) are coupled through $\tau(t)$. The noise projection coefficients $[\sigma_1 \dots \sigma_n]^\top$ are obtained using the inner product $\langle \epsilon(\chi_O \frac{c\kappa}{2} s_p(x), -\chi_O \frac{c\kappa}{3} s_p(x), 0), \Psi_\gamma \rangle$ as shown in Figure (5.2) where $\gamma = 1, \dots, 4$. The values of σ are shown in the Table (5.1).

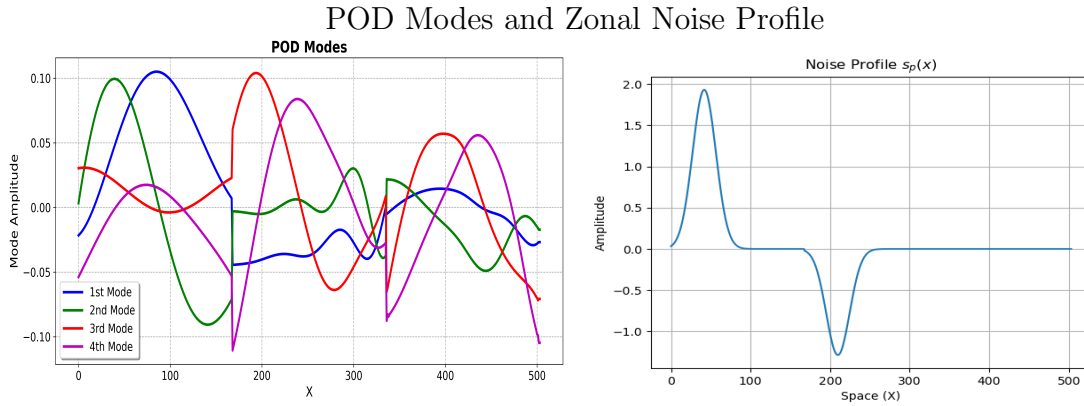


Figure 5.2: POD modes of deterministic model (left) and noise profile $\epsilon(\chi_O \frac{c\kappa}{2} s_p(x), -\chi_O \frac{c\kappa}{3} s_p(x), 0)$ as a function of space (right).

σ_1	σ_2	σ_3	σ_4
0.54463	0.62471	-0.24323	-0.16448

Table 5.1: Values of projected noise intensities σ .

In Figure (5.3), the scaled Ornstein-Uhlenbeck process with the corresponding σ values is shown where each one contributes as randomness to the differential equations (5.9) for each ensemble ω .

Ornstein–Uhlenbeck Process Ensembles Scaled by σ for Modes $N = 1, \dots, 4$

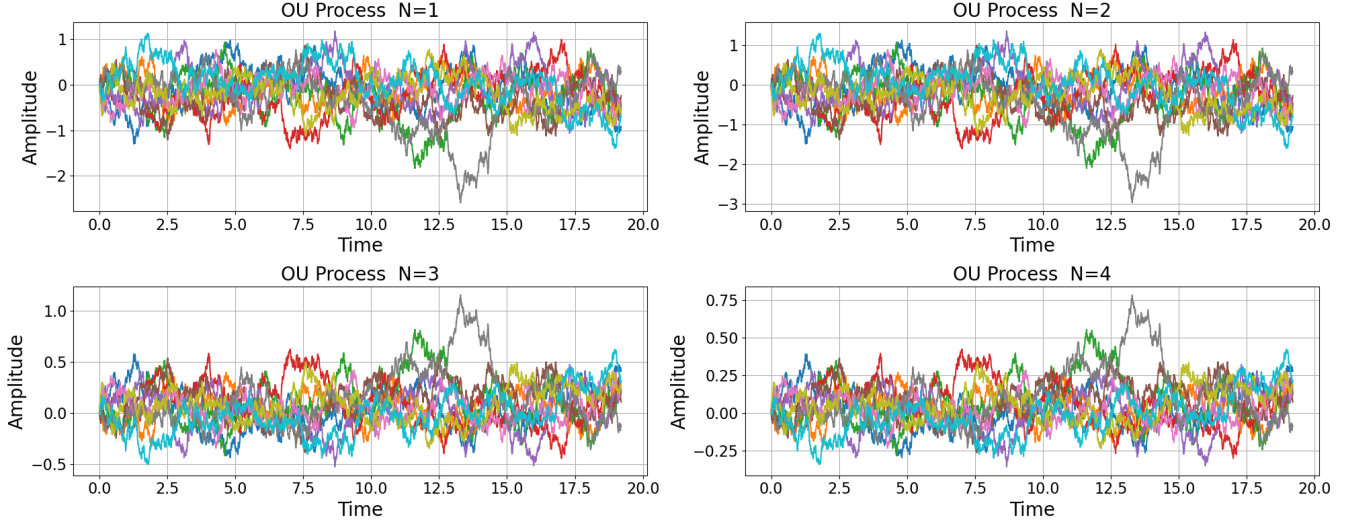


Figure 5.3: Scaled noise amplitude ensembles of the Ornstein-Uhlenbeck process with σ values.

Figure (5.4) depicts the time evolution of $q^s(t, \omega)$ and $a^s(t, \omega)$. While the first one is obtained using the projection of $u_S(t, x, \omega)$ onto POD modes $\Psi^D(x)$ using the formula (5.8), the latter is obtained by simulating SDEs (5.9) and (5.10). We use ten ensembles, but only the first four of them are shown for visualization purposes. In Figure (5.4), the comparison of modes for each ensemble is shown from left to right while ensembles are shown from top to bottom. The results validate that $q_\alpha^s(t, \omega)$ and $a_\alpha^s(t, \omega)$ align very well for $\alpha = 1, 2$ while we observe some errors for $\alpha = 3, 4$.

POD Coefficients and Stochastic Reduced-Order Dynamics

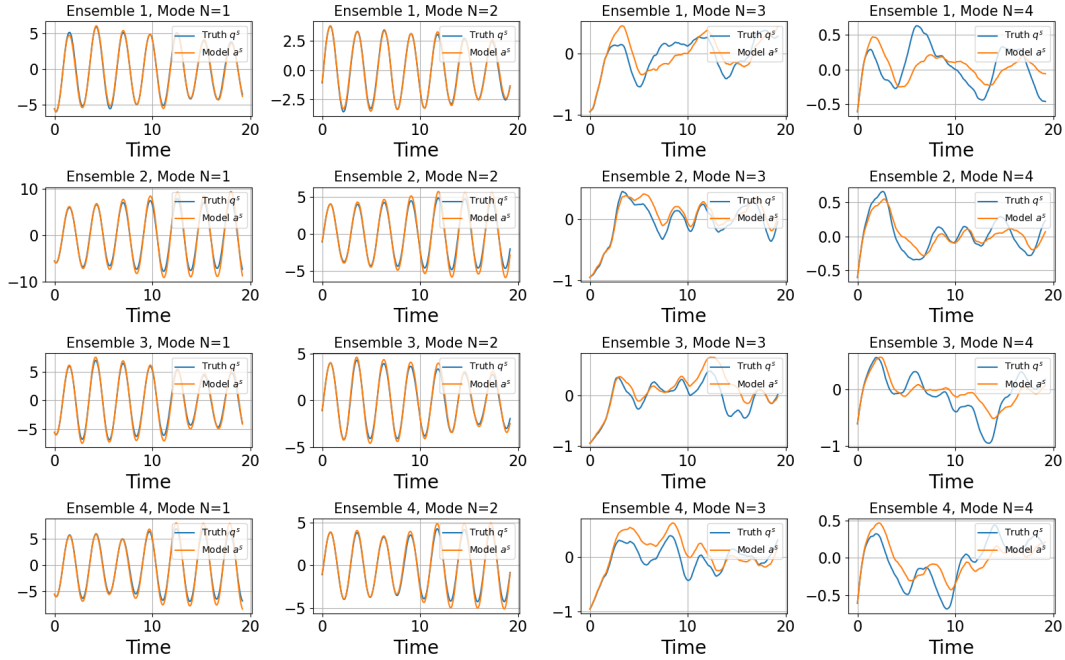


Figure 5.4: Comparison of POD time coefficients $q^s(t, \omega)$ and simulation of stochastic reduced order equations $a^s(t, \omega)$.

Figure (5.5) illustrates the full-order SPDE model as well as the reconstruction using POD and POD-Galerkin. The first figure is the average of ten ensembles of the simulation of SPDE model (5.1). The second one depicts the POD reconstruction using $\{q_\alpha^S(t)\}_{\alpha=1}^{\alpha=4}$ and $\{\Psi_\alpha^D(x)\}_{\alpha=1}^{\alpha=4}$. Final figure shows the reconstruction using $\{a_\alpha^S(t)\}_{\alpha=1}^{\alpha=4}$ and $\{\Psi_\alpha^D(x)\}_{\alpha=1}^{\alpha=4}$.

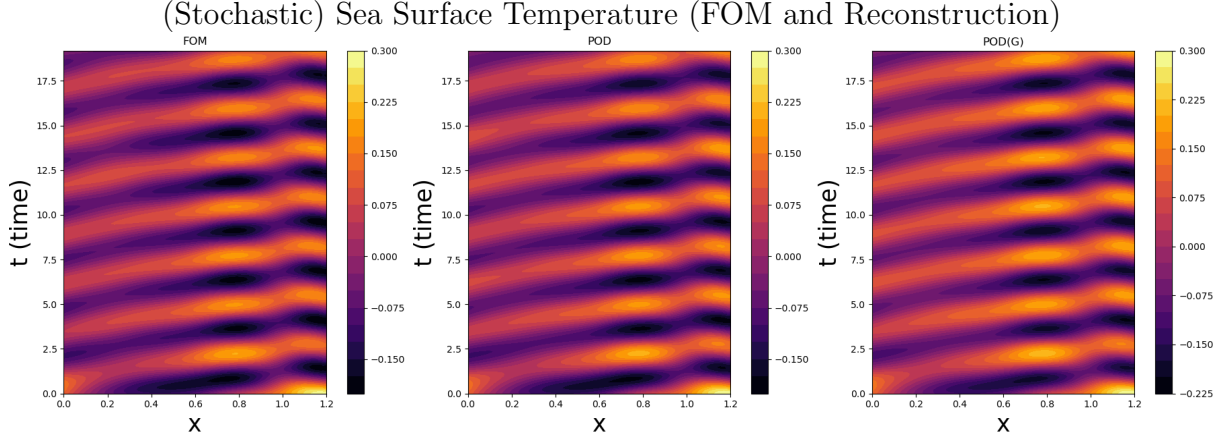


Figure 5.5: Stochastic full order model (FOM), POD and POD-Galerkin (POD-G) reconstructions of SST $T(t, x)$, each obtained using the average of ten ensemble simulations and four modes.

Figure (5.6) illustrates the reconstruction errors between the full-order models and POD as well as POD-Galerkin reconstructed models. Similar to the deterministic case, it should be noted that while FOM and reconstructed models are on the scale 10^{-1} , errors are on the scale 10^{-2} . This verifies that the accuracy of each method is always above 95% with the POD approach being slightly better. These results validate that the ROM (5.9) and (5.10) are appropriate candidates for reduced order filters, which are introduced in the next section. Moreover, it should be noted that $q_\alpha(t)$ and $a_\alpha(t)$ represent the stochastic time coefficient without use of superscript s in the next sections.

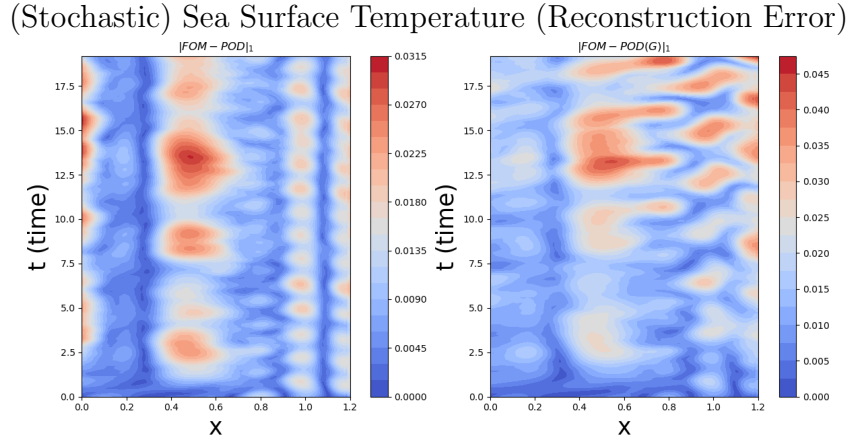


Figure 5.6: L1 errors between POD and POD-G reconstructions, each obtained using the average of ten ensemble simulations.

5.4 Nonlinear Filtering Theory

Nonlinear filtering, a subfield of data assimilation, has its origins in the mid-20th century, arising from efforts to extend the (linear) Kalman filter to nonlinear systems. Early theoretical foundations of nonlinear filtering were laid by Kushner (1964) and Stratonovich (1960), who independently derived the fundamental equation governing the evolution of the conditional probability density function, now known as the Kushner-Stratonovich equation [80, 150]. This equation describes the evolution of the posterior distribution of the state given observations, but is often computationally intractable for high-dimensional or complex systems. This limitation inspired the development of approximate methods, including the extended Kalman filter (EKF) and the unscented Kalman filter (UKF), which linearize or approximate the system dynamics and observation models to achieve practical implementations [145].

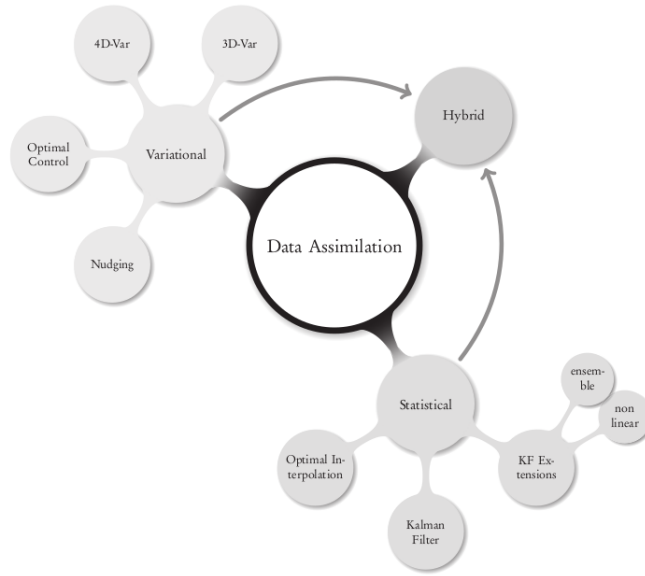


Figure 5.7: The big picture for DA methods and algorithms [2].

Over time, advancements in computational power and numerical techniques facilitated the development of more sophisticated methods for nonlinear filtering. The Fokker-Planck-based approaches, which use partial differential equations to describe the evolution of the probability density function, gained traction in certain applications but remained computationally expensive. We provide some results of Fokker-Planck-based methods on nonlinear filtering in the Appendix (D). Stochastic simulation methods, such as Monte Carlo approaches, emerged as powerful alternatives, enabling approximate solutions to the nonlinear filtering problem. The ensemble Kalman filter (EnKF), introduced by Evensen in 1994, became a popular choice in geophysical data assimilation, leveraging ensemble-based approximations of the covariance [37]. However, a major breakthrough occurred with the introduction of the particle filter (PF), which revolutionized nonlinear filtering by providing a flexible framework for approximating posterior distributions for high-dimensional systems [44, 86]. The particle filter, which relies on random sampling techniques to represent a probability distribution, is also known as the sequential Monte Carlo (SMC) method. These data assimilation methods and algorithms are illustrated in Figure (5.7). In the next sections, we provide the formulation of discrete time filtering and particle filters.

5.4.1 Nonlinear Filtering Problem

The nonlinear filtering problem involves estimating the state of a stochastic dynamical system using noisy and potentially incomplete observations. To define the problem, consider the evolution of the state space $\{\mathbf{x}_k, k \in \mathbb{N}\}$ given by

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \beta_k \quad (5.11)$$

in which $f(\cdot)$ denotes the deterministic model, and β_k is a random forcing representing unresolved physics, discretization, etc. The objective of filtering is to recursively estimate $\mathbf{x}_k$ from observations

$$\mathbf{y}_k = \mathbf{H}(\mathbf{x}_k) + \eta_k \quad (5.12)$$

in which the observation errors η_k are random vectors representing measurement errors and possibly representation error [79].

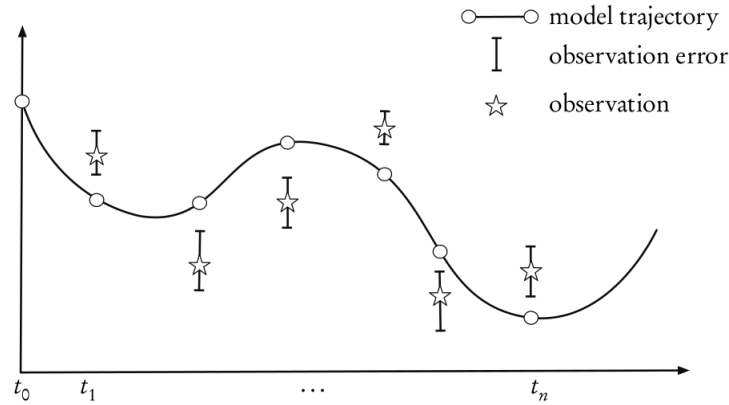


Figure 5.8: Sequential assimilation [2].

