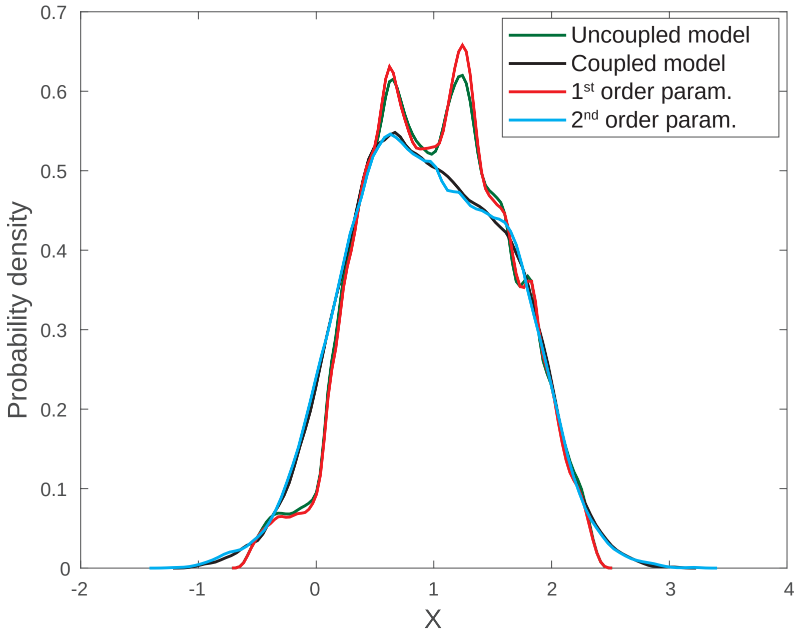
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Statistical Mechanical Methods for Parametrization in Geophysical Fluid Dynamics



Gabriele Vissio

Hamburg 2018

Berichte zur Erdsystemforschung
Reports on Earth System Science

212
2018

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

name.surname@mpimet.mpg.de
www.mpimet.mpg.de

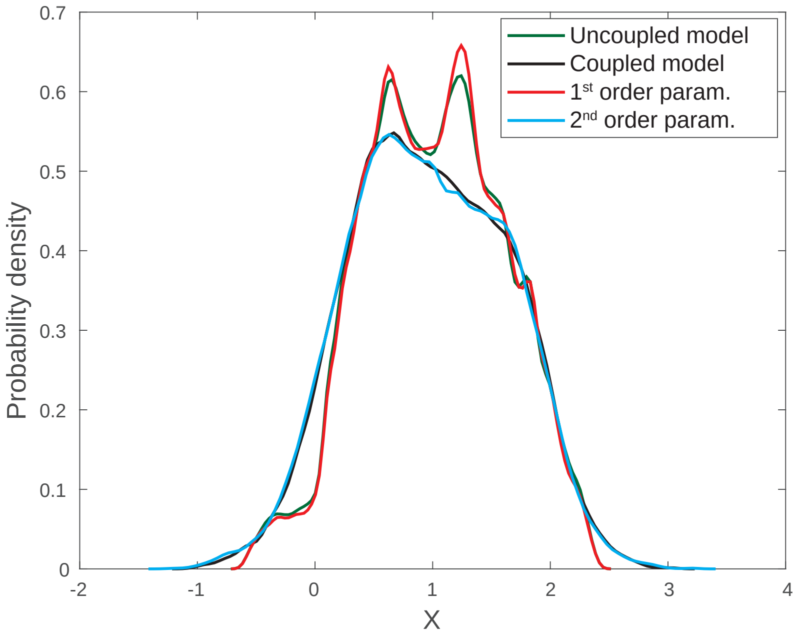
Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Photos below: ©MPI-M
Photos on the back from left to right:
Christian Klepp, Jochem Marotzke,
Christian Klepp, Clotilde Dubois,
Christian Klepp, Katsumasa Tanaka

Statistical Mechanical Methods for Parametrization in Geophysical Fluid Dynamics



Dissertation with the aim of achieving a doctoral degree at the
Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences of Universität Hamburg
submitted by

Gabriele Vissio

Hamburg 2018

Gabriele Vissio

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Universität Hamburg
Geowissenschaften
Meteorologisches Institut
Bundesstr. 55
20146 Hamburg

Tag der Disputation: 08.11.2018

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Valerio Lucarini
Prof. Dr. Gualtiero Badin

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 212
Reports on Earth System Science / Max Planck Institute for Meteorology 2018

ISSN 1614-1199

"Mathematicians seem to have no difficulty in creating new concepts faster than the old ones become well understood."

Edward N. Lorenz

UNIVERSITÄT HAMBURG

Abstract

Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences

Statistical Mechanical Methods for Parametrization in Geophysical Fluid Dynamics

by Gabriele VISSIO

Constructing accurate, flexible, and efficient parametrizations of sub-grid scale processes is a central area of interest in the numerical modelling of geophysical fluids. Here a recently introduced scale-adaptive approach constructed using statistical mechanical arguments and composed by deterministic, stochastic and non-markovian contributions is studied. It is applied first on a modified version of the two-level Lorenz 96 model and then on the Lorenz 84 model forced by the Lorenz 63 model.

The former application focuses on the flexibility of the model and its performance compared with a data driven approach. While the parametrization proposed is universal and can be easily analytically adapted to changes in the parameters' values by a simple rescaling procedure, the parametrization constructed with the ad-hoc approach used as benchmark needs to be recomputed each time the parameters of the systems are changed. The price of the higher flexibility of the method proposed here is having a lower accuracy in each individual case.

The latter application, lacking a memory term, focuses on the stochastic forcing provided by the Wouters-Lucarini methodology. It is shown that the approach allows for dealing seamlessly with the case of the Lorenz 63 being a fast as well as a slow forcing compared to the characteristic time scales of the Lorenz 84 model. The results are tested using both standard metrics based on the moments of the variables of interest as well as Wasserstein distance between the projected measure of the original system on the Lorenz 84 model variables and the measure of the parametrized one. By testing these methods on reduced phase spaces obtained by projection, it is shown that comparisons based on the Wasserstein distance might be of relevance in many applications despite the curse of dimensionality.

The results obtained are encouraging: the Wouters-Lucarini's parametrization demonstrates its reliability and flexibility with respect to the models used which, despite their extremely simplified representations of the atmosphere, allow to highlight the feasibility of the approach for nonlinear dynamics which presents similarities - in form of advection, forcing, dissipation, coupling terms - with more complex and realistic models.

German version:

Die Erstellung genauer, flexibler und effizienter Parametrisierungen von Sub-Gitter-Skalenprozessen ist ein zentraler Bereich der numerischen Modellierung geophysikalischer Fluide. Hier wird ein kürzlich eingeführter skalenadaptiver Ansatz untersucht, der unter Verwendung statistisch-mechanischer Argumente und aus deterministischen, stochastischen und nicht-markovischen Teilen zusammengesetzt wurde. Dieser Ansatz wird zuerst auf eine modifizierte Version des zweistufigen Lorenz 96-Modells und dann auf das Lorenz 84-Modell angewandt, das vom Lorenz 63-Modell angetrieben wurde ist.

Die erste Anwendung konzentriert sich auf die Flexibilität des Modells und seiner Leistung im Vergleich zu einem datengesteuerten Ansatz. Während die vorgeschlagene Parametrisierung universell ist und leicht durch analytische Verfahren an Änderungen der Parameterwerte durch eine einfache Neuskalierung angepasst werden kann, muss die mit dem als Benchmark verwendeten Adhoc-Ansatz konstruierte Parametrisierung jedes Mal neu berechnet werden, wenn die Parameter der Systeme geändert werden. Der Preis für die höhere Flexibilität des hier vorgeschlagenen Verfahrens ist in jedem Einzelfall eine geringere Genauigkeit.

Die zweite Anwendung, der einen Speicherterm fehlt, konzentriert sich auf den stochastischen Antrieb, welcher durch die Wouters-Lucarini-Methode bereitgestellt wird. Es wird gezeigt, dass der Ansatz ermöglicht, nahtlos mit dem Fall des Lorenz 63 umzugehen, der sowohl ein schneller als auch ein langsamer Antrieb ist, verglichen mit den charakteristischen Zeitskalen des Lorenz 84 Modells. Die Ergebnisse werden unter Verwendung sowohl von auf den Momenten der betreffenden Variablen basierten Standardmetriken als auch von der Wasserstein-Distanz zwischen dem projizierten Maß des Originalsystems auf den Lorenz 84-Modellvariablen und dem Maß der parametrisierten Variablen getestet. Durch Testen dieser Verfahren auf reduzierte Phasenräume, die durch Projektion erhalten werden, wird gezeigt, dass auf der Wasserstein-Distanz basierte Vergleiche in vielen Anwendungen trotz des Problems der Dimensionalität von Relevanz sein können.