In a Bayesian formulation, the filtering problem is to recursively calculate the state estimate $\mathbf{x}_k$, at time k given the measurement $\mathbf{y}_{1:k}$ up to time k . Thus, it is essential to derive the probability density function (pdf) $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ where the initial pdf $p(\mathbf{x}_0|\mathbf{y}_0) \equiv p(\mathbf{x}_0)$ of the state vector, which is also known as the prior, is available. Then, in principle, the pdf $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ may be obtained, recursively, in two stages: prediction and update. The sequential scheme of model and observations are shown in Figure (5.8).

Suppose that the required pdf $p(\mathbf{x}_k|\mathbf{y}_{1:k-1})$ at time $k-1$ is available. The prediction stage involves using the system model (1) to obtain the prior pdf of the state at time k via the Chapman–Kolmogorov equation

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (5.13)$$

Note that in (5.13), use has been made of the fact that $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{1:k-1}) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$ as (5.11) describes a Markov process of order one. The probabilistic model of the state evolution $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is defined by the system equation (5.13) and the known statistics of $\mathbf{v}_{k-1}$.

At time step k , a measurement $\mathbf{y}_k$ becomes available, and this may be used to update the prior (update stage) via Bayes' rule

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} \quad (5.14)$$

where the normalizing constant

$$p(\mathbf{y}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) d\mathbf{x}_k \quad (5.15)$$

depends on the likelihood function $p(\mathbf{y}_k|\mathbf{x}_k)$ defined by the measurement model (5.12) and the known statistics of η_k . In the update stage (5.14), the measurement $\mathbf{y}_k$ is used to modify the prior density to obtain the required posterior density of the current state [1]. In the next section, we derive the Bayesian formulation to estimate the state $\mathbf{x}_k$ using particle filters.

5.4.2 Particle Filter

The particle filter was introduced in the 1990s by Gordon, Salmond Smith [45]. It uses a set of weighted particles to calculate the posterior distribution and updates these particles recursively as new observations become available. This method is particularly well-suited for systems with strong nonlinearity or non-Gaussian noise [1] as it avoids the linearization or Gaussian assumptions of Kalman-based filters. Applications of particle filters span a wide range of fields, including robotics (for simultaneous localization and mapping), finance (for volatility estimation), geophysics (for tracking atmospheric and oceanic states), and biomedical engineering (for neural signal processing).

In order to develop the details of the algorithm, let $\{\mathbf{x}_{0:k}^i, w_k^i\}_{i=1}^{N_s}$ denote a random measure that characterizes the posterior pdf $p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$, where $\{\mathbf{x}_{0:k}^i, i = 0, \dots, N_s\}$ is a set of support points with associated weights $\{w_k^i, i = 1, \dots, N_s\}$ and $\mathbf{x}_{0:k} = \{\mathbf{x}_j, j = 0, \dots, k\}$ is the set of all states up to time k . The weights are normalized such that $\sum_i w_k^i = 1$. Then, the posterior density at k can be approximated as

$$p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^i) \quad (5.16)$$

We therefore have a discrete weighted approximation of the true posterior, $p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$. The weights are chosen using the principle of importance sampling [11, 34]. This principle relies on the following: Suppose $p(x) \propto \pi(x)$ is a probability density from which it is difficult to draw samples but for which $\pi(x)$ can be evaluated (as well as $p(x)$ up to proportionality). In addition, let $x^i \sim q(x)$, $i = 1, \dots, N_s$ be samples that are easily generated from a proposal $q(\cdot)$ called an importance density. Then, a weighted approximation to the density $p(\cdot)$ is given by

$$p(x) \approx \sum_{i=1}^{N_s} w^i \delta(x - x^i) \quad (5.17)$$

where

$$w^i \propto \frac{\pi(x^i)}{q(x^i)} \quad (5.18)$$

is the normalized weight of the i th particle.

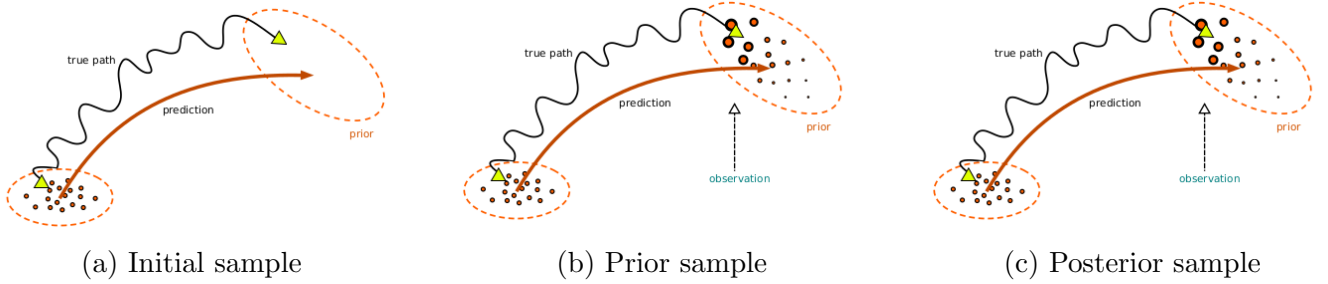


Figure 5.9: Evolution of signal density using particles [169].

Therefore, if the samples $\mathbf{x}_{0:k}^i$ were drawn from an importance density $q(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$, then the weights in (5.16) are defined by (5.18) to be

$$w_k^i \propto \frac{p(\mathbf{x}_{0:k}^i|\mathbf{y}_{1:k})}{q(\mathbf{x}_{0:k}^i|\mathbf{y}_{1:k})} \quad (5.19)$$

Returning to the sequential case, at each iteration, one could have samples constituting an approximation to $p(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1})$ and want to approximate $p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$ with a new set of samples. If the importance density is chosen to factorize such that

$$q(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) = q(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k})q(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1}) \quad (5.20)$$

then one can obtain samples $\mathbf{x}_{0:k}^i \sim q(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$ by augmenting each of the existing samples $\mathbf{x}_{0:k-1}^i \sim q(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1})$ with the new state $\mathbf{x}_k^i \sim q(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k})$. To derive the weight update equation, $p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$ is first expressed in terms of $p(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1})$, $p(\mathbf{y}_k|\mathbf{x}_k)$, and $p(\mathbf{x}_k|\mathbf{x}_{k-1})$. Note that (5.14) can be derived by integrating (5.22)

$$p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k-1})p(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} \quad (5.21)$$

$$= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} \times p(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1}) \quad (5.22)$$

$$= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} p(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1}) \quad (5.23)$$

By substituting (5.20) and (5.23) into (5.19), the weight update equation can then be shown to be

$$w_k^i = w_{k-1}^i \frac{p(\mathbf{y}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i|\mathbf{x}_{0:k-1}^i, \mathbf{y}_{1:k})} \quad (5.24)$$

Furthermore, if $q(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}) = q(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_k)$, then the importance density becomes only dependent on $\mathbf{x}_{k-1}$ and $\mathbf{y}_k$. This is particularly useful in the common case when only a filtered estimate of $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ is required at each time step. From this point on, we will assume such a case, except when explicitly stated otherwise. In such scenarios, $\mathbf{x}_{0:k-1}$ need not be stored; therefore, one can discard the path $\mathbf{x}_{0:k-1}^i$ and history of observations $\mathbf{y}_{1:k-1}$. The modified weight is then

$$w_k^i \propto w_{k-1}^i \frac{p(\mathbf{y}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i, \mathbf{y}_k)} \quad (5.25)$$

and the posterior filtered density $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ can be approximated as

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i) \quad (5.26)$$

where the weights are defined in (5.25). It can be shown that as $N_s \rightarrow \infty$, the approximation (5.26) approaches the true posterior density $p(\mathbf{x}_k|\mathbf{y}_{1:k})$.

This approach is called sequential importance sampling (SIS) algorithm, and consists of recursive propagation of the weights and support points as each measurement is received sequentially. Iteratively, the particle filtering algorithm is described by using the initial, prior and posterior samples as illustrated in Figure (5.9), and is summarized in Appendix (D).

5.5 Prediction of ENSO-Related SST Using a Particle Filter

In this section, we provide the prediction (estimation) results of ENSO-related SST anomalies using a particle filter. The model equations are given by

$$\begin{bmatrix} \dot{a}_1 \\ \dot{a}_2 \\ \dot{a}_3 \\ \dot{a}_4 \end{bmatrix} = \begin{bmatrix} f_1(a_1, a_2, a_3, a_4) \\ f_2(a_1, a_2, a_3, a_4) \\ f_3(a_1, a_2, a_3, a_4) \\ f_4(a_1, a_2, a_3, a_4) \end{bmatrix} + \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \end{bmatrix} \tau(t) \quad (5.27)$$

and coupled with the Ornstein-Uhlenbeck process $\tau(t)$:

$$\dot{\tau}(t) = -d_p(\tau(t) - \hat{\tau}) + \Xi_\tau \dot{W}(t) \quad (5.28)$$

These model equations are obtained by POD-Galerkin reduction as shown in Section (5.3). Moreover, the observation model is defined as:

$$dy(t) = \mathbf{H}ad t + dV(t)$$

where $\mathbf{a} = [a_1, a_2, a_3, a_4]$ and $V(t)$ is the observation noise. Observation matrix $\mathbf{H}$ is defined as:

$$\mathbf{H} = \begin{bmatrix} \psi_1^S(x_1) & \psi_2^S(x_1) & \psi_3^S(x_1) & \psi_4^S(x_1) \\ \psi_1^S(x_2) & \psi_2^S(x_2) & \psi_3^S(x_2) & \psi_4^S(x_2) \\ \psi_1^S(x_3) & \psi_2^S(x_3) & \psi_3^S(x_3) & \psi_4^S(x_3) \\ \psi_1^S(x_4) & \psi_2^S(x_4) & \psi_3^S(x_4) & \psi_4^S(x_4) \end{bmatrix}$$

where $\psi^S(x) = \{\psi_1^S(x), \psi_2^S(x), \psi_3^S(x), \psi_4^S(x)\}$ represent the SST modes as shown in Figure (5.10).

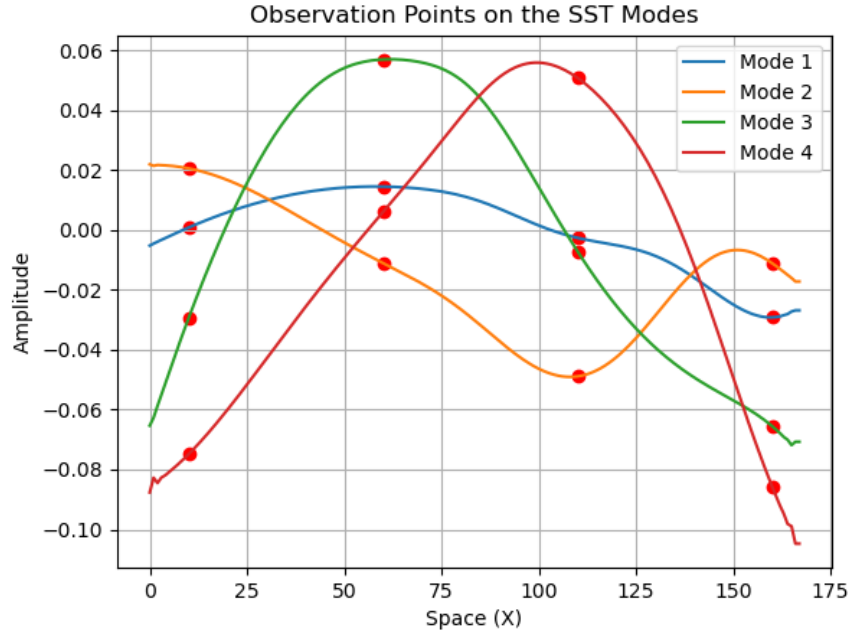


Figure 5.10: Four (artificial) sensors located on the SST modes $(\psi_1^S(x), \psi_2^S(x), \psi_3^S(x), \psi_4^S(x))$.

The artificial sensor locations on SST modes are chosen based the measurement locations along the equator in the Pacific Ocean, where Tropical Atmosphere Ocean (TAO) or PIRATA array sensors are often deployed to monitor sea surface temperatures (SSTs), ocean currents, and atmospheric conditions [108].

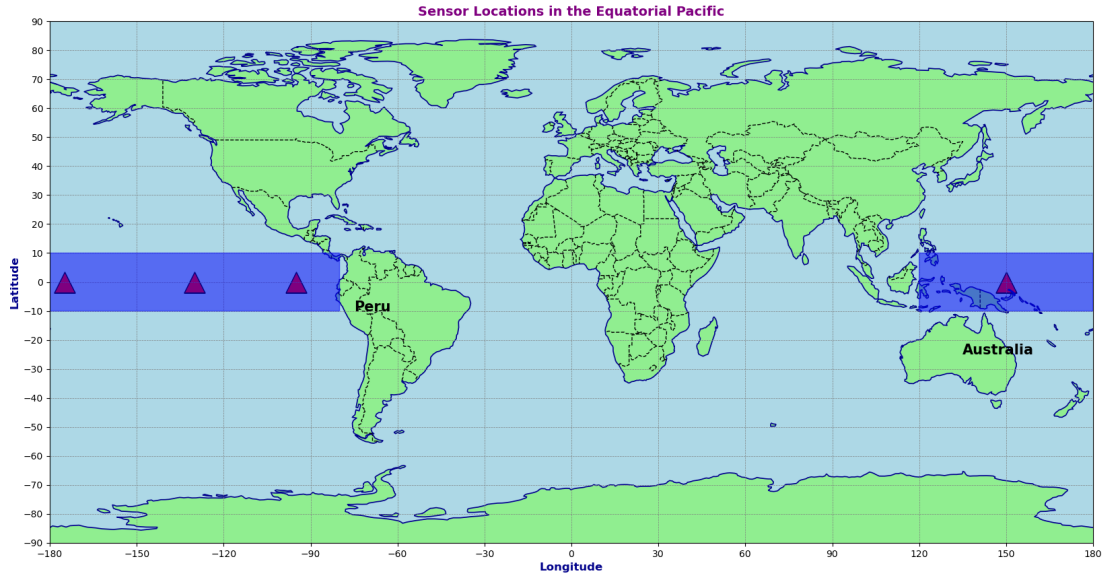


Figure 5.11: Four (artificial) sensors located between Australia (west) and South America (east) for ENSO-related SST prediction.

The locations on the SST modes are $x_1 = 10$, $x_2 = 60$, $x_3 = 110$ and $x_4 = 160$ are (roughly) corresponding to $165^\circ E$, $170^\circ W$, $125^\circ W$ and $95^\circ W$ in the physical map between Peru and Australia as shown in Figure (5.11). The ENSO regions are highlighted with dark blue, while sensor locations are highlighted with red. At $95^\circ W$, in the Eastern Pacific off the coast of Ecuador and near the Galápagos Islands, observations contribute to understanding ocean-atmosphere interactions. Moving westward, $125^\circ W$ lies within the Niño 3 region, crucial for ENSO monitoring. Further along, $170^\circ W$ marks the western part of the Niño 4 region, near the International Date Line, while $165^\circ E$ is situated in the far western Pacific within the warm pool region, close to Papua New Guinea [39]. These locations are key points in the equatorial Pacific used for climate monitoring.

Table (5.2), Table (5.3) and Table (5.4) show root mean square error (RMSE) between truth $q_\alpha(t)$ and $\hat{a}_\alpha(t)$ for each $\alpha = 1, \dots, 4$ and $N = 1, \dots, 10$ using 2, 10 and 100 particles, respectively. These results illustrate that the average RMSE between $q_\alpha(t)$ and $\hat{a}_\alpha(t)$ is highest when we use 2 particles as shown in Table (5.2). We observe a decrease in the RMSE as the number of particles increases to 10 as Table (5.3) depicts. Finally, 100 particles provides the least RMSE between truth and estimation as depicted in Table (5.4).

$N \setminus \alpha$	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$
1	2.15041	1.32658	0.19607	0.32399
2	1.27993	0.77188	0.38540	0.80816
3	1.49222	0.95776	0.34001	0.50791
4	1.31384	0.79834	0.35769	0.65422
5	3.95279	2.44826	0.59709	0.96357
6	1.43829	0.87249	0.52204	0.50693
7	1.81068	1.12634	0.31083	0.24162
8	0.70814	0.41439	0.36192	0.34583
9	5.71670	3.55058	0.38825	0.35637
10	1.32547	0.84686	0.61840	0.52932
N_{avg}	2.11885	1.31135	0.40777	0.52379

Table 5.2: Root mean square (RMSE) between truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 2 particles where α is the corresponding mode and N is the ensemble.

$N \setminus \alpha$	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$
1	2.50553	1.58457	0.27835	0.26389
2	1.37071	0.83383	0.40497	0.56726
3	1.69041	1.10872	0.32415	0.39766
4	1.32182	0.79636	0.29877	0.39169
5	2.22055	1.33722	0.21619	0.32237
6	1.88935	1.18696	0.25200	0.30577
7	4.18982	2.54341	0.30655	0.28061
8	0.85589	0.50584	0.36409	0.36558
9	2.47465	1.53720	0.17990	0.25208
10	2.01953	1.27629	0.36595	0.22006
N_{avg}	2.05383	1.27104	0.29909	0.33670

Table 5.3: Root mean square (RMSE) between truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 10 particles where α is the corresponding mode and N is the ensemble.

$N \setminus \alpha$	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$
1	0.91869	0.57305	0.18204	0.24005
2	1.14308	0.74703	0.17617	0.16347
3	0.99282	0.59680	0.24051	0.24119
4	0.68051	0.44013	0.25134	0.19244
5	0.45638	0.33377	0.15078	0.22276
6	0.60706	0.38026	0.24701	0.23554
7	0.43131	0.24686	0.13730	0.19837
8	1.31637	0.83361	0.36978	0.34164
9	0.42808	0.24496	0.18150	0.21906
10	1.26820	0.77991	0.15054	0.17587
N_{avg}	0.82478	0.52298	0.19752	0.22384

Table 5.4: Root mean square (RMSE) between truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 100 particles where α is the corresponding mode and N is the ensemble.