Die Ergebnisse sind ermutigend: Die Parametrisierung von Wouters-Lucarini zeigt ihre Zuverlässigkeit und Flexibilität mit Bezug auf die verwendeten Modelle, die trotz ihrer extrem vereinfachten Darstellung der Atmosphär die Machbarkeit des Ansatzes für nichtlineare Dynamik hervorheben, der Ähnlichkeiten - in Form von Advektion, Antrieben, Dissipation, Kopplungstermen - mit komplexeren und realistischeren Modellen aufweist.

Acknowledgements

First of all, I am grateful to my supervisor Valerio Lucarini: he gave me the possibility to work in an extraordinarily interesting field and his support both on personal and on professional level was very important for me. This just partially includes our discussions on English language.

I am very grateful to Gualtiero Badin and Christian Franzke, respectively my internal supervisor and the co-advisor in my Panel Chair, always available for help, clarifications and discussions.

I wish to thank the Theoretical Meteorology group and, in particular, Sebastian Schubert, Vera Melinda Galfi, Valerio Lembo, Federica Gugole, Guannan Hu, Denny Gohlke and the former members Markus Kilian, Tamás Bódai and Maida Zahid.

On the Max Planck Institute side I wish to thank the HErZ/ARCS group and especially Cathy Hohenegger, Daniel Klocke and Mirjana Sakradzija. I am grateful to Juan-Pedro Mellado for the useful discussions about turbulence. Among the IMPRS-ESM students I would especially like to thank the whole 2015 and 2016 years - my antisocial behaviour did not prevent me to notice they are delightful people. Most of them, at least.

The journey through German bureaucracy would not have been possible without the help of the whole IMPRS-ESM office in MPI (Antje Weitz, Cornelia Kampmann, Michaela Born and the never forgotten Wiebke Boehm) and of Sabine Ehrenreich in University Geschaeftszimmer, I thank them for easing the burden.

I just want to steal some space in the technical acknowledgements to state that I am extremely grateful to my friends and family for supporting me in this task, especially to Sara Borella - and not just for the three figures in this thesis she helped me with.

Lastly, I would like to thank Steve Gunn and Sunil Patel for the template used for this thesis and Gabriel Peyré for making the Matlab software related to Wasserstein distance publicly available.

Contents

Abstract vii

Acknowledgements xi

1 Introduction 1

- 1.1 What is a parametrization? 1
- 1.2 A practical example 7
- 1.3 Mass flux parametrization 8
- 1.4 Parametrizations so far 10
- 1.5 Objectives of this thesis 13
- 1.6 Structure of the thesis 14

2 Mathematical Background 17

- 2.1 Wouters-Lucarini's Parametrization 17
  - 2.1.1 The method 18
    - Deterministic, stochastic, and non-markovian terms 21
    - Independent coupling case 22
    - Forcing in the fast dynamics 23
- 2.2 Lorenz Models 24
  - 2.2.1 Lorenz 96 24
    - Modifications to the model 27
    - Wilks Parametrization 30
  - 2.2.2 Lorenz 84 32
  - 2.2.3 Lorenz 63 33
  - 2.2.4 Lorenz 84 forced by Lorenz 63 35
- 2.3 Wasserstein Distance 37

3 Results 41

- 3.1 Wouters-Lucarini's parametrization on Lorenz 96 41
  - 3.1.1 Constructing the Parametrization 41
  - 3.1.2 Performance on Lorenz 96 44
    - Sensitivity to the Strength of the Coupling 47

Scale adaptivity 49
Forcing the fast scale dynamics 51
3.2 Wouters-Lucarini's parametrization on Lorenz 84 forced by Lorenz 63 53
3.2.1 Constructing the Parametrization 53
3.2.2 Performance on Lorenz 84 forced by Lorenz 63 55
Parametrizing the fast system 55
Parametrizing the slow system 61

4 Conclusions **69**

Bibliography **73**

List of Figures

1.1 Schematic view of the components of the climate system, their processes and interactions (Solomon et al., 2007). 2
1.2 Global mean energy budget under present-day climate conditions. Numbers state magnitudes of the individual energy fluxes in $\frac{W}{m^2}$, adjusted within their uncertainty ranges to close the energy budgets. Numbers in parentheses attached to the energy fluxes cover the range of values in line with observational constraints (Stocker et al., 2013, see also for the sources of the data). 3
1.3 Idealized atmospheric cells around a rotating Earth (Lutgens and Tarbuck, 2001). 4
1.4 Typical space and time scales of atmospheric phenomena (Cullen and Brown, 2009). 5

2.1 A schematic representation of Lorenz 96 model in the case $K = 8$ and $J = 8$, with the inner circle composed by the slow X variables while the outer circle represents the fast J variables (adapted from Wilks, 2005). 25
2.2 Typical longitudinal profiles of X_k and $Y_{j,k}$. 26
2.3 A schematic representation of the modified Lorenz 96 model used in this thesis (adapted from Wilks, 2005). Note the presence of the external forcing F_2 and the division of the Y variables in non interacting subsections. 28
2.4 Probability density of the X variable for the original (black line) and the modified (red line) Lorenz 96 model. 29
2.5 a) Time autocorrelation; b) Spatial autocorrelation of the X variable for the original (black line) and the modified (red line) Lorenz 96 model. 30
2.6 Example of scatterplot with related fit in the case of Lorenz 96 model. 31

2.7 Schematic representation of the Lorenz 84 model. The blue arrow represents the westerlies, therefore the X variable, while the orange arrow represents the poleward heat transport and is therefore related to the Y and the Z variables. 32
2.8 Poincaré section in $Z = 0$ of Lorenz 84 model. 33
2.9 Convective roll in Rayleigh Bénard convection model. The flux of rising fluid in the right convective plume is balanced by a downward flux of fluid in the descending plume on the left. . . 34
2.10 Lorenz attractor drawn by Lorenz 63 model. 35
2.11 Poincaré section in $Z = 0$ of Lorenz 84 model coupled with Lorenz 63 model. 37

3.1 Time lagged autocovariance of the noise term $\sigma(t)$ with $b = 10$ and $h = 1$. 43
3.2 Memory effects as measured by the factor H, see Eq. (3.7), with $b = 10$ and $h = 1$. 44
3.3 Probability density of the X variable for the different models considered. 45
3.4 Temporal autocorrelation of the X variable for the different models considered. 46
3.5 Spatial autocorrelation of the X variable for the different models considered. 46
3.6 a) First moment; b) Second centered moment; c) Third centered moment; d) Fourth centered moment as functions of the coupling strength for the different models considered. 48
3.7 a) Probability density of the X variable in the case of $c = 1$, $b = 10$ and $h = 1$. b) Zoom on the peak of the distribution. . . . 49
3.8 a) Probability density of the X variable in the case of $c = 100$, $b = 10$ and $h = 0.1$. b) Zoom on the peak of the distribution. . . 50
3.9 Probability density of the X variable in the case of $c = 5$, $b = 8$ and $h = 1.1$. 51
3.10 Probability density of the X variable calculated adding G to uncoupled Y equation. The standard case is the one shown in Fig.3.3. 52
3.11 Time autocorrelation of the X variable calculated adding G to uncoupled Y equation. The standard case is the one shown in Fig.3.4. 52
3.12 Poincaré section in $Z = 0$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization. 55

3.13 Poincaré section in $X = 1$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization. 56
3.14 $3D$ view of the attractor of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization. . . . 57
3.15 Probability density of the X variable. 59
3.16 Probability density of the Y variable. 59
3.17 Probability density of the Z variable. 60
3.18 Wasserstein distances from the coupled model with respect to the number of cubes per side: a) $3D$ case; b) Projection on XY plane; c) Projection on XZ plane; d) Projection on YZ plane. . . 62
3.19 Poincaré section in $Z = 0$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization. 63
3.20 Poincaré section in $X = 1$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization. 64
3.21 $3D$ view of the attractor of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization. . . . 65
3.22 Probability density of the X variable. 66
3.23 Probability density of the Y variable. 66
3.24 Probability density of the Z variable. 67
3.25 Wasserstein distances from the coupled model with respect to the number of cubes per side: a) $3D$ case; b) Projection on XY plane; c) Projection on XZ plane; d) Projection on YZ plane. . . 68