Figures (5.12), (5.13) and (5.14) illustrate the comparison of time series $q_\alpha(t)$ and $\hat{a}(t)$ using 2, 10 and 100 particles, respectively. Each time series is obtained using the average of 10 ensembles with the corresponding modes $\alpha = 1, \dots, 4$. We observe good agreement between $q_\alpha(t)$ and $\hat{a}(t)$ for $\alpha = 1, 2$ while we observe some errors depending on the number of particles for $\alpha = 3, 4$.

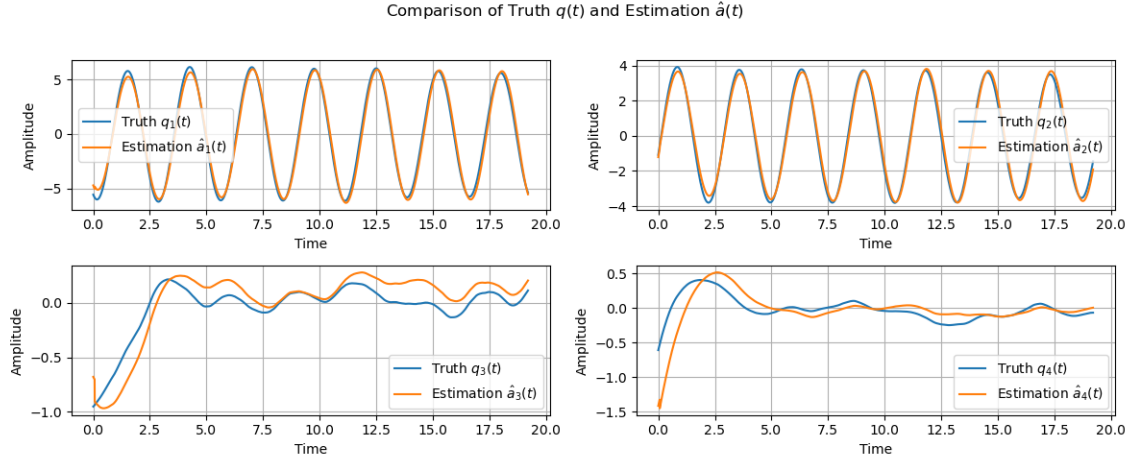


Figure 5.12: Average of truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 2 particles and 10 ensembles.

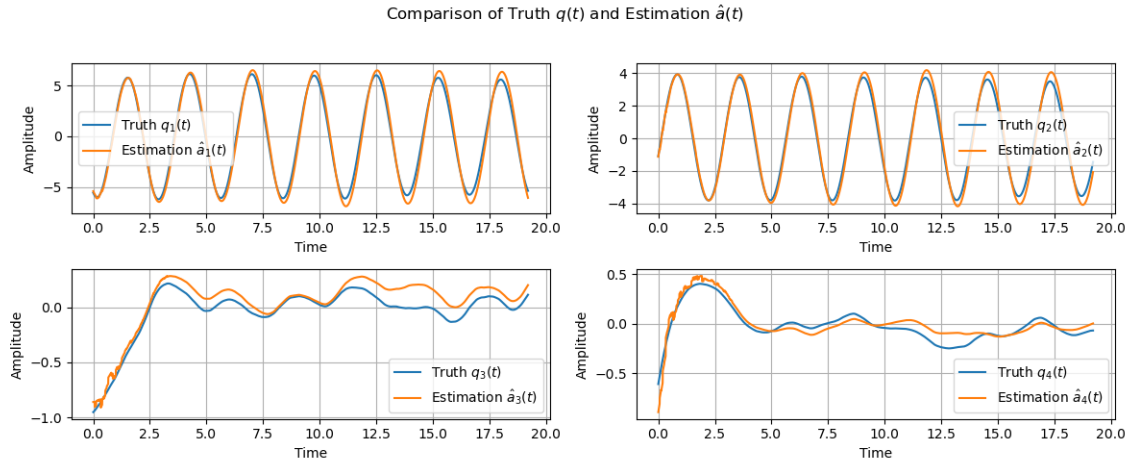


Figure 5.13: Average of truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 10 particles and 10 ensembles.

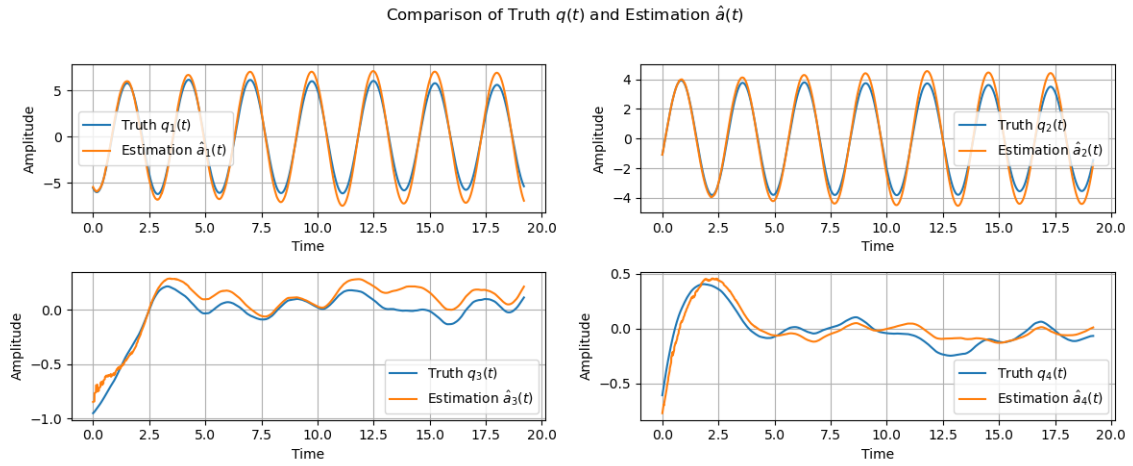


Figure 5.14: Average of truth $q_\alpha(t)$ and estimation $\hat{a}_\alpha(t)$ using 100 particles and 10 ensembles.

5.6 Conclusion and Future Directions

In this chapter, we explored the data assimilation of the stochastic ENSO model using particle filter. First, we have shown the validity of POD-Galerkin reduced order modeling, which was introduced in Chapter (4), for the stochastic partial differential equations driven by an Ornstein-Uhlenbeck process. We projected SPDEs onto POD modes coming from the deterministic model and introduced reduced order stochastic equations (SDEs). The validity and accuracy of this approach is shown by comparing the POD time coefficients $q_\alpha(t)$ and simulations of SDEs $a_\alpha(t)$. Our results show that the corresponding time coefficients align closely with high accuracy. To demonstrate the validity of these results, we generated 10 ensembles and showed that the method is valid for each ensemble.

The second part of this chapter focused on the prediction of ENSO-related SST anomalies using particle filters. In this part, a theoretical background of nonlinear filtering and particles filters has been introduced. After setting up the filtering problem, we combined the reduced order stochastic equations and artificial observations in the Pacific ocean, which is based on realistic experiments, to estimate the SST state. We tested the efficiency of the particle filter using a different number of particles and ensembles. Our results show that as we increase the number of particles, we obtain a more accurate estimation. Combining POD-Galerkin reduced order modeling of stochastic ENSO model and particle filtering, this chapter provides a promising approach for the accurate and efficient predictions of climate variables.

Future work will focus on the ENSO-related SST estimation using different data assimilation techniques such as ensemble Kalman filter (EnKF) to compare these results with the particle filter. Moreover, variational data assimilation methods such 4D-Var will be used to estimate more accurate initial conditions. Combining these methods with real ENSO measurements, we hope to be able to make more realistic SST predictions.

Chapter 6

Conclusion and Future Work

6.1 Conclusions and Future Work

6.1.1 Concluding Remarks

In this thesis, we explored different aspects of climate modeling and simulation with a focus on El Niño-Southern Oscillation (ENSO) and proposed efficient and accurate solutions by employing, developing and improving known mathematical techniques. The main challenges of climate modeling arise from uncertainty, nonlinearity and high-dimensionality. The main purpose of this work was to propose solutions to these challenges. We first proposed a new modeling and numerical framework for known ENSO models by adjusting the modeling domain and boundary conditions. This provided us with a modeling and simulation domain confined to the equatorial Pacific ocean for atmosphere, ocean and sea surface temperature, which are the main components of the ENSO mechanism. After producing appropriate simulation data, we attempted to improve well-known methods for efficient and accurate schemes. This work has been presented in Chapter (2).

One of the main challenges in climate modeling is the formulation and quantification of uncertainty. As a first step to overcome these difficulties, we applied Wiener Chaos Expansion (WCE) and Monte Carlo (MC) methods to the linear stochastic ENSO model in Chapter (3). The utilization of WCE provided a comprehensive study based on a spectral approach while MC simulations enhanced our ability to analyze stochastic models through ensemble methods. Although WCE and MC-based methods have been applied to many problems, this chapter investigated a relatively uncommon application to high-dimensional coupled SPDE models. We have shown that WCE is an accurate and cost-efficient spectral approach for linear high-dimensional problems compared to MC ensemble applications. However, we encountered some difficulties for the nonlinear stochastic ENSO model when using WCE due to the high-dimensionality and non-Gaussian nature of the problem. These challenges led us to develop dimensionality reduction techniques as the next stage of this thesis.

Building upon the fact that cost-efficient and accurate algorithms are essential for high-dimensional, nonlinear and stochastic models, we delved into reduced-order modeling (ROM) frameworks in Chapter (4). In this part of the thesis, we have shown the efficacy, stability and accuracy of the POD-Galerkin ROM in modeling ENSO phenomena. Our results demonstrated that by capturing

the full-order model with only four modes and four reduced-order equations, a substantial reduction in computational complexity is achieved without significant loss of accuracy. Due to the special properties of the model, we introduced a novel approach using different POD bases, but the same time coefficients for all model components. Moreover, we applied neural networks and SINDy to explore different ROM and model discovery techniques. After developing stable and accurate reduced order models, we presented a reduced-order particle filter in the final chapter of this thesis.

The final part of this thesis covered the data assimilation of the nonlinear stochastic ENSO model. We first demonstrated the validity of POD-Galerkin reduced order modeling, which was introduced in Chapter (4), for the stochastic partial differential equations driven by the Ornstein-Uhlenbeck process. We projected SPDEs onto POD modes derived from the deterministic model and introduced reduced-order stochastic equations (SDEs). After setting up the filtering framework, we combined the reduced-order stochastic equations and artificial observations in the Pacific ocean, which is based on realistic experiments, to estimate the ENSO-related SST anomalies. We employed particle filters and tested the efficiency using different numbers of particles and ensembles in Chapter (5).

To conclude, from novel ENSO modeling to uncertainty quantification, from reduced order modeling to nonlinear filtering, this thesis provided a promising approach for accurate and efficient predictions of ENSO-related climate variables. In the next section, we present suggestions and future directions for each stage of this thesis to improve the efficiency, reliability and accuracy of ENSO-related climate variables.

6.1.2 Future Work

In this thesis, we have presented the different aspects of ENSO from modeling to simulations (both deterministic and stochastic) and from reduced-order modeling to SST estimations. The scenarios described below are ones that can be studied in the future for more accurate and realistic results.

Advanced numerical methods and coupled ENSO-MJO models

In Chapter (2), we introduced a novel ENSO model by adjusting the model domain. To develop more realistic modeling approaches, a natural step would be to couple the modified ENSO model with another important climate phenomena such as the Madden-Julian Oscillation (MJO). The coupled ENSO and MJO models can be given as:

$$\mathcal{M}_1 \frac{\partial u_1}{\partial t}(t, x) = \mathcal{A}_1(\kappa)u_1(t, x) + \mathcal{N}_1(u_1, u_1, \mu)(t, x) + C_{21}u_2(t, x) + \dot{W}_1(t) \quad (6.1)$$

$$\mathcal{M}_2 \frac{\partial u_2}{\partial t}(t, x) = \mathcal{A}_2(\kappa)u_2(t, x) + \mathcal{N}_2(u_2, u_2, \mu)(t, x) + C_{12}u_1(t, x) + \dot{W}_2(t) \quad (6.2)$$

where model (6.1) and model (6.2) represent the ENSO and MJO models with coupling constants C_{12} and C_{21} , respectively. The importance of ENSO-MJO coupling and some proposed models are introduced in [52, 116, 122, 152]. Moreover, we presented simulation results of ENSO in (2), where we considered the model with a small nonlinearity. As we increased the nonlinearity, we encountered

numerical instabilities. To address this issue, future work should explore more advanced numerical methods, such as finite volume or finite element, to capture high nonlinearity more effectively.

Nonlinear Stochastic ENSO Model Using Sparse Wiener Chaos Expansion

We demonstrated the application of the stochastic linear ENSO model using WCE and compared it with the Monte Carlo method in Chapter (3). The WCE can be separated into two components:

$$u(t, x, \omega) = \sum_{|\alpha| \leq 1} u_\alpha(t, x) V_\alpha(\xi) + \sum_{|\alpha| \geq 2} u_\alpha(t, x) V_\alpha(\xi) \quad (6.3)$$

The first component provides a Gaussian approximation of $u(t, x, \omega)$, while the second component captures the higher order non-Gaussian terms. Since the problem in Chapter (3) is linear, we focused only on the first component. We have also emphasized the challenges of the WCE method when applied to a nonlinear model due to its non-Gaussian feature. Sparse truncation methods have been proposed in the literature to address these challenges [57]. The next step in this stage of the research would be to apply an appropriate sparse truncation approach to solve the nonlinear ENSO model using WCE.

POD-Galerkin Reduced Order Modeling Using H^1 (Sobolev) Norm

In Chapter (4), we presented POD-Galerkin reduced-order modeling applied to the nonlinear ENSO model. Our results were based on the L^2 norm when computing the correlation matrix and applying Galerkin projection. We obtained stable, efficient and accurate results using the L^2 norm. However, as the nonlinearity (parameter μ) increases, the problem's complexity grows. In such cases, stabilization methods may be required to achieve a stable solution. Various stabilization techniques in POD-Galerkin methods have proposed in the literature using H^1 (Sobolev) norm [38, 61]. By considering the derivatives of the solution, H^1 (Sobolev) norms ensure that the reduced basis functions approximate the solution accurately while also capturing its smoothness and gradient information. Using the H^1 norm, the correlation matrix is defined as:

$$\mathcal{C}_{ij} = \frac{1}{M} (\langle \bar{u}_i, \bar{u}_j \rangle + \epsilon \langle \nabla \bar{u}_i, \nabla \bar{u}_j \rangle) \quad 1 \leq i, j \leq M \quad (6.4)$$

where ϵ and M are the weight given on the gradient terms and number of snapshots, respectively. The eigenvectors and eigenvalues of the matrix (6.4) are used to construct the POD modes. Moreover, the reduced order equations have the following form using the H^1 norm:

$$\dot{a}_\gamma(t) = \sum_{\alpha=1}^n a_\alpha(t) \left(\langle \mathcal{A}_0 \hat{u}_\alpha, \tilde{u}_\gamma \rangle + \epsilon \langle \nabla (\mathcal{A}_0 \hat{u}_\alpha), \nabla \tilde{u}_\gamma \rangle \right) \quad (6.5)$$

$$+ \sum_{\alpha, \beta=1}^n a_\alpha(t) a_\beta(t) \left(\langle \mathcal{N}_0(\hat{u}_\alpha, \hat{u}_\beta), \tilde{u}_\gamma \rangle + \epsilon \langle \nabla (\mathcal{N}_0(\hat{u}_\alpha, \hat{u}_\beta)), \nabla (\tilde{u}_\gamma) \rangle \right) \quad (6.6)$$

Depending on the complexity of the problem, either L^2 or H^1 norm can be more suitable.

Data Assimilation Using Real-Time Observation and Variational Methods

In Chapter (5), we demonstrated that ENSO-related SST anomalies can be predicted using data assimilation algorithms. We employed particle filters and presented prediction results. Accurate initial conditions are essential for numerical climate prediction. Data assimilation algorithms, such as 4D-Var, help ensure accurate initial conditions, which are particularly important for chaotic systems like the atmosphere and ocean, where small errors in initial conditions grow rapidly [7, 30]. Given the observations

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\varepsilon}_r \quad (6.7)$$

and the background vector

$$\mathbf{x}_b = \mathbf{x} + \boldsymbol{\varepsilon}_b \quad (6.8)$$

the objective is to minimize the following 4D-Var cost function using the least squares method to obtain an estimation:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^\top \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \quad (6.9)$$

Here, $\mathbf{B}$ is the covariance matrix of the background $\boldsymbol{\varepsilon}_b$, and $\mathbf{R}$ the covariance matrix of the observation error $\boldsymbol{\varepsilon}_r$, which accounts for both instrument and representativeness errors [132]. $\mathbf{H}$ is the observation operator, mapping the variables from the model space to the observation space as defined in Chapter (5). By integrating reduced-order modeling with accurate initial conditions and real-time data, this approach may yield more precise and reliable climate predictions.