List of Tables

3.1 Values of the constants determining the parametrization according to the Wilks method for various values of the model's parameter c. 49
3.2 Expectation values for the ensemble average of the first two moments of the variables X, Y, and Z. The uncertainty is indicated as standard deviation (std) σ over the ensemble of realizations. All the values are multiplied by 10^2. 61
3.3 Expectation values for the ensemble average of the first two moments of the variables X, Y, and Z. The uncertainty is indicated as standard deviation (std) σ over the ensemble of realizations. All the values are multiplied by 10^2. 64

To Sara

Chapter 1

Introduction

1.1 What is a parametrization?

The climate is a chaotic system featuring both external forcings and internal dissipation. Due to the extreme variety of its different subsystems - *atmosphere, hydrosphere, cryosphere, lithosphere and biosphere* -, each composed by countless parts, and the complexity of the phenomena describing the interactions within and among these components, the range of the space and time scales covered by it stretches over several orders of magnitude (Peixoto and Oort, 1992; Hartmann, 1994; Holton, 2004; McGuffie and Henderson-Sellers, 2005; Palmer and Hagedorn, 2006).

The interplay among the components is shown in Fig.1.1. Everything starts from the Sun which, emitting short wave radiations, acts as an external forcing for the Earth system. This radiation, weakly absorbed by the atmosphere and eventually reflected by clouds, is then absorbed or reflected by the surface, which wide diversity of composition - soil, vegetation, water, ice, urban areas, etc. - gives rise to an intricate set of exchanges of energy, momentum and chemical components among biosphere, hydrosphere, cryosphere and atmosphere. The latter is activated by the re-emission in long waves radiation. The motion of the air triggers new phenomena, like the generation of clouds - which greatly affects the energy balance and produce precipitations - and the wind stress. The complex interrelations among all these subcomponents rapidly develop with nonlinear dynamics, preventing the climate system to be predictable after a relatively low time span.

Fig.1.2 focuses the attention on the energy budget of the atmosphere. The above mentioned solar radiation is partially reflected in the space by clouds and absorbed by the atmosphere. The energy that reaches the ground is partly reflected - mainly by bright cover like ice -; the rest is absorbed by the ground and re-emitted in long waves. The atmosphere is not transparent to this kind

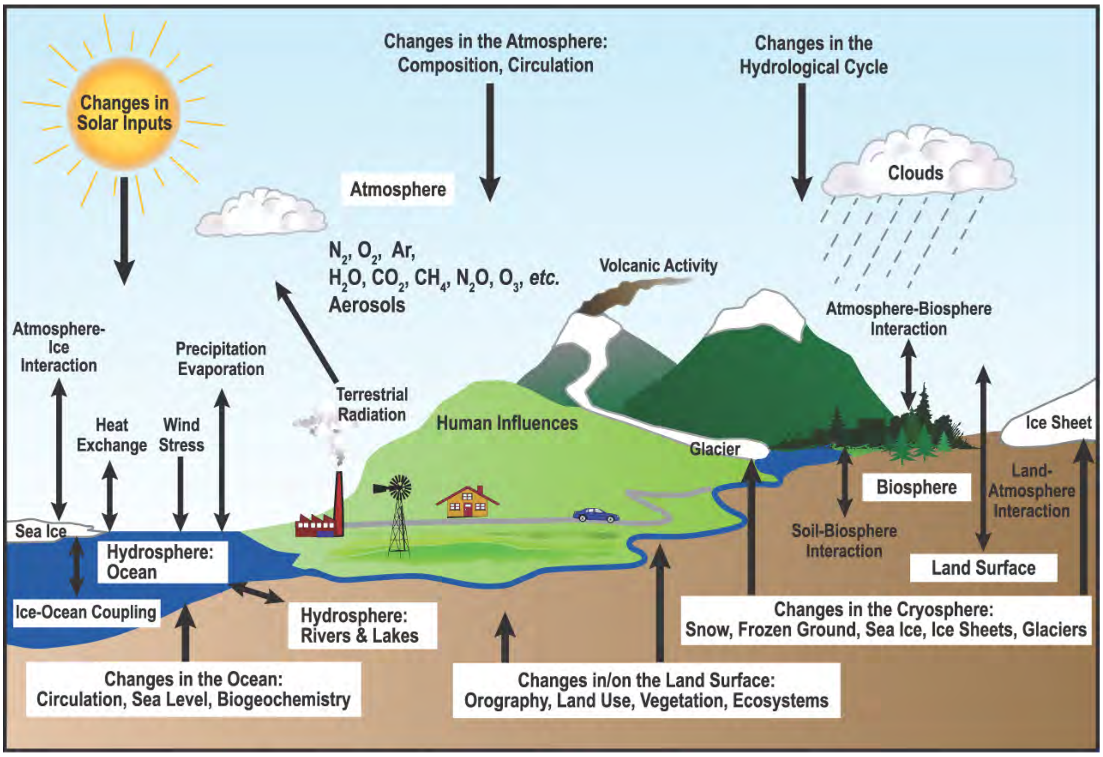

Changes in Solar Inputs
Changes in the Atmosphere: Composition, Circulation
Changes in the Hydrological Cycle
Atmosphere
Clouds
N_2, O_2, Ar, H_2O, CO_2, CH_4, N_2O, O_3, etc.
Aerosols
Volcanic Activity
Atmosphere-Biosphere Interaction
Atmosphere-Ice Interaction
Precipitation Evaporation
Terrestrial Radiation
Human Influences
Heat Exchange
Wind Stress
Glacier
Ice Sheet
Biosphere
Land-Atmosphere Interaction
Sea Ice
Hydrosphere: Ocean
Soil-Biosphere Interaction
Land Surface
Ice-Ocean Coupling
Hydrosphere: Rivers & Lakes
Changes in the Cryosphere: Snow, Frozen Ground, Sea Ice, Ice Sheets, Glaciers
Changes in the Ocean: Circulation, Sea Level, Biogeochemistry
Changes in/on the Land Surface: Orography, Land Use, Vegetation, Ecosystems

FIGURE 1.1: Schematic view of the components of the climate system, their processes and interactions (Solomon et al., 2007).

of radiation and absorbs it re-emitting energy both towards the ground and the space. The energy received by the ground is used also for the release of sensible heat and evaporation, two important engines for the atmosphere dynamics.

The atmosphere is the most important component of the climate with respect to the aims of this thesis. The majority of weather and climate phenomena affects the first 10^4 meters in the vertical direction - the troposphere - while the scale in the horizontal direction is 10^7 meters. This thin gas layer is forced and set into motion mainly by the energy received by the Sun, which differential heating is due to a plethora of surface conditions - e.g. water or soil, mountains or planes, bare soil or vegetation, etc. - and thermodynamical inhomogeneities due to different components with diversified physical and chemical properties.

As a general rule, the tropical area receives more energy than the rest of the Earth, thus creating a net flow of rising air, heated by the surface below; on the other hand, the poles on average emit more energy than the one absorbed and the air above them tends to descend. The differential heating causes a net flux of energy from the hottest regions - the tropics - to the coldest - the poles - carried on by the ocean and the atmosphere. At the planetary scale even the Earth rotation plays an important role, since in this non-inertial system the motion of