References

- [1] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, *A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking*, IEEE Transactions on Signal Processing **50** (2002), no. 2, 174–188.
- [2] M. Asch, M. Bocquet, and M. Nodet, *Data assimilation: methods, algorithms, and applications*, Society for Industrial and Applied Mathematics, 2016.
- [3] N. Aubry, P. Holmes, J. L. Lumley, and E. Stone, *The dynamics of coherent structures in the wall region of a turbulent boundary layer*, Journal of Fluid Mechanics **192** (1988), 115–173.
- [4] Y. Aydogdu, P. Baxendale, and N. S. Namachchivaya, *Stochastic perturbations of el nino southern oscillations (enso): a wiener chaos approach*, Egu general assembly conference abstracts, 2023, pp. EGU23–8733.
- [5] Y. Aydogdu and N. S. Namachchivaya, *A numerical study of stochastic el niño southern oscillations using wiener chaos expansion and monte carlo methods*, Tellus A: Dynamic Meteorology and Oceanography **76** (2024), no. 1.
- [6] M. A. Balmaseda, K. Mogensen, and A. T. Weaver, *Evaluation of the ecmwf ocean reanalysis system oras4*, Quarterly Journal of the Royal Meteorological Society **139** (2013), no. 674, 1132–1161.
- [7] R. N. Bannister, *A review of operational methods of variational and ensemble-variational data assimilation*, Quarterly Journal of the Royal Meteorological Society **143** (2017), no. 703, 607–633.
- [8] P. Bauer, A. Thorpe, and G. Brunet, *The quiet revolution of numerical weather prediction*, Nature **525** (2015), no. 7567, 47–55.
- [9] P. H. Baxendale, *Snapshots and pod*, 2024. Unpublished notes.
- [10] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, *Greedy layer-wise training of deep networks*, Advances in neural information processing systems, 2006.
- [11] N. Bergman, *Recursive bayesian estimation: Navigation and tracking applications*, Ph.D. Thesis, 1999.
- [12] J. Bjerknes, *Atmospheric teleconnections from the equatorial pacific*, Monthly Weather Review **97** (1969), no. 3, 163–172.
- [13] J. Blum, F. X. Le Dimet, and I. M. Navon, *Data assimilation for geophysical fluids*, Handbook of numerical analysis, 2009, pp. 385–441.
- [14] J. P. Boyd, *Equatorial solitary waves. part i: Rossby solitons*, Journal of Physical Oceanography **10** (1980), no. 11, 1699–1717.
- [15] J. Brajard, A. Carrassi, M. Bocquet, and L. Bertino, *Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the lorenz 96 model*, Journal of Computational Science **44** (2020), 101171.
- [16] S. L. Brunton, J. L. Proctor, and J. N. Kutz, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, Proceedings of the National Academy of Sciences **113** (2016), no. 15, 3932–3937.
- [17] R. H. Cameron and W. T. Martin, *The orthogonal development of non-linear functionals in series of fourier-hermite functionals*, Annals of Mathematics **48** (1947), 385–392.
- [18] A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen, *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*, Wiley Interdisciplinary Reviews: Climate Change **9** (2018), no. 5, e535.

- [19] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, *Data-driven discovery of coordinates and governing equations*, Proceedings of the National Academy of Sciences **116** (2019), no. 45, 22445–22451.
- [20] D. Chen, M. A. Cane, A. Kaplan, S. E. Zebiak, and D. Huang, *Predictability of el niño over the past 148 years*, Nature **428** (2004), no. 6984, 733–736.
- [21] N. Chen, F. Gilani, and J. Harlim, *A bayesian machine learning algorithm for predicting enso using short observational time series*, Geophysical Research Letters **48** (2021), no. 17, e2021GL093704.
- [22] N. Chen and A. J. Majda, *Predicting the cloud patterns for the boreal summer intraseasonal oscillation through a low-order stochastic model*, Mathematics of Climate and Weather Forecasting **1** (2015), no. 1.
- [23] N. Chen, A. J. Majda, and S. Thual, *Observations and mechanisms of a simple stochastic dynamical model capturing el niño diversity*, Journal of Climate **31** (2018), no. 1, 449–471.
- [24] A. J. Clarke, *An introduction to the dynamics of el niño and the southern oscillation*, Elsevier, 2008.
- [25] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, *Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps*, Proceedings of the National Academy of Sciences **102** (2005), no. 21, 7426–7431.
- [26] R. Courant, K. Friedrichs, and H. Lewy, *Über die partiellen differenzengleichungen der mathematischen physik*, Mathematische Annalen **100** (1928), no. 1, 32–74.
- [27] D. N. Daescu and I. M. Navon, *Efficiency of a pod-based reduced second-order adjoint model in 4d-var data assimilation*, International Journal for Numerical Methods in Fluids **53** (2007), no. 6, 985–1004.
- [28] B. M. de Silva, K. Champion, M. Quade, J. C. Loiseau, J. N. Kutz, and S. L. Brunton, *Pysindy: A python package for the sparse identification of nonlinear dynamics from data*, arXiv preprint arXiv:2004.08424 (2020).
- [29] H. A. Dijkstra, *The enso phenomenon: theory and mechanisms*, Advances in Geosciences **6** (2006), 3–15.
- [30] F. X. Le Dimet and O. Talagrand, *Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects*, Tellus A: Dynamic Meteorology and Oceanography **38** (1986), no. 2, 97–110.
- [31] S. M. Djouadi and S. Sahyoun, *On a generalization of the proper orthogonal decomposition and optimal construction of reduced order models*, 2012 american control conference (acc), 2012, pp. 1436–1441.
- [32] D. L. Donoho and C. Grimes, *Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data*, Proceedings of the National Academy of Sciences **100** (2003), no. 10, 5591–5596.
- [33] A. Dosovitskiy, *An image is worth 16x16 words: Transformers for image recognition at scale*, arXiv preprint arXiv:2010.11929 (2020).
- [34] A. Doucet and A. Lee, *Sequential monte carlo methods*, Handbook of graphical models, 2018, pp. 165–188.
- [35] P. N. Edwards, *History of climate modeling*, Wiley Interdisciplinary Reviews: Climate Change **2** (2011), no. 1, 128–139.
- [36] A. Einstein, *On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat*, Annalen der Physik **17** (1905), 549–560.
- [37] G. Evensen, *Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics*, Journal of Geophysical Research: Oceans **99** (1994), no. C5, 10143–10162.
- [38] F. Fang, C. C. Pain, I. M. Navon, G. J. Gorman, M. D. Piggott, P. A. Allison, and A. J. H. Goddard et al., *A pod reduced order unstructured mesh ocean modelling method for moderate reynolds number flows*, Ocean Modelling **28** (2009), no. 1-3, 127–136.
- [39] University Corporation for Atmospheric Research (UCAR), *Niño sst indices: Niño 1+2, 3, 3.4, 4, oni, and tni*, n.d. Climate Data Guide.
- [40] C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry and the natural sciences*, Springer series in synergetics, Springer, 1985.
- [41] R. G. Ghanem and P. D. Spanos, *Stochastic finite elements: a spectral approach*, Courier Corporation, 2003.
- [42] A. E. Gill, *Some simple solutions for heat-induced tropical circulation*, Quarterly Journal of the Royal Meteorological Society **106** (1980), no. 449, 447–462.

- [43] M. Girfoglio, A. Quaini, and G. Rozza, *A pod-galerkin reduced order model for a les filtering approach*, Journal of Computational Physics **436** (2021), 110260.
- [44] S. Godsill, *Particle filtering: the first 25 years and beyond*, Icassp 2019-2019 ieee international conference on acoustics, speech and signal processing (icassp), 2019, pp. 7760–7764.
- [45] N. J. Gordon, D. J. Salmond, and A. F. Smith, *Novel approach to nonlinear/non-gaussian bayesian state estimation*, Iee proceedings f (radar and signal processing), 1993April, pp. 107–113.
- [46] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Vol. 42, Springer Science & Business Media, 2013.
- [47] E. Guilyardi, A. Capotondi, M. Lengaigne, S. Thual, and A. T. Wittenberg, *Enso modeling: History, progress, and challenges*, El niño southern oscillation in a changing climate, 2020, pp. 199–226.
- [48] B. Gustafsson, H. O. Kreiss, and J. Oliger, *Time dependent problems and difference methods*, Vol. 24, John Wiley & Sons, 1995.
- [49] E. Halley, *An historical account of the trade winds, and monsoons, observable in the seas between and near the tropicks, with an attempt to assign the physical cause of the said winds*, Philosophical Transactions of the Royal Society of London **16** (1753), no. 183, 153–168.
- [50] K. Hasselmann, *Stochastic climate models part i. theory*, Tellus **28** (1976), no. 6, 473–485.
- [51] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, Proceedings of the ieee conference on computer vision and pattern recognition, 2016, pp. 770–778.
- [52] H. H. Hendon, M. C. Wheeler, and C. Zhang, *Seasonal dependence of the mjo–enso relationship*, Journal of Climate **20** (2007), no. 3, 531–543.
- [53] G. E. Hinton and R. R. Salakhutdinov, *Reducing the dimensionality of data with neural networks*, Science **313** (2006), no. 5786, 504–507.
- [54] T. V. Hoang, S. Krumscheid, H. G. Matthies, and R. Tempone, *Machine learning-based conditional mean filter: a generalization of the ensemble kalman filter for nonlinear data assimilation*, arXiv preprint (2021).
- [55] P. Holmes, *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge University Press, 2012.
- [56] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, Journal of Educational Psychology **24** (1933), no. 6, 417.
- [57] T. Y. Hou, W. Luo, B. Rozovskii, and H. M. Zhou, *Wiener chaos expansions and numerical solutions of randomly forced equations of fluid mechanics*, Journal of Computational Physics **216** (2006), no. 2, 687–706.
- [58] J. W. Hurrell, *Decadal trends in the north atlantic oscillation: Regional temperatures and precipitation*, Science **269** (1995), no. 5224, 676–679.
- [59] J. W. Hurrell and C. Deser, *North atlantic climate variability: the role of the north atlantic oscillation*, Journal of Marine Systems **79** (2010), no. 3-4, 231–244.
- [60] C. E. Iles, R. Vautard, J. Strachan, S. Joussaume, B. R. Eggen, and C. D. Hewitt, *The benefits of increasing resolution in global and regional climate simulations for european climate extremes*, Geoscientific Model Development **13** (2020), no. 11, 5583–5607.
- [61] A. Iollo, S. Lanteri, and J. A. Désidéri, *Stability properties of pod–galerkin approximations for the compressible navier–stokes equations*, Theoretical and Computational Fluid Dynamics **13** (2000), no. 6, 377–396.
- [62] M. Ishii and N. Mori, *d4pdf: large-ensemble and high-resolution climate simulations for global warming risk assessment*, Progress in Earth and Planetary Science **7** (2020), no. 1, 1–22.
- [63] K. Itô, *109. stochastic integral*, Proceedings of the Imperial Academy **20** (1944), no. 8, 519–524.
- [64] ———, *On a formula concerning stochastic differentials*, Nagoya Mathematical Journal **3** (1951), 55–65.
- [65] S. Janson, *Gaussian hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, 1997.

- [66] F. F. Jin, *An equatorial ocean recharge paradigm for enso. part i: Conceptual model*, Journal of the Atmospheric Sciences **54** (1997), no. 7, 811–829.
- [67] ———, *An equatorial ocean recharge paradigm for enso. part ii: A stripped-down coupled model*, Journal of the Atmospheric Sciences **54** (1997), no. 7, 830–847.
- [68] M. R. Jovanović, P. J. Schmid, and J. W. Nichols, *Sparsity-promoting dynamic mode decomposition*, Physics of Fluids **26** (2014), no. 2.
- [69] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME–Journal of Basic Engineering **82** (1960), no. 1, 35–45.
- [70] E. Kalnay, *Atmospheric modeling, data assimilation and predictability*, Vol. 341, Cambridge University Press, 2003.
- [71] E. A. Kalpinelli, N. E. Frangos, and A. N. Yannacopoulos, *Numerical methods for hyperbolic spdes: a wiener chaos approach*, Stochastic Partial Differential Equations: Analysis and Computations **1** (2013), 606–633.
- [72] K. Karhunen, *Über lineare methoden in der wahrscheinlichkeitsrechnung*, Annales Academiae Scientiarum Fennicae Series A1: Mathematica Physica **47** (1947).
- [73] E. Kharazmi, Z. Zhang, and G. E. Karniadakis, *Variational physics-informed neural networks for solving partial differential equations*, arXiv preprint arXiv:1912.00873 (2019).
- [74] K. Kira and L. A. Rendell, *A practical approach to feature selection*, Machine learning proceedings, 1992, pp. 249–256.
- [75] R. Kleeman and A. M. Moore, *A theory for the limitation of enso predictability due to stochastic atmospheric transients*, Journal of the Atmospheric Sciences **54** (1997), no. 6, 753–767.
- [76] P. E. Kloeden and E. Platen, *A survey of numerical methods for stochastic differential equations*, Stochastic Hydrology and Hydraulics **3** (1989), 155–178.
- [77] A. Kolmogoroff, *Über die analytischen methoden in der wahrscheinlichkeitsrechnung*, Mathematische Annalen **104** (1931), 415–458.
- [78] M. A. Kramer, *Nonlinear principal component analysis using autoassociative neural networks*, AIChE Journal **37** (1991), no. 2, 233–243.
- [79] H. R. Kunsch, *Particle filters*, Bernoulli **19** (2013), no. 4, 1391–1403.
- [80] H. J. Kushner, *Dynamical equations for optimal nonlinear filtering*, Journal of Differential Equations **3** (1967), no. 2, 179–190.
- [81] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: Data-driven modeling of complex systems*, Society for Industrial and Applied Mathematics, 2016.
- [82] P. S. Laplace, *Théorie analytique des probabilités*, Courcier, 1820.
- [83] M. Latif and T. P. Barnett, *Causes of decadal climate variability over the north pacific and north america*, Science **266** (1994), no. 5185, 634–637.
- [84] K. Law, A. Stuart, and K. Zygalakis, *Data assimilation*, Springer, Cham, Switzerland, 2015.
- [85] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, *Backpropagation applied to handwritten zip code recognition*, Neural Computation **1** (1989), no. 4, 541–551.
- [86] P. J. Van Leeuwen, H. R. Künsch, L. Nerger, R. Potthast, and S. Reich, *Particle filters for high-dimensional geoscience applications: A review*, Quarterly Journal of the Royal Meteorological Society **145** (2019), no. 723, 2335–2365.
- [87] M. Lengaigne, J. P. Boulanger, C. Menkes, P. Delecluse, and J. Slingo, *Westerly wind events in the tropical pacific and their influence on the coupled ocean-atmosphere system*, Earth Climate: The Ocean-Atmosphere Interaction, Geophys. Monogr **147** (2004), 49–69.
- [88] Y. Li, J. Li, W. Zhang, Q. Chen, J. Feng, F. Zheng, and X. Zhou, *Impacts of the tropical pacific cold tongue mode on enso diversity under global warming*, Journal of Geophysical Research: Oceans **122** (2017), no. 11, 8524–8542.

- [89] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, *Fourier neural operator for parametric partial differential equations*, arXiv preprint arXiv:2010.08895 (2020).
- [90] B. Liu, B. Gan, W. Cai, L. Wu, T. Geng, H. Wang, and F. Jia, *Will increasing climate model resolution be beneficial for enso simulation?*, Geophysical Research Letters **49** (2022), no. 11, e2021GL096932.
- [91] C. Liu, F. Wang, A. Köhl, X. Wang, C. Wang, and K. J. Richards, *Subsurface ocean turbulent mixing enhances central pacific enso*, Nature Communications **16** (2025), no. 1, 2315.
- [92] M. Loève, *Probability theory, vol. ii*, 4th ed., Graduate Texts in Mathematics, vol. 46, Springer-Verlag, 1978.
- [93] E. N. Lorenz, *Deterministic nonperiodic flow*, Journal of Atmospheric Sciences **20** (1963), no. 2, 130–141.
- [94] S. Lototsky, R. Mikulevicius, and B. L. Rozovskii, *Nonlinear filtering revisited: a spectral approach*, SIAM Journal on Control and Optimization **35** (1997), no. 2, 435–461.
- [95] S. Lototsky and B. Rozovskii, *Stochastic differential equations: a wiener chaos approach*, From stochastic calculus to mathematical finance, 2006, pp. 433–506.
- [96] S. V. Lototsky and B. L. Rozovsky, *Stochastic partial differential equations*, Vol. 11, Springer, Berlin, 2017.
- [97] J. L. Lumley, *Toward a turbulent constitutive relation*, Journal of Fluid Mechanics **41** (1970), no. 2, 413–434.
- [98] W. Luo, *Wiener chaos expansion and numerical solutions of stochastic partial differential equations*, Ph.D. Thesis, 2006.
- [99] R. A. Madden and P. R. Julian, *Observations of the 40–50-day tropical oscillation—a review*, Monthly Weather Review **122** (1994), no. 5, 814–837.
- [100] A. Majda, *Introduction to pdes and waves for the atmosphere and ocean*, Vol. 9, American Mathematical Society, 2003.
- [101] A. J. Majda and M. Branicki, *Lessons in uncertainty quantification for turbulent dynamical systems*, Discrete and Continuous Dynamical Systems-Series A **32** (2012), no. 9, 3133–3221.
- [102] A. J. Majda, C. Franzke, and B. Khouider, *An applied mathematics perspective on stochastic modelling for climate*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **366** (2008), no. 1875, 2427–2453.
- [103] A. J. Majda, I. Timofeyev, and E. Vanden Eijnden, *Models for stochastic climate prediction*, Proceedings of the National Academy of Sciences **96** (1999), no. 26, 14687–14691.
- [104] J. E. Marsden and M. McCracken, *The hopf bifurcation and its applications*, Vol. 19, Springer Science & Business Media, 2012.
- [105] G. Maruyama, *Continuous markov processes and stochastic equations*, Rendiconti del Circolo Matematico di Palermo **4** (1955), 48–90.
- [106] A. Masud and L. A. Bergman, *Solution of the four dimensional fokker-planck equation: Still a challenge*, Proceedings of icossar, 2005June, pp. 1911–1916.
- [107] K. McGuffie and A. Henderson-Sellers, *The climate modelling primer*, John Wiley & Sons, 2014.
- [108] M. J. McPhaden, *The tropical atmosphere ocean array is completed*, Bulletin of the American Meteorological Society **76** (1995), no. 5, 739–741.
- [109] M. J. McPhaden, A. Santoso, and W. Cai, *Introduction to el niño southern oscillation in a changing climate*, El niño southern oscillation in a changing climate, 2020, pp. 1–19.
- [110] C. R. Mechoso, J. D. Neelin, and J. Y. Yu, *Testing simple models of enso*, Journal of the Atmospheric Sciences **60** (2003), no. 2, 305–318.
- [111] M. A. Mendez, M. Raiola, A. Masullo, S. Discetti, A. Ianaro, R. Theunissen, and J. M. Buchlin, *Pod-based background removal for particle image velocimetry*, Experimental Thermal and Fluid Science **80** (2017), 181–192.
- [112] N. Metropolis and S. Ulam, *The monte carlo method*, Journal of the American Statistical Association **44** (1949), no. 247, 335–341.