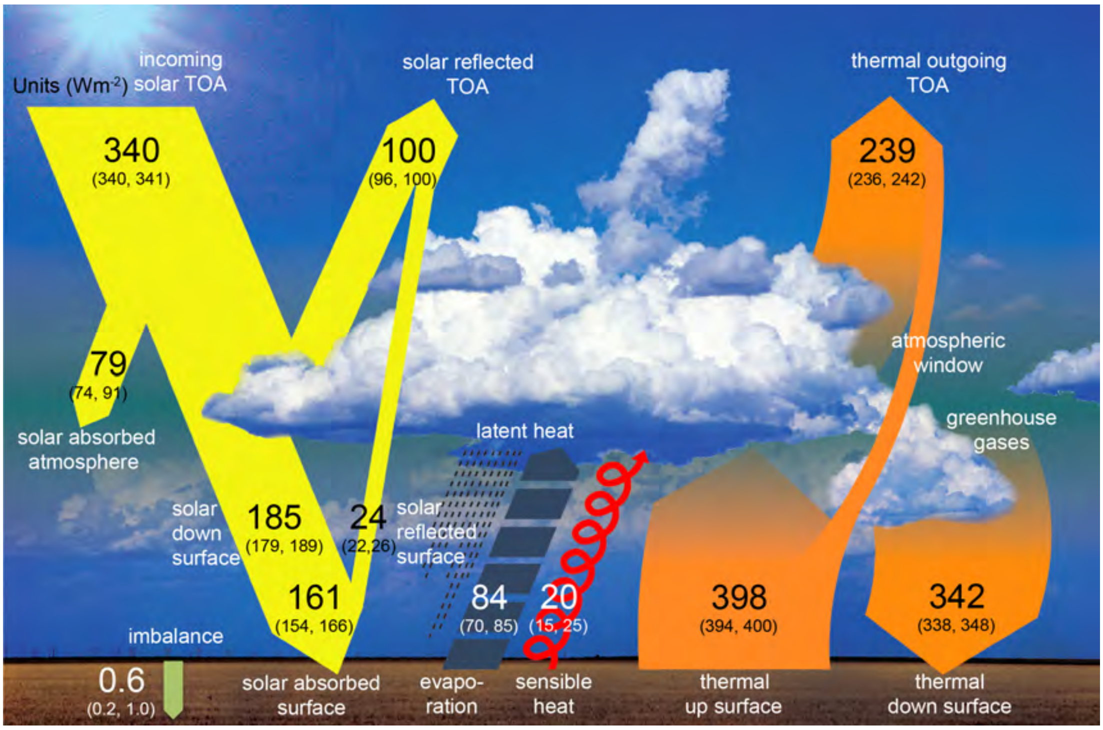

Units (Wm-2)
incoming solar TOA
340
(340, 341)
solar reflected TOA
100
(96, 100)
thermal outgoing TOA
239
(236, 242)
79
(74, 91)
solar absorbed atmosphere
atmospheric window
greenhouse gases
latent heat
solar down surface
185
(179, 189)
24
(22,26)
solar reflected surface
161
(154, 166)
84
(70, 85)
20
(15, 25)
398
(394, 400)
342
(338, 348)
imbalance
0.6
(0.2, 1.0)
solar absorbed surface
evapo-ration
sensible heat
thermal up surface
thermal down surface

FIGURE 1.2: Global mean energy budget under present-day climate conditions. Numbers state magnitudes of the individual energy fluxes in $\frac{W}{m^2}$, adjusted within their uncertainty ranges to close the energy budgets. Numbers in parentheses attached to the energy fluxes cover the range of values in line with observational constraints (Stocker et al., 2013, see also for the sources of the data).

a point on the surface is affected by an apparent force, called the Coriolis force. The subsequent acceleration breaks the general circulation of the atmosphere, which would otherwise be composed by just one cell of rising tropical air and descending polar air, in three distinct cells for each hemisphere - the equatorial *Hadley cell*, the midlatitudinal *Ferrel cell* and the *polar cell* (Fig.1.3). The conservation of momentum prevents these cells to be simply oriented in a meridional direction, giving to the wind stress a zonal component and creating the easterlies in the tropical belt and in the polar area and the westerlies in midlatitudes. Between the cells, at $8 - 15$ km of altitude, the temperature gradient sets up an extremely fast flow of air called *jet stream*. At synoptic scales ($\sim 10^6$ meters), small disturbances induced into the jet by variations in the meridional temperature gradient tend to amplify, developing synoptic scale systems of high and low pressure - this phenomenon is called *baroclinic instability*.