- [113] I. Mezić, *Spectral properties of dynamical systems, model reduction and decompositions*, Nonlinear Dynamics **41** (2005), 309–325.
- [114] G. N. Mil'shtejn, *Approximate integration of stochastic differential equations*, Theory of Probability & Its Applications **19** (1975), no. 3, 557–562.
- [115] R. Mizuta, A. Murata, M. Ishii, H. Shiogama, K. Hibino, N. Mori, and M. Kimoto et al., *Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models*, Bulletin of the American Meteorological Society **98** (2017), no. 7, 1383–1398.
- [116] C. Moser, N. Chen, and Y. Zhang, *A stochastic conceptual model for the coupled enso and mjo*, arXiv preprint **2411.05264** (2024).
- [117] A. Mukherjee, Y. Aydogdu, T. Ravichandran, and N. Sri Namachchivaya, *Stochastic parameterization using compressed sensing: Application to the lorenz-96 atmospheric model*, Tellus A: Dynamic Meteorology and Oceanography **74** (2022), no. 1, 300–317.
- [118] M. T. Nakao, M. Plum, and Y. Watanabe, *Numerical verification methods and computer-assisted proofs for partial differential equations*, Springer, 2019.
- [119] J. D. Neelin, *Climate change and climate modeling*, Cambridge University Press, 2010.
- [120] J. D. Neelin, D. S. Battisti, A. C. Hirst, F. F. Jin, Y. Wakata, T. Yamagata, and S. E. Zebiak, *Enso theory*, Journal of Geophysical Research: Oceans **103** (1998), no. C7, 14261–14290.
- [121] J. D. Neelin, M. Latif, and F. F. Jin, *Dynamics of coupled ocean-atmosphere models: The tropical problem*, Annual Review of Fluid Mechanics **26** (1994), no. 1, 617–659.
- [122] M. Newman, P. D. Sardeshmukh, and C. Penland, *How important is air–sea coupling in enso and mjo evolution?*, Journal of Climate **22** (2009), no. 11, 2958–2977.
- [123] A. Novelia, Y. Aydogdu, T. Ravichandran, and N. Namachchivaya, *Compressed compressor*, arXiv preprint arXiv:2107.08938 (2021).
- [124] D. Nualart, *The malliavin calculus and related topics*, Vol. 1995, Springer, Berlin, 2006.
- [125] T. Palmer and B. Stevens, *The scientific challenge of understanding and estimating climate change*, Proceedings of the National Academy of Sciences **116** (2019), no. 49, 24390–24395.
- [126] K. Pearson, *Liii. on lines and planes of closest fit to systems of points in space*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **2** (1901), no. 11, 559–572.
- [127] S. G. Penny and T. M. Hamill, *Coupled data assimilation for integrated earth system analysis and prediction*, Bulletin of the American Meteorological Society **98** (2017), no. 7, ES169–ES172.
- [128] L. Pichler, A. Masud, and L. A. Bergman, *Numerical solution of the fokker–planck equation by finite difference and finite element methods—a comparative study*, Computational methods in stochastic dynamics: Volume 2, 2013, pp. 69–85.
- [129] J. L. Proctor, S. L. Brunton, and J. N. Kutz, *Generalizing koopman theory to allow for inputs and control*, SIAM Journal on Applied Dynamical Systems **17** (2018), no. 1, 909–930.
- [130] H. Qian, G. Yin, and Q. Zhang, *A new computational method for nonlinear filtering*, Collections **56** (2023), no. 05.
- [131] F. Rabier, *Overview of global data assimilation developments in numerical weather-prediction centres*, Quarterly Journal of the Royal Meteorological Society **131** (2005), no. 613, 3215–3233.
- [132] F. Rabier and Z. Liu, *Variational data assimilation: theory and overview*, Proceedings of the ecmwf seminar on recent developments in data assimilation for atmosphere and ocean, 2003September, pp. 29–43.
- [133] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, Journal of Computational Physics **378** (2019), 686–707.
- [134] D. A. Randall, R. A. Wood, S. Bony, R. Colman, T. Fiechfet, J. Fyfe, et al., *Climate models and their evaluation*, Climate change 2007: The physical science basis, 2007, pp. 589–662.

- [135] C. P. Robert and G. Casella, *Monte carlo statistical methods*, Vol. 2, Springer, New York, 1999.
- [136] M. J. Roberts, P. L. Vidale, C. Senior, H. T. Hewitt, C. Bates, S. Berthou, and M. F. Wehner et al., *The benefits of global high resolution for climate simulation: process understanding and the enabling of stakeholder decisions at the regional scale*, Bulletin of the American Meteorological Society **99** (2018), no. 11, 2341–2359.
- [137] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, *Spectral analysis of nonlinear flows*, Journal of Fluid Mechanics **641** (2009), 115–127.
- [138] R. Y. Rubinstein and D. P. Kroese, *Simulation and the monte carlo method*, John Wiley & Sons, 2016.
- [139] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, *Data-driven discovery of partial differential equations*, Science Advances **3** (2017), no. 4, e1602614.
- [140] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature **323** (1986), no. 6088, 533–536.
- [141] O. San and T. Iliescu, *Proper orthogonal decomposition closure models for fluid flows: Burgers equation*, arXiv preprint arXiv:1308.3276 (2013).
- [142] P. J. Schmid, *Dynamic mode decomposition of numerical and experimental data*, Journal of Fluid Mechanics **656** (2010), 5–28.
- [143] G. A. Schmidt and S. Sherwood, *A practical philosophy of complex climate modelling*, European Journal for Philosophy of Science **5** (2015), 149–169.
- [144] T. Schneider, S. Lan, A. Stuart, and J. Teixeira, *Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations*, Geophysical Research Letters **44** (2017), no. 24, 12–396.
- [145] D. Simon, *Optimal state estimation: Kalman, h infinity, and nonlinear approaches*, John Wiley & Sons, 2006.
- [146] J. Sirignano and K. Spiliopoulos, *Dgm: A deep learning algorithm for solving partial differential equations*, Journal of Computational Physics **375** (2018), 1339–1364.
- [147] L. Sirovich, *Turbulence and the dynamics of coherent structures. i. coherent structures*, Quarterly of Applied Mathematics **45** (1987), no. 3, 561–571.
- [148] M. A. Spall and J. Pedlosky, *Reflection and transmission of equatorial rossby waves*, Journal of Physical Oceanography **35** (2005), no. 3, 363–373.
- [149] B. F. Spencer and L. A. Bergman, *On the numerical solution of the fokker-planck equation for nonlinear stochastic systems*, Nonlinear Dynamics **4** (1993), 357–372.
- [150] R. L. E. Stratonovich, *Conditional markov processes*, Non-linear transformations of stochastic processes, 1965, pp. 427–453.
- [151] M. J. Suarez and P. S. Schopf, *A delayed action oscillator for enso*, Journal of Atmospheric Sciences **45** (1988), no. 21, 3283–3287.
- [152] Y. Tang and B. Yu, *Mjo and its relationship to enso*, Journal of Geophysical Research: Atmospheres **113** (2008), no. D14.
- [153] J. B. Tenenbaum, V. D. Silva, and J. C. Langford, *A global geometric framework for nonlinear dimensionality reduction*, Science **290** (2000), no. 5500, 2319–2323.
- [154] C. J. Thompson and D. S. Battisti, *A linear stochastic dynamical model of enso. part i: Model development*, Journal of Climate **13** (2000), no. 15, 2818–2832.
- [155] S. Thual, A. J. Majda, N. Chen, and S. N. Stechmann, *Simple stochastic model for el niño with westerly wind bursts*, Proceedings of the National Academy of Sciences **113** (2016), no. 37, 10245–10250.
- [156] G. K. Vallis, *Atmospheric and oceanic fluid dynamics*, Cambridge University Press, 2017.
- [157] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, *Extracting and composing robust features with denoising autoencoders*, Proceedings of the 25th international conference on machine learning, 2008, pp. 1096–1103.

- [158] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P. A. Manzagol, and L. Bottou, *Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion*, Journal of Machine Learning Research **11** (2010), no. 12.
- [159] B. Wang, X. Luo, Y. M. Yang, W. Sun, M. A. Cane, W. Cai, and J. Liu, *Historical change of el niño properties sheds light on future changes of extreme el niño*, Proceedings of the National Academy of Sciences **116** (2019), no. 45, 22512–22517.
- [160] Z. Wang, I. Akhtar, J. Borggaard, and T. Iliescu, *Proper orthogonal decomposition closure models for turbulent flows: A numerical comparison*, Computer Methods in Applied Mechanics and Engineering **237** (2012), 10–26.
- [161] T. T. Warner, *Numerical weather and climate prediction*, Cambridge University Press, 2010.
- [162] C. Wengel, S. S. Lee, M. F. Stuecker, A. Timmermann, J. E. Chu, and F. Schloesser, *Future high-resolution el niño/southern oscillation dynamics*, Nature Climate Change **11** (2021), no. 9, 758–765.
- [163] N. Wiener, *Differential-space*, Journal of Mathematics and Physics **2** (1923), no. 1-4, 131–174.
- [164] ———, *The homogeneous chaos*, American Journal of Mathematics **60** (1938), no. 4, 897–936.
- [165] ———, *Extrapolation, interpolation, and smoothing of stationary time series: With engineering applications*, The MIT Press, 1949.
- [166] ———, *Nonlinear problems in random theory*, 1966.
- [167] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, *A data-driven approximation of the koopman operator: Extending dynamic mode decomposition*, Journal of Nonlinear Science **25** (2015), 1307–1346.
- [168] D. Xiu and G. E. Karniadakis, *The wiener-asky polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing **24** (2002), no. 2, 619–644.
- [169] H. C. Yeong, *Dimensional reduction in nonlinear estimation of multiscale systems*, Ph.D. Thesis, 2017.
- [170] M. Zakai, *On the optimal filtering of diffusion processes*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **11** (1969), no. 3, 230–243.
- [171] S. E. Zebiak and M. A. Cane, *A model el niño-southern oscillation*, Monthly Weather Review **115** (1987), no. 10, 2262–2278.
- [172] C. Zhang, *Madden-julian oscillation*, Reviews of Geophysics **43** (2005), no. 2.

APPENDICES

Appendix A

A Complete El Niño-Southern Oscillation Model

A.1 A Complete Coupled Model

This section provides a more complete model of ENSO used in this thesis. The model is introduced by Majda and co-workers [23, 155] and are defined as:

$$\mathbf{U} \stackrel{\text{def}}{=} [u(t, x, y), v(t, x, y), \theta(t, x, y), U(t, x, Y), V(t, x, Y), H(t, x, Y), T(t, x, y)]$$

and can be written in differential form as

$$\mathcal{M} \frac{\partial \mathbf{U}}{\partial t}(t, x, y) = \mathcal{A}(\kappa) \mathbf{U}(t, x, y) + \mathcal{N}(\mathbf{U}, \mathbf{U}, \mu) \quad (\text{A.1})$$

where $\mathcal{M}$ is a 7×7 constant matrix with diagonals $[0, 0, 0, 1, 1, 1, 1]$, the linear operator $\mathcal{A}(\kappa)$ defined as

$$\mathcal{A}(\kappa) = \begin{bmatrix} 0 & y & \partial_x & 0 & 0 & 0 & 0 \\ -y & 0 & \partial_y & 0 & 0 & 0 & 0 \\ \partial_x & \partial_y & 0 & 0 & 0 & 0 & \alpha_q/(1 - \overline{Q}) \\ \kappa c_1 & 0 & 0 & 0 & c_1 Y & -c_1 \partial_x & 0 \\ 0 & 0 & 0 & -Y & -\partial_Y & 0 & 0 \\ 0 & 0 & 0 & -c_1 \partial_x & -c_1 \partial_Y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_1 \eta & c_1 \zeta \alpha_q \end{bmatrix}$$

depends on a parameter κ , the wind stress coefficient, and nonlinearity $\mathcal{N}(u, u, \mu)$ with advection coefficient μ , is defined by

$$\mathcal{N}(\mathbf{U}, \mathbf{U}, \mu) \stackrel{\text{def}}{=} [0, 0, 0, 0, 0, 0, -\mu \frac{\partial}{\partial x}(UT)]^T.$$

In the above model (A.1), x is zonal direction. y and Y are meridional direction in the atmosphere and ocean, respectively. They are different in the atmosphere and ocean as they each scale to a suitable Rossby radius. u and v are zonal and meridional atmospheric winds, respectively, while θ is potential temperature. U and V are zonal and meridional oceanic currents, while H is thermocline depth. Finally, T represent sea surface temperature (SST) anomalies. The atmosphere extends over

the entire equatorial belt $0 \leq x \leq L_A$ with periodic conditions $u(t, 0, y) = u(t, L_A, y)$ whereas the Pacific Ocean extends over $0 \leq x \leq L_O$ with reflective boundary condition $\int_{-\infty}^{\infty} U(t, 0, y) dY = 0$ and $U(t, L_O, Y) = 0$. We modify these boundary conditions and provide a slightly novel ENSO model in this research. The model parameters c_1 , κ , η , ξ and α_q are defined in Chapter (2). After some calculations, it can be shown that ENSO model can be projected onto parabolic cylinder functions due to non-constant Coriolis constant at low-latitudes. These functions are provided in the next section.

A.2 Parabolic Cylinder Functions

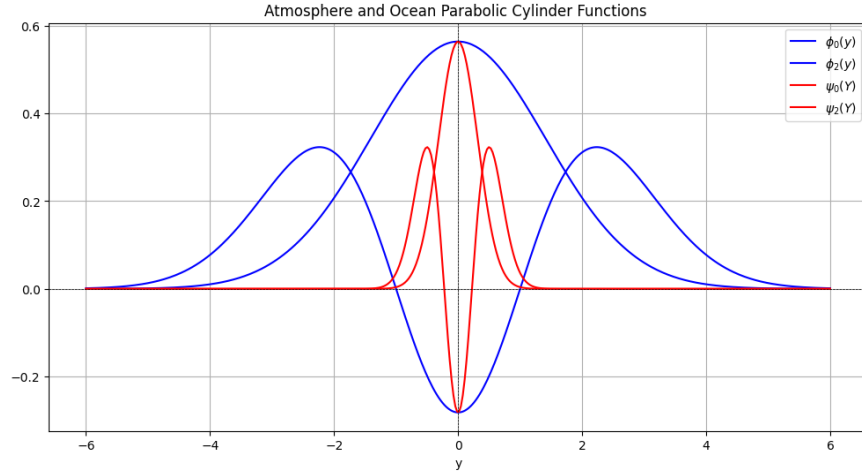


Figure A.1: Meridional profiles of atmosphere ($\phi_{0,2}(y)$) and ocean ($\Psi_{0,2}(Y)$) parabolic cylinder functions.

After solving the system (2.3) in Chapter (2) solution $\mathbf{U}(t, x, y)$ can be constructed by using corresponding atmospheric and oceanic parabolic cylinder functions, as illustrated in Figure (A.1). The first atmospheric parabolic functions read $\phi_0(y) = (\pi)^{-1/4} \exp(-y^2/2)$, $\phi_2(y) = (4\pi)^{-1/4} (2y^2 - 1) \exp(-y^2/2)$, while the ocean parabolic functions read $\Psi_m(Y)$ where $Y = y/\sqrt{c}$. The parabolic cylinder functions in the ocean and atmosphere differ due to the Rossby radius. The reconstructed variables reads :

$$\begin{aligned}
 u(t, x, y) &= (K^A(t, x) - R^A(t, x))\phi_0(y) + (R^A(t, x)/\sqrt{2})\phi_2(y) \\
 \theta(t, x, y) &= -(K^A(t, x) - R^A(t, x))\phi_0(y) - (R^A(t, x)/\sqrt{2})\phi_2(y) \\
 U(t, x, Y) &= (K^O(t, x) - R^O(t, x))\Psi_0(Y) + (R^O(t, x)/\sqrt{2})\Psi_2(Y) \\
 H(t, x, Y) &= (K^O(t, x) + R^O(t, x))\Psi_0(Y) + (R^O(t, x)/\sqrt{2})\Psi_2(Y)
 \end{aligned} \tag{A.2}$$

It should be noted that we focus on the solution of

$$\mathbf{u}(t, x) \stackrel{\text{def}}{=} [K^A(t, x), R^A(t, x), K^O(t, x), R^O(t, x), T(t, x)]^T$$

in this thesis.

Appendix B

Hermite Polynomials and Wiener Chaos Expansion

B.1 Hermite Polynomials

This section provides the properties of Hermite polynomials, which is used in Chapter (3), where we describe the deterministic and stochastic components of the WCE. Hermite polynomials are closely related to Gaussian random variables through their role as orthogonal polynomials with respect to the Gaussian probability measure. Specifically, Hermite polynomials form an orthogonal basis for functions of Gaussian random variables, which makes them essential in expansions like the Wiener Chaos Expansion (WCE). Consider functions on the real axis $\mathbb{R} = (-\infty, \infty)$ equipped with the Gaussian measure

$$\mu(dx) = \rho(x)dx \quad \rho(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \quad (\text{B.1})$$

where dx is the Lebesgue measure. The space of square integrable functions with Gaussian measure μ can be denoted as

$$L^2(\mathbb{R}, \mu) = \left\{ f(x); \int_{-\infty}^{\infty} f(x)^2 \mu(dx) < \infty \right\}$$

The inner product on this space is defined as

$$(f, g)_{\mu} = \int_{-\infty}^{\infty} f(x)g(x)\mu(dx) = \int_{-\infty}^{\infty} f(x)g(x)\rho(x)dx$$

If ξ is a Gaussian random variable with distribution $N(0, 1)$ then

$$(f, g)_{\mu} = \mathbb{E}[f(\xi)g(\xi)]$$

where $\mathbb{E}$ denotes the expectation operator. The un-normalized Hermite polynomials are defined as

$$P_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right) \quad n = 0, 1, 2, \dots \quad (\text{B.2})$$

$P_n(x)$ are orthogonal polynomials with respect to the Gaussian measure

$$(P_n, P_m)_\mu = \mathbb{E}[P_n(\xi)P_m(\xi)] = n!\delta_{n,m}$$

Hence the normalized Hermite polynomials are defined as

$$H_n(x) = \frac{P_n(x)}{\sqrt{n!}} = (n!)^{-\frac{1}{2}}(-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right) \quad (\text{B.3})$$

It is well known that $\{H_n(x); n = 0, 1, \dots\}$ are a complete orthonormal basis in the Hilbert space $L^2(\mathbb{R}, \mu)$. Since $V_0(x) = 1$ particularly we have

$$\mathbb{E}[H_n(\xi)] = \int_{-\infty}^{\infty} H_n(x) d\mu(x) = (H_n, 1)_\mu = 0 \quad \text{if } n \neq 0 \quad (\text{B.4})$$

So a Hermite polynomial of a standard Gaussian random variable has zero mean if its order is bigger than zero. This fact is helpful for further calculations. Like most orthogonal polynomials, Hermite polynomials have a generating function

$$\psi(x, z) = e^{-\frac{z^2}{2} + xz} \quad (\text{B.5})$$

Expanding $\psi(x, z)$ into a Taylor series of variable z , we have

$$\psi(x, z) = \sum_{n=0}^{\infty} \frac{\partial^n \psi(x, z)}{\partial z^n} \Big|_{z=0} \left(\frac{z^n}{n!} \right)$$

On the other hand,

$$\psi(x, z) = e^{\frac{x^2}{2}} e^{-\frac{(z-x)^2}{2}} = e^{\frac{x^2}{2}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right) z^n = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} z^n \quad (\text{B.6})$$

So the coefficients of the Taylor expansion of $\psi(x, z)$ are exactly the un-normalized Hermite polynomials.

$$P_n(x) = \frac{\partial^n \psi(x, z)}{\partial z^n} \Big|_{z=0} \quad (\text{B.7})$$

Differentiating (B.2) directly we get

$$P'_n(x) = xP_n(x) - P_{n+1}(x) \quad (\text{B.8})$$

On the other hand, by differentiating both sides of (B.6) with respect to x we have

$$\frac{\partial}{\partial x} \psi(x, z) = z\psi(x, z) = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} z^{n+1} = \sum_{n=0}^{\infty} \frac{P'_n(x)}{n!} z^n$$

Shifting the summation index and comparing the coefficients of z^n suggests that

$$P'_n(x) = nP_{n-1}(x) \quad (\text{B.9})$$

Combining (B.8) and (B.9) together, we get the recursive relation of the un-normalized Hermite

polynomials

$$P_{n+1}(x) - xP_n(x) + nP_{n-1}(x) = 0 \quad n = 1, 2, \dots \quad (\text{B.10})$$

with $P_{-1}(x) = 0$, $P_0(x) = 1$. From the recursive relation, we can easily get the first six un-normalized Hermite polynomials

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = x^2 - 1$$

$$P_3(x) = x^3 - 3x$$

$$P_4(x) = x^4 - 6x^2 + 3$$

$$P_5(x) = x^5 - 10x^3 + 15x$$

For the normalized Hermite polynomials, the recursive relation becomes

$$\sqrt{n+1}H_{n+1}(x) - xH_n(x) + \sqrt{n}H_{n-1}(x) = 0 \quad (\text{B.11})$$

$$H_{-1}(x) = 0 \quad V_0(x) = 1 \quad (\text{B.12})$$

and the differential rule is

$$H'_n(x) = \sqrt{n}H_{n-1}(x) \quad (\text{B.13})$$

Appendix C

Reduced-Order Equations of the El Niño-Southern Oscillation Model

C.1 Invariance of the POD Modes

In this section, we provide supplementary results in addition to those obtained in Chapter (4). We generate five ensemble simulations of model (4.2) using different initial conditions. Figure (C.1) validates that POD modes $\hat{\Psi}_\alpha(x)$ are invariant as they maintain similar pattern regardless of initial conditions. Such modes only change the direction due to the eigenvectors, but they preserve the magnitude. Depending on the initial condition POD coefficients $q_\alpha(t)$ have different magnitude while maintaining the same frequency, as shown in Figure (C.2).

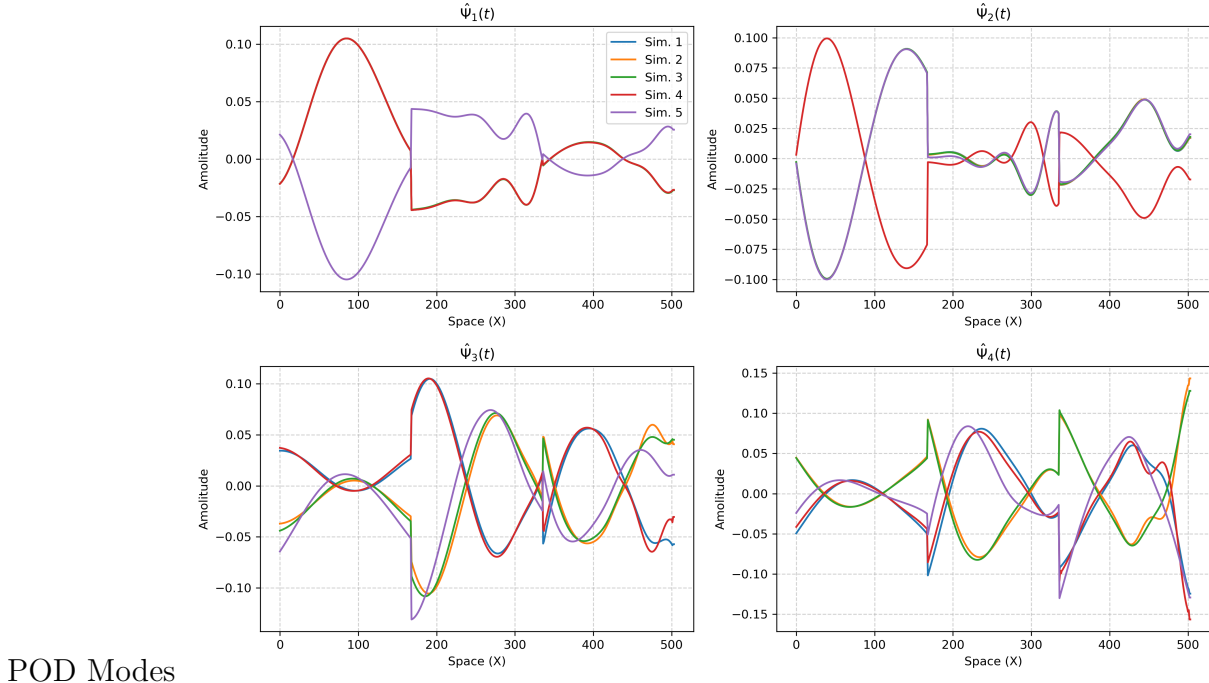


Figure C.1: POD modes $\hat{\Psi}_\alpha(x)$ for $\alpha = 1, \dots, 4$ obtained from simulations with different initial conditions.

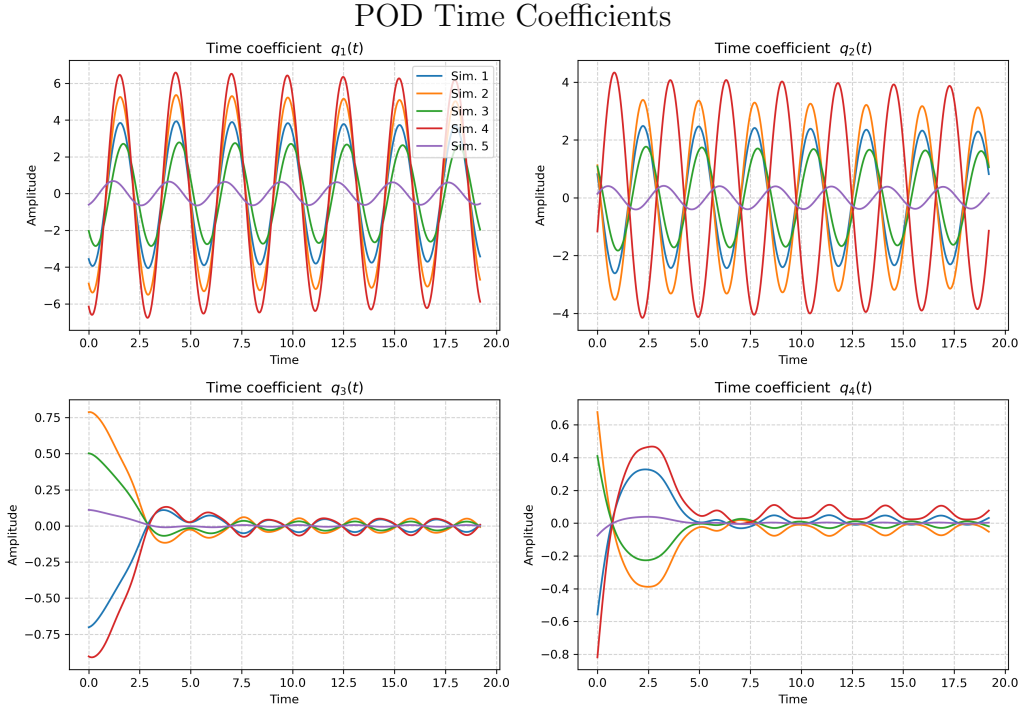


Figure C.2: POD time coefficients $q_\alpha(t)$ for $\alpha = 1, \dots, 4$ obtained from simulations with different initial conditions.

C.2 Reduced Order Equations

In this section, we provide the explicit form of the reduced-order equations of the ENSO model (4.2) introduced in Chapter (4). These ROMs are obtained using POD-Galerkin, POD-SINDy and Autoencoder-SINDy methods.

C.2.1 POD-Galerkin Reduced Order Equations

One of the main results of this thesis is the reduced order modeling based on the POD-Galerkin method. The ROM given by (4.21) has the following explicit form:

$$\begin{bmatrix} \dot{a}_1 \\ \dot{a}_2 \\ \dot{a}_3 \\ \dot{a}_4 \end{bmatrix} = \begin{bmatrix} -0.0073 & 3.6551 & 1.0371 & 0.5446 \\ -1.4313 & 0.0159 & 0.2480 & -0.6435 \\ -0.0208 & -0.0072 & -0.4296 & 0.6164 \\ -0.0057 & -0.0045 & -0.9101 & -0.5193 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \mu \begin{bmatrix} f_1(a_1, a_2, a_3, a_4) \\ f_2(a_1, a_2, a_3, a_4) \\ f_3(a_1, a_2, a_3, a_4) \\ f_4(a_1, a_2, a_3, a_4) \end{bmatrix} \quad (\text{C.1})$$

$f_1 =$		a_1	a_2	a_3	a_4	$f_2 =$		a_1	a_2	a_3	a_4
	a_1	0.0027	0.0080	0.0108	0.0240		a_1	-0.0152	0.0204	0.0079	0.0056
	a_2	0.0097	-0.0038	-0.0072	-0.0183		a_2	-0.0114	0.0163	-0.0075	-0.0221
	a_3	-0.0073	0.0064	0.0002	0.0071		a_3	-0.0726	0.0088	0.0014	0.0253
	a_4	-0.0416	0.0178	0.0002	0.0169		a_4	-0.0145	0.0275	-0.0122	-0.0033

$f_3 =$		a_1	a_2	a_3	a_4	$f_4 =$		a_1	a_2	a_3	a_4
	a_1	0.0079	0.0271	0.0120	0.0514		a_1	0.0444	0.0105	0.0447	0.0500
	a_2	0.0533	-0.0108	-0.0069	-0.0754		a_2	-0.0059	-0.0272	-0.0430	-0.0706
	a_3	-0.0056	0.0338	-0.0194	0.0367		a_3	0.1138	0.0522	-0.0120	0.0430
	a_4	-0.1363	0.0437	0.0007	0.0778		a_4	-0.0133	0.0601	-0.0304	0.0829

The coefficients in Equation (C.1) are calculated using the inner products $\langle \mathcal{A}_0 \hat{\Psi}_\alpha, \tilde{\Psi}_\gamma \rangle$ for the linear part and $\langle \mathcal{N}_0(\hat{\Psi}_\alpha, \hat{\Psi}_\beta), \tilde{\Psi}_\gamma \rangle$ for the nonlinear part, where $\alpha = 1, \dots, 4$ and $\gamma = 1, \dots, 4$. The comparison of these equations $a_\alpha(t)$ and POD coefficients $q_\alpha(t)$ are shown in Figure (4.3).

C.2.2 POD-SINDy Reduced Order Equations

The second reduced-order equations in this research is explored through SINDy method, as introduced in Section (4.6). To apply SINDy, we first use the POD time coefficients $q_\alpha(t)$, which are time series, and plug them into the Equation (4.27). This algorithm is processed using built-in Python *Pysindy* package [28].

	$\dot{s}_1$	$\dot{s}_2$	$\dot{s}_3$	$\dot{s}_4$
s_1	-0.0224	-1.4335	-0.0197	-0.0105
s_2	3.6573	-0.0205	-0.0083	-0.0058
s_3	0.6418	0.3688	-0.5479	-0.5332
s_4	0	0	0	0
s_1^2	0.0007	-0.0001	0.0015	0.0021
$s_1 s_2$	0.0019	0	0.0036	-0.0004
$s_1 s_3$	-0.0564	0.0770	-0.0713	0.0651
$s_1 s_4$	-0.0739	0	0	0.0723
s_2^2	0.0001	0.0006	0.0039	-0.0006
$s_2 s_3$	0.0996	-0.0025	0.0101	-0.0829
$s_2 s_4$	0	0.1020	-0.0941	0
s_3^2	0	0	0	0
$s_3 s_4$	0	0	0	0
s_4^2	0	0	0	0

Table C.1: Coefficients of linear and quadratic terms in the system of the POD-SINDy equations.

It should be noted that there is a good agreement between the largest coefficients of Equations (C.1) and Table (C.1). While the coefficient of a_2 term in $\dot{a}_1$ (C.1) is 3.6551, the coefficient of s_2

term in $\dot{s}_1$ Table (C.1) is 3.6573. Moreover, while the coefficient of a_1 term in $\dot{a}_2$ (C.1) is -1.4313 , the coefficient of s_1 term in $\dot{s}_2$ in Table (C.1) is -1.4335 . These findings validate the use of these different ROM approaches as other coefficients are relatively small to compare. The comparison of POD coefficients $q_\alpha(t)$ and simulation of SINDy equations $s_\alpha(t)$ is illustrated in Figure (C.3).

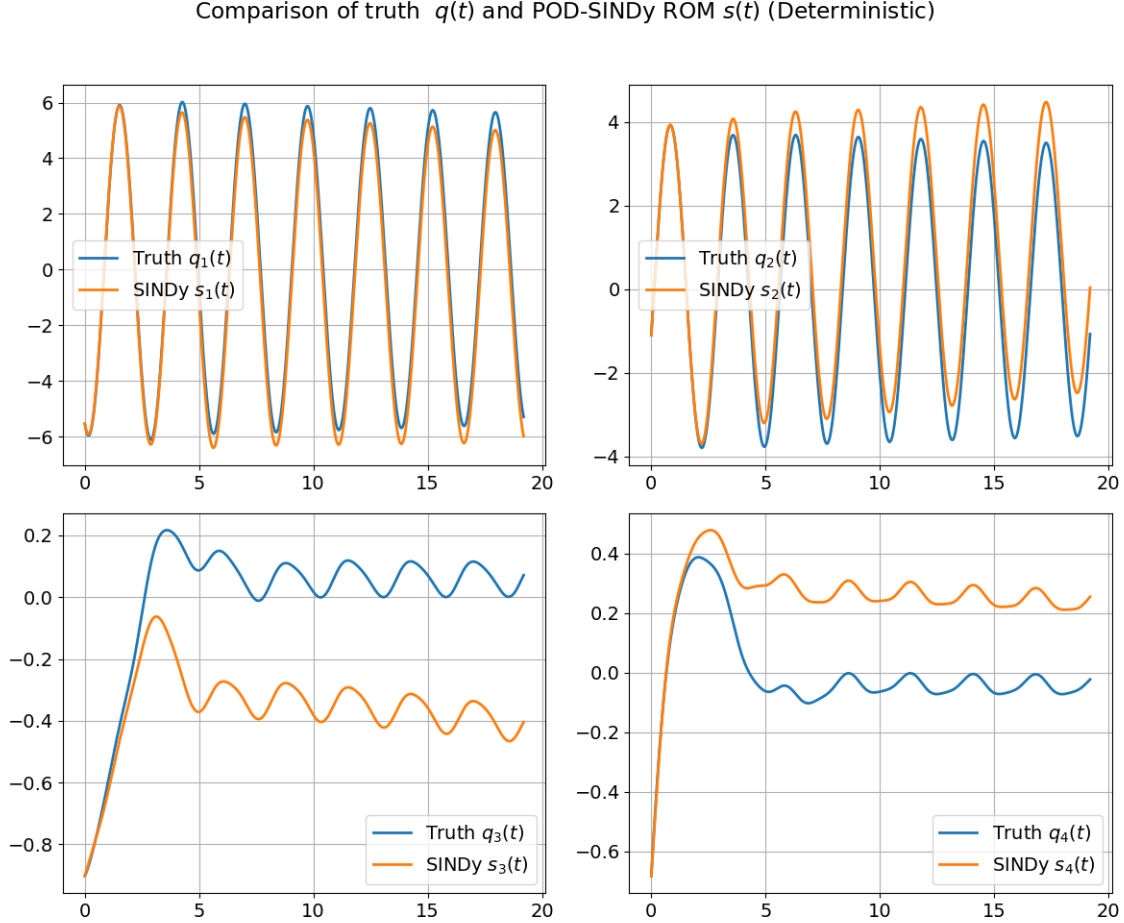


Figure C.3: Comparison of POD time coefficients $q_\alpha(t)$ and the simulation of the POD-SINDy equations $s_\alpha(t)$.

C.2.3 Autoencoder-SINDy Reduced Order Equations

The final reduced-order equations in this research is explored through autoencoders and SINDy methods, as introduced in Section (4.5) (4.6). While the full-order model (4.2) is reduced using autoencoders, the model equations are explored by employing SINDy. The AE-SINDy reduced-order equations are the followings:

	$\dot{s}_1$	$\dot{s}_2$	$\dot{s}_3$	$\dot{s}_4$
s_1	0	0	0	0
s_2	0	0	0	0
s_3	0.6009	-0.7532	-0.3643	-2.5726
s_4	-1.7731	0.7783	2.0848	0.3260
s_1^2	0	0	0	0
$s_1 s_2$	0	0	0	0
$s_1 s_3$	0	0	0	0
$s_1 s_4$	0	0	0	0
s_2^2	0	0	0	0
$s_2 s_3$	0	0	0	0
$s_2 s_4$	0	0	0	0
s_3^2	-0.0005	0	0.0005	0.0006
$s_3 s_4$	0.0013	0	-0.00162	0
s_4^2	0.0015	-0.0008	-0.0017	-0.0012

Table C.2: Coefficients of linear and quadratic terms in the system of the AE-SINDy equations.

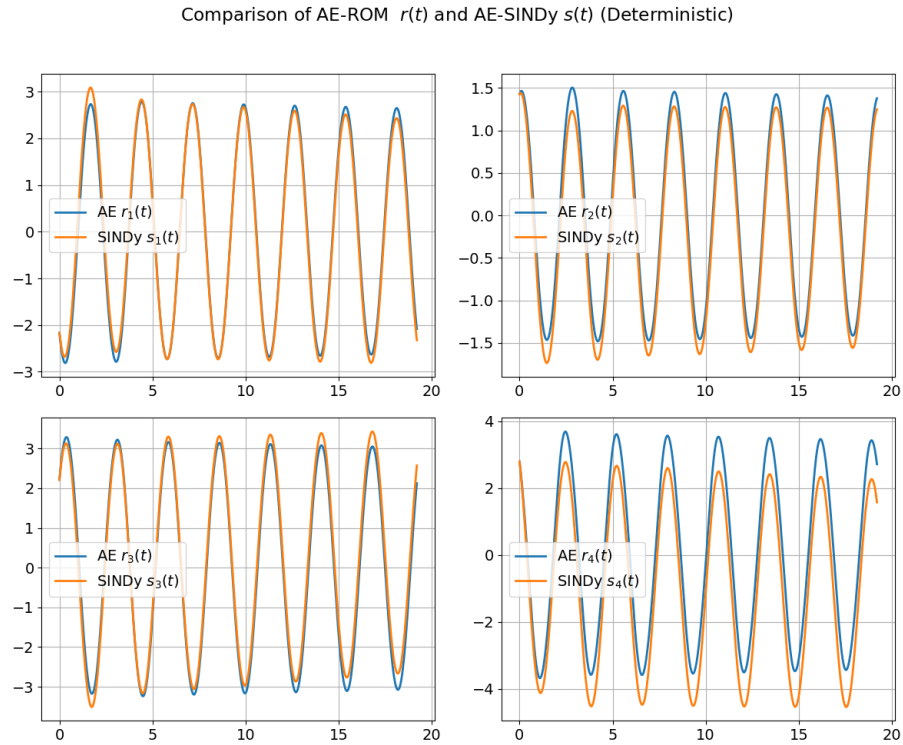


Figure C.4: Comparison of autoencoder latent space $r_\alpha(t)$ and the simulation of the AE-SINDy equations $s_\alpha(t)$.

The comparison of autoencoder reduced time series $r_\alpha(t)$ and simulation of SINDy equations $s_\alpha(t)$ is illustrated in Figure (C.4).

C.3 Hopf Bifurcation

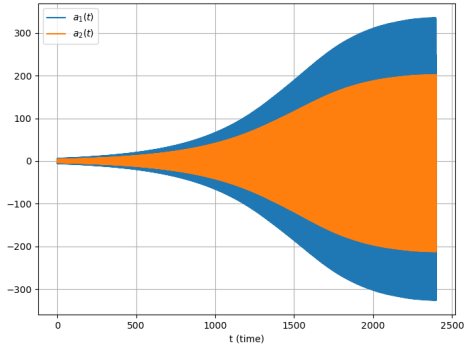
A Hopf bifurcation occurs when a system transitions from a stable equilibrium to a periodic orbit as a system parameter is varied. This bifurcation is typically observed in nonlinear dynamical systems and is characterized by the appearance or disappearance of oscillatory behavior. Consider a dynamical system governed by a set of ordinary differential equations:

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, \mu)$$

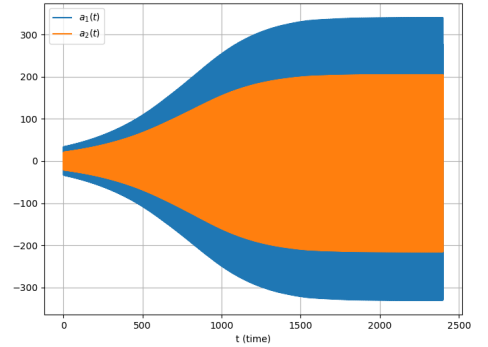
where $\mathbf{x} \in \mathbb{R}^n$ is the state vector, μ is a real parameter, and f is a smooth vector field. A Hopf bifurcation occurs under the following conditions: (i) The system has a fixed point $\mathbf{x}_0$ for all values of μ and (ii) a pair of complex-conjugate eigenvalues $\lambda(\mu)$ and $\bar{\lambda}(\mu)$ of the Jacobian matrix $Df(\mathbf{x}_0, \mu)$ cross the imaginary axis as μ passes through a critical value μ_0 . Specifically,

$$\Re(\lambda(\mu)) = 0 \quad \text{at} \quad \mu = \mu_0$$

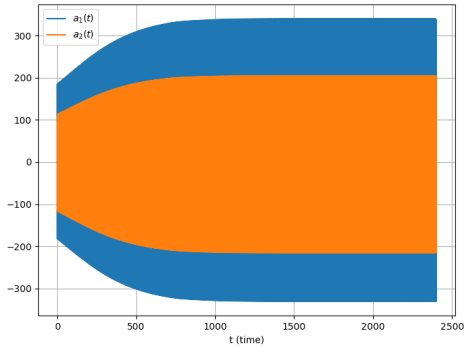
which leads to oscillatory behavior as μ changes [46]. Figure (C.5) illustrates the dynamics of reduced-order model (C.1), which has critical eigenvalues and is captured by the nonlinearity μ as $t \rightarrow \infty$.



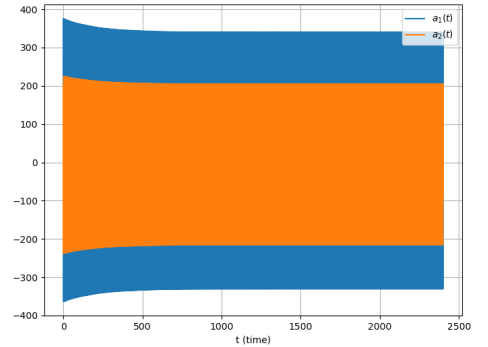
(a) $[a_1(0), a_2(0)] = [-5.5706, -1.0802]$



(b) $[a_1(0), a_2(0)] = [14.4293, 18.9197]$



(c) $[a_1(0), a_2(0)] = [94.4293, 98.9197]$



(d) $[a_1(0), a_2(0)] = [194.4293, 198.9197]$

Figure C.5: Dynamics of reduced-order model (C.1) as $t \rightarrow \infty$.

Appendix D

Nonlinear Filtering of the El Niño-Southern Oscillation (ENSO)

D.1 Validation of POD-Galerkin ROMs of the Stochastic ENSO Model

In this section, we provide supplementary results in addition to those obtained in Chapter (5). Since we are working with SPDEs, it is important to generate ensembles and validate the proposed method. For this purpose, we generated ten ensembles of the model (5.1) using different random seeds of the Ornstein-Uhlenbeck process, as shown in Figure (5.3).

(Stochastic) POD Time Coefficients

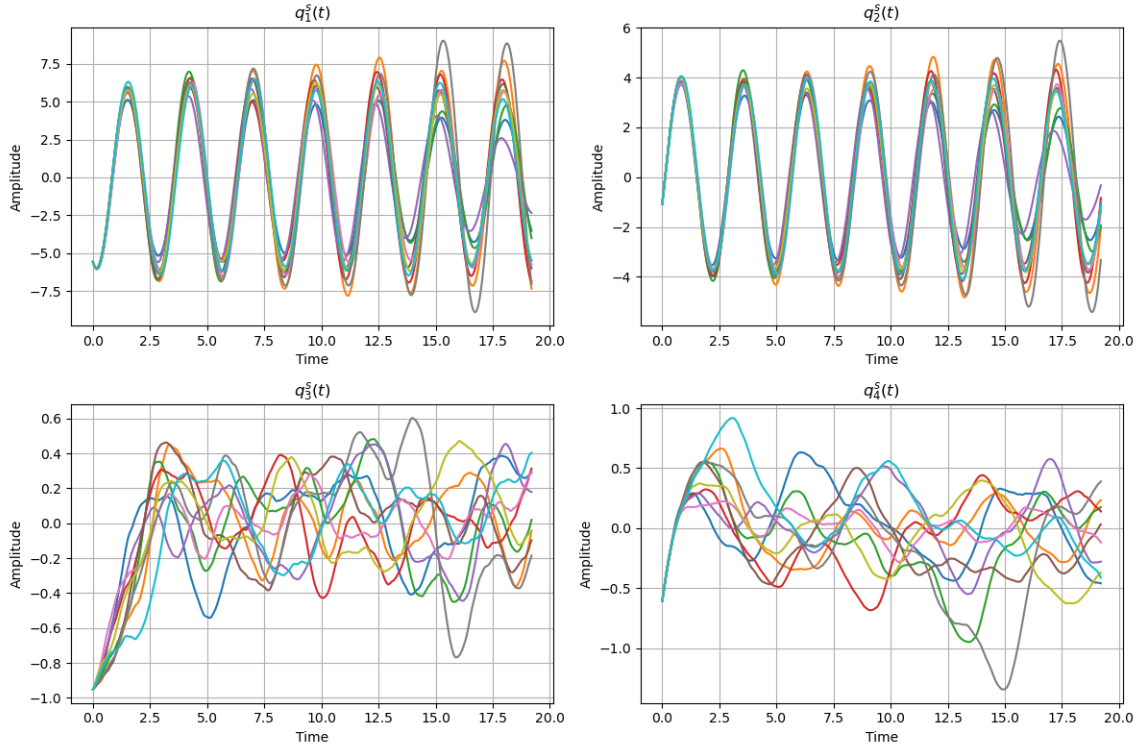


Figure D.1: Ten ensembles of POD time coefficients $q_\alpha^s(t)$ using Equation (5.8).

Every ensemble creates distinct solutions due to the randomness. Therefore, we obtain different POD time coefficients $q_\alpha^N(t, \omega)$ and reduced-order stochastic equations $a_\alpha^N(t, \omega)$ for each ensemble, where $\alpha = 1, \dots, 4$ is the corresponding mode, and $N = 1, \dots, 10$ is the ensemble. These results are shown in Figures (D.1) and (D.2). To validate the results, we compare each $q_\alpha^N(t, \omega)$ and $a_\alpha^N(t, \omega)$. The first four ensembles are depicted in the main part of the thesis, as shown in Figure (5.4). Our findings show that each ensemble from $N = 1$ to $N = 10$ show a good agreement between $q_\alpha^N(t, \omega)$ and $a_\alpha^N(t, \omega)$ for each $\alpha = 1, \dots, 4$.

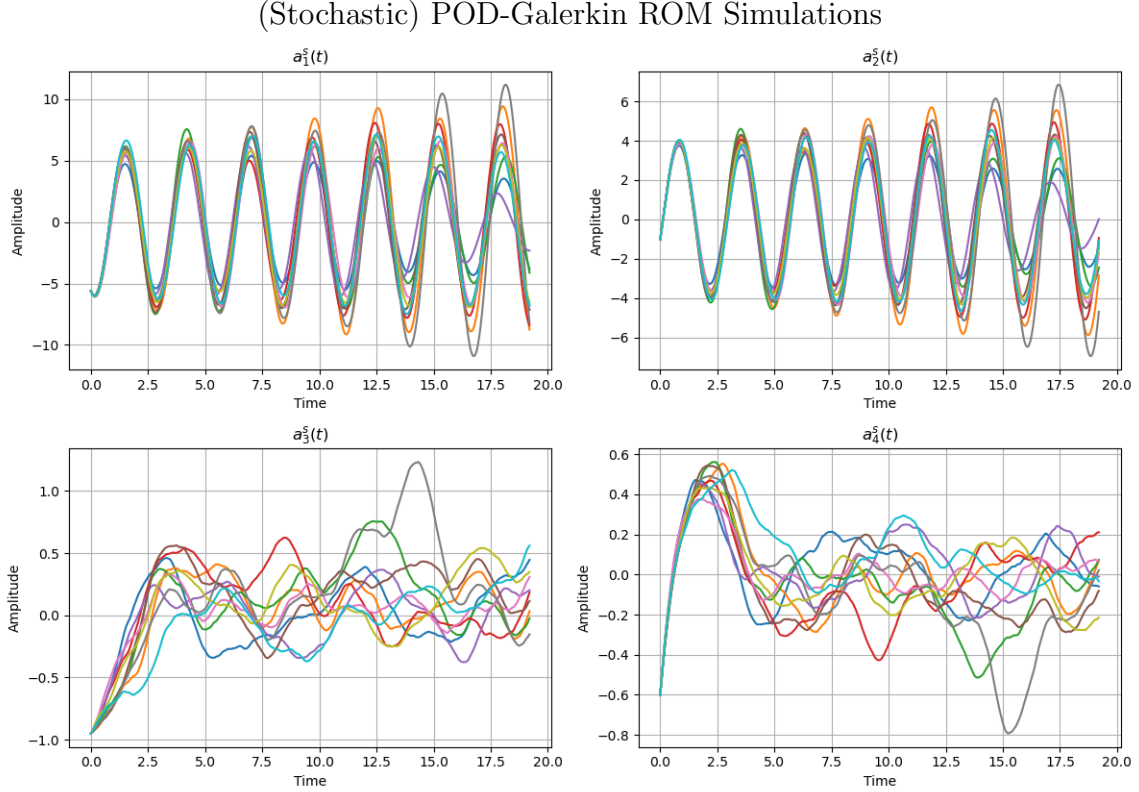


Figure D.2: Ten ensembles of POD-Galerkin reduced-order stochastic equations $a_\alpha^s(t)$ using Equations (5.9)-(5.10).

Figure (D.3) illustrates the full-order model of the first three ensembles together with POD reconstruction using $q_\alpha^s(t, \omega)$ and $\Psi_\alpha(x)$. For each ensemble, we achieve high accuracy and a small reconstruction error for SST $T(t, x, \omega)$. Figure (D.4) depicts the full-order model of the first three ensembles together with POD reconstruction using $a_\alpha^s(t, \omega)$ and $\Psi_\alpha(x)$. Similar to the POD case, we obtain high accuracy and a small reconstruction error for each ensemble of SST $T(t, x, \omega)$.

(Stochastic) SST and POD Reconstructions

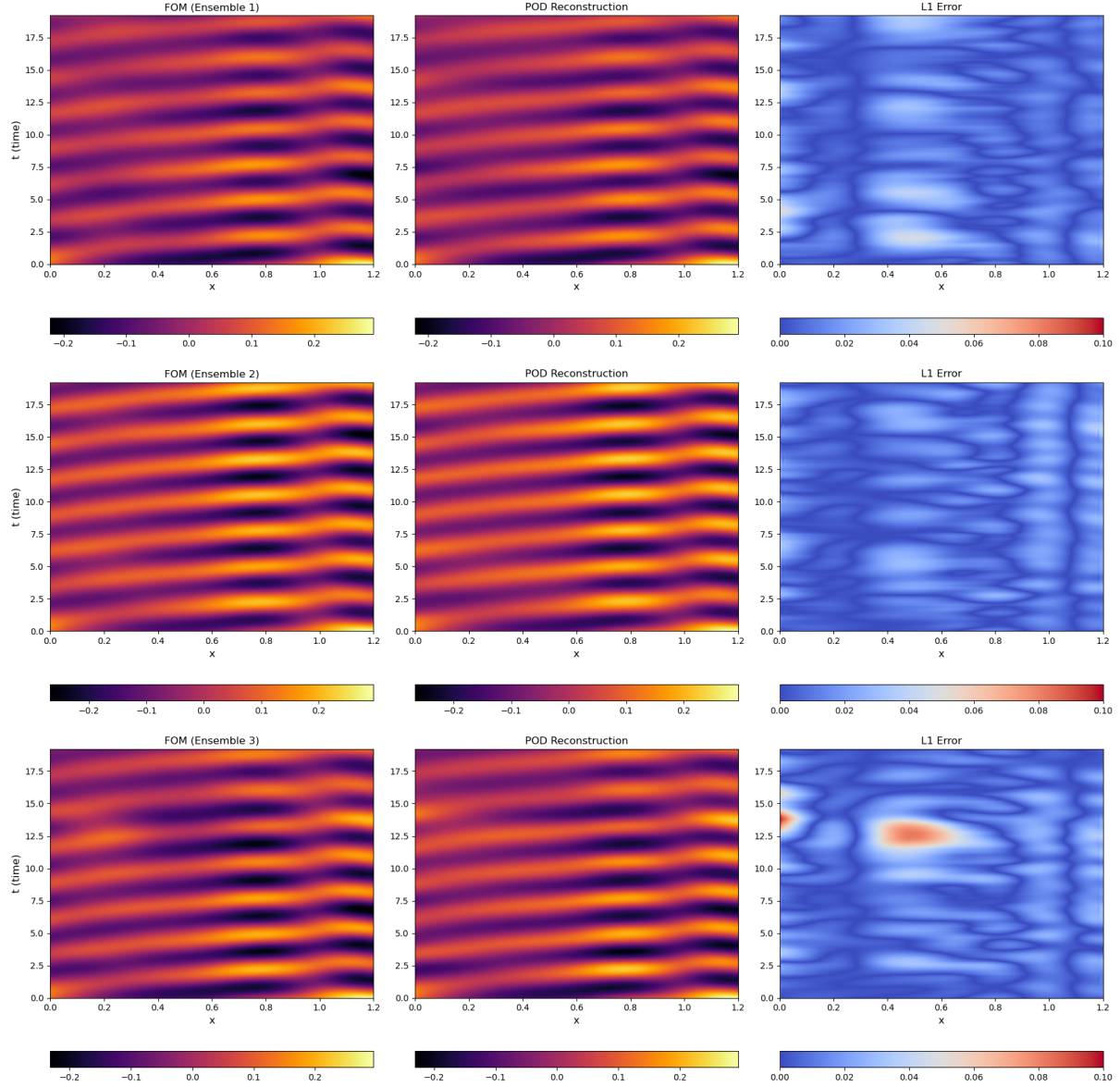


Figure D.3: First three ensembles of the SST component of full-order stochastic (5.1) (left), POD reconstruction of full-order model using $q_\alpha^s(t, \omega)$ and $\Psi_\alpha(x)$ (middle) and L1 error between them (right).

(Stochastic) SST and POD-Galerkin Reconstructions

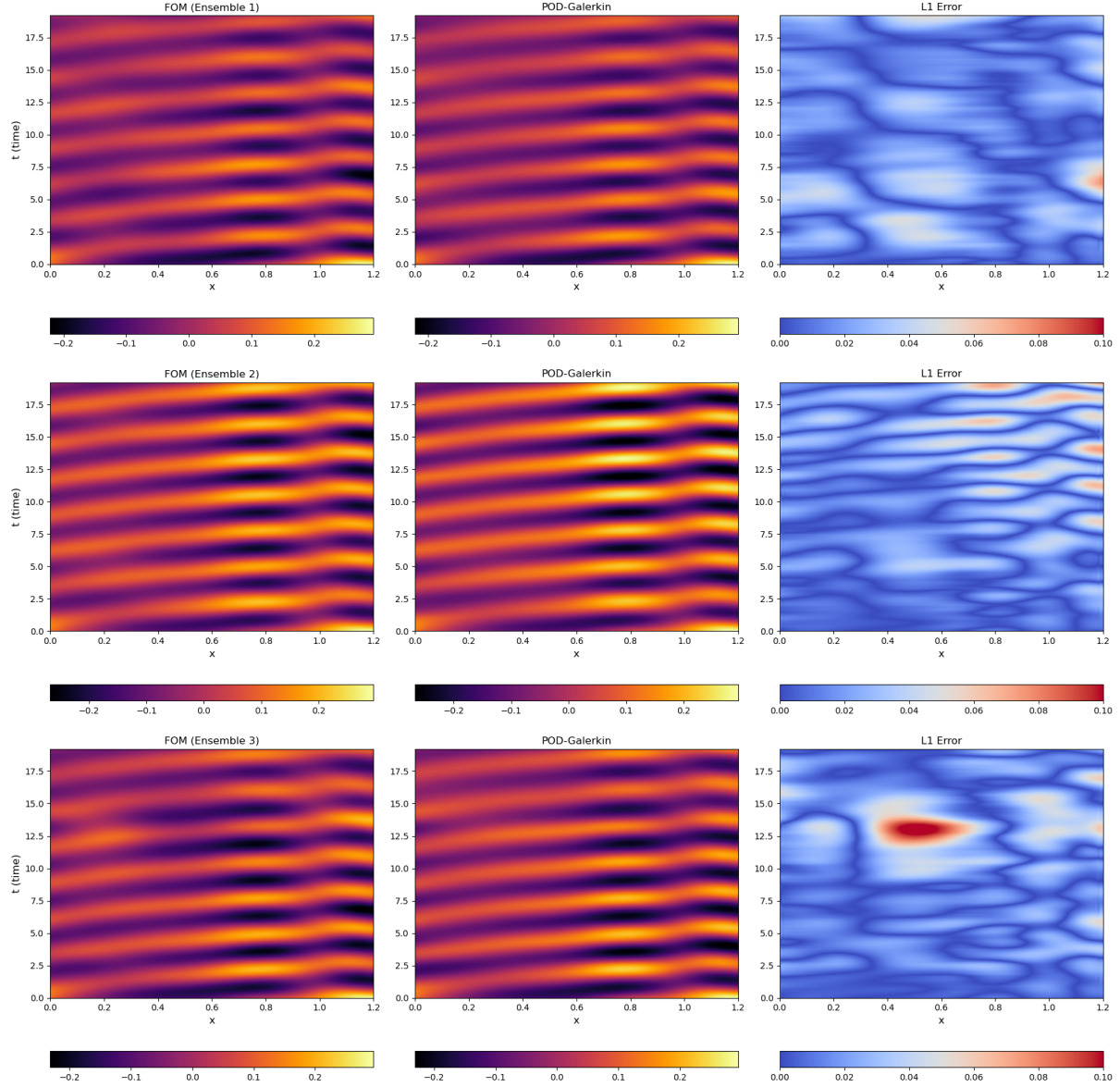


Figure D.4: First three ensembles of the SST component of full-order stochastic (5.1) (left), POD reconstruction of full-order model using $a_\alpha^s(t)$ and $\Psi_\alpha(x)$ (middle) and L1 error between them (right).

D.2 Particle Filter Algorithm

The particle filter algorithm described in Chapter (5) using the initial, prior and posterior samples is illustrated in Figure (5.9), and is summarized in the following table:

Step	Description
Step 1: Initialization	• Sample N_s particles $\{x_0^i\}_{i=1}^{N_s}$ from the prior $p(x_0)$ (Figure (5.9a)). • Set their weights $w_0^i = \frac{1}{N_s}$.
Step 2: Prediction	• For each particle i, propagate (Figure (5.9b)): $x_t^i \sim p(x_t \mid x_{t-1}^i).$
Step 3: Update	• Compute the weights based on the observation y_t (Figure (5.9c)): $w_t^i \propto w_{t-1}^i \cdot p(y_t \mid x_t^i).$ • Normalize the weights: $w_t^i \leftarrow \frac{w_t^i}{\sum_{j=1}^{N_s} w_t^j}.$
Step 4: Resampling	• Resample N_s particles $\{x_t^i\}_{i=1}^{N_s}$ based on their weights to obtain equally weighted particles.
Step 5: Repeat	• Repeat Steps 2–4 for $t = 1, 2, \dots, T$. • It should be noted that, in this chapter, N and N_s represent number of ensembles and particles, respectively.

Table D.1: Steps of the Particle Filter Algorithm.

D.3 A Spectral Approach to Nonlinear Filtering : Revisiting Wiener Chaos Expansion

In this section, we provide a spectral approach on nonlinear filtering by extending the knowledge on Wiener Chaos Expansion, as shown in Chapter (3). We consider the filtering scheme where the signal process $x(t)$ is a Markov diffusion process and the observation process is of the form

$$y(t) = y_0 + \int_0^t h(x(s))ds + w(t) \quad (\text{D.1})$$

where $w(t)$ is a Brownian motion independent of the process. And, $x(t)$ has the following form:

$$\begin{aligned} dx(t) &= b(x(t))dt + \sigma(x(t))dw(t), \quad 0 < t \leq T, \\ x(0) &= x_0 \end{aligned} \tag{D.2}$$

Let f be a given bounded function R^d and $\hat{f}(x(t))$ be the optimal filter. A fundamental result of filtering theory says that the optimal filter is given by the formula

$$\hat{f}(x(t)) = \frac{\int_{R^d} f(x)u(t, x)dx}{\int_{R^d} u(t, x)dx} \tag{D.3}$$

where $u(t, x)$ is the so-called unnormalized filtering density (UFD). A standard way to study the UFD is to treat it as a solution of the Zakai equation

$$du(t, x) = \mathcal{L}^*u(t, x)dt + h(x)u(t, x)dy(t) \tag{D.4}$$

where $\mathcal{L}^*$ is the formally adjoint operator to the generator of the Markov process $x(t)$, and is defined as (for 2-dimensional system):

$$\mathcal{L}^* = \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 \frac{\partial^2}{\partial x_i \partial x_j} [(\sigma \sigma^T)_{ij} u] - \sum_{i=1}^2 \frac{\partial}{\partial x_i} [\mu_i(x) u]$$

The solution of the Zakai equation (D.4) can be written in the form [94] :

$$u(t, x) = \sum_{\alpha} \frac{1}{\sqrt{\alpha!}} \varphi_{\alpha}(t, x) V_{\alpha}(y) \tag{D.5}$$

where $V_{\alpha}(y)$ are Wick polynomials of Wiener integrals $\int_0^t m_i(s)dy(s)$ where $\{m_k\}$ is a complete orthonormal system $L_2([0, t])$ and $\varphi_{\alpha}(t, x)$ are deterministic Hermite-Fourier coefficients in the Cameron-Martin orthogonal decomposition of $u(t, x)$ [17]. One important feature of the expansion is that it separates observations and parameters in that the Wick polynomials are completely defined by the observation process $y(t)$ but Hermite-Fourier coefficients $\varphi_{\alpha}(t, x)$ are determined only by the coefficients of the signal process $x(t)$, its initial distribution and the observation function h [94].

D.3.1 Recursive Equations

Plugging the expansion (D.5) into the Zakai equation (D.4) results in the following equation:

$$\frac{\partial \varphi_{\alpha}(s, x)}{\partial s} = \mathcal{L}^* \varphi_{\alpha}(s, x) + \sum_{k, l} \alpha_k^l m_k(s) h^l(x) \varphi_{\tilde{\alpha}(k, l)}(s, x) \quad 0 < s \leq t$$

where $\alpha = (\alpha_k^l)_{1 \leq l \leq 2; k \geq 1 \in \mathcal{J}}$, as defined in Chapter (3). $\varphi_0(s, x)$ satisfies the forward Kolmogorov equation (Fokker-Planck equation) corresponding to the state process:

$$\begin{aligned} \frac{\partial \varphi_0(s, x)}{\partial s} &= \mathcal{L}^* \varphi_0(s, x) \\ \varphi_0(0, x) &= p(x) \end{aligned} \quad (\text{D.6})$$

where $p(x)$ is the initial distribution. Fokker-Planck equation describes the time evolution of the probability density function (PDF) for a stochastic process. For $|\alpha| = 1$ with $\alpha_i = 1$, where we write φ_i for φ_α , the corresponding equation has the following form:

$$\begin{aligned} \frac{\partial \varphi_i(s, x)}{\partial s} &= \mathcal{L}^* \varphi_i(s, x) + m_i(s) h(x) \varphi_0(s, x) \\ \varphi_i(0, x) &= 0 \end{aligned} \quad (\text{D.7})$$

These equations (forced Fokker-Planck equations) have zero initial condition (distribution), and are forced by $\varphi_0(s, x)$ and $m_j(s)$. The orthonormal basis functions $m_j(s)$ are defined as:

$$m_1(s) = \frac{1}{\sqrt{T}}, \quad m_i(s) = \sqrt{\frac{2}{T}} \cos\left(\frac{\pi(i-1)s}{T}\right) \quad (\text{D.8})$$

where $2 \leq i \leq K$, $0 \leq s \leq T$ in which K depends on the order of truncation.

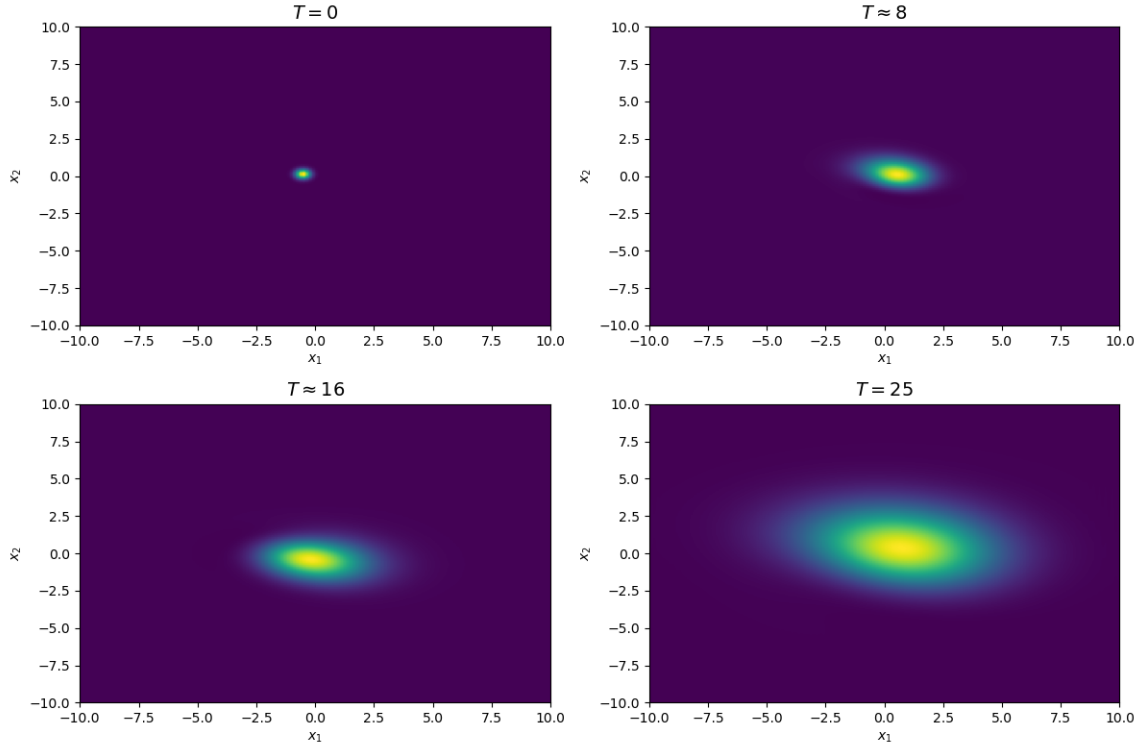


Figure D.5: Evolution of probability densities over time.

Figure (D.5) depicts the solution of Equation (D.6) by using the 2D state model obtained from the POD-Galerkin reduced order model forced by white noise. We start with an initial distribution

at $T = 0$. This probability density diffuses over time. It should be noted that this is an initial application of WCE to nonlinear filtering, and there are many points that need to be considered carefully.

Solving Fokker-Planck type equations numerically presents several challenges. In multi-variable systems, the equation becomes high-dimensional, leading to an exponential increase in computational cost—commonly known as the curse of dimensionality. For example, producing Figure (D.5) was produced by using an open software, and required 1GB storage, which is extremely expensive. When we include $|\alpha| = 1$ equations, the solution is infeasible, and requires high computational resources, which are beyond the scope of this research. Additionally, the equation often exhibits both fast and slow dynamics, which leads to stiffness. Explicit methods require very small time steps for stability, making them inefficient. In contrast, implicit methods, such as Alternating Direction Implicit (ADI) or Crank-Nicolson, offer stability for stiff problems and allow for larger time steps [106, 128, 149].