The Earth rotation does not take on such an important role when the focus shifts on smaller scales as the ones interested by convection (the horizontal

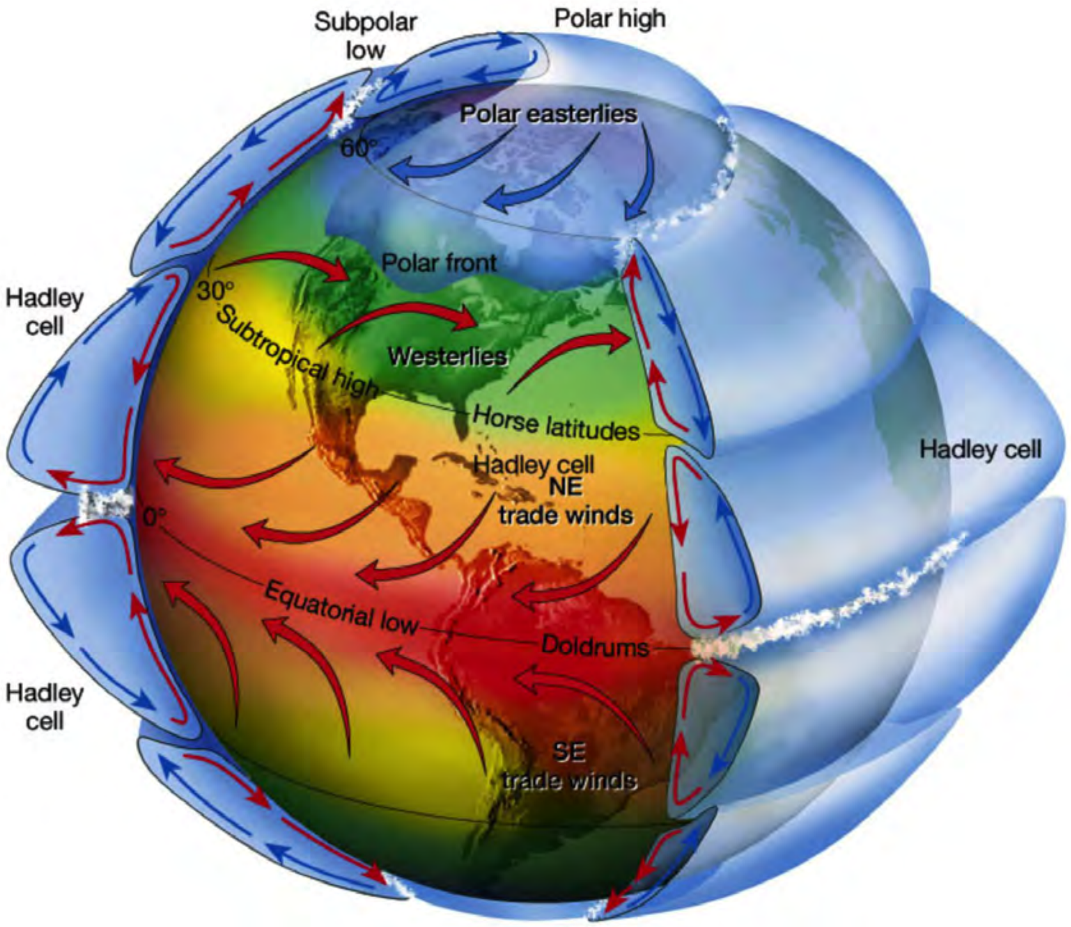

Subpolar low
Polar high
Polar easterlies
60°
Polar front
30°
Subtropical high
Westerlies
Horse latitudes
Hadley cell
NE trade winds
Hadley cell
0°
Equatorial low
Doldrums
SE trade winds
Hadley cell
Hadley cell

FIGURE 1.3: Idealized atmospheric cells around a rotating Earth (Lutgens and Tarbuck, 2001).

scale for deep convection, for example, is $\sim 10^3$ meters), since the deflection due to Coriolis force is negligible for dynamics on a short time scale. Indeed, the convection is a local physical phenomenon where masses of air lift for various reasons - heated by the ground or mechanically forced by mountains or other air masses with different thermodynamical properties - and leads to the formation of clouds and, eventually, precipitation. The response time - i.e. the time scale needed for a system to re-equilibrate to a new state after the application of a small perturbation - changes from weeks for the large scale circulation to hours for the convection.

The energy received from the Sun flows from the larger - planetary, synoptic, convective - to smaller scales. On very small scales (<10 meters) it is possible to observe the onset of a purely dissipative phenomenon, the *turbulence* (Pope, 2002; Kundu and Cohen, 2010). The *energy cascade* - i.e. the transfer of energy from larger to smaller scales of motion - is dissipated on the molecular level by the viscosity, that is, the resistance opposed to the flow by a fluid due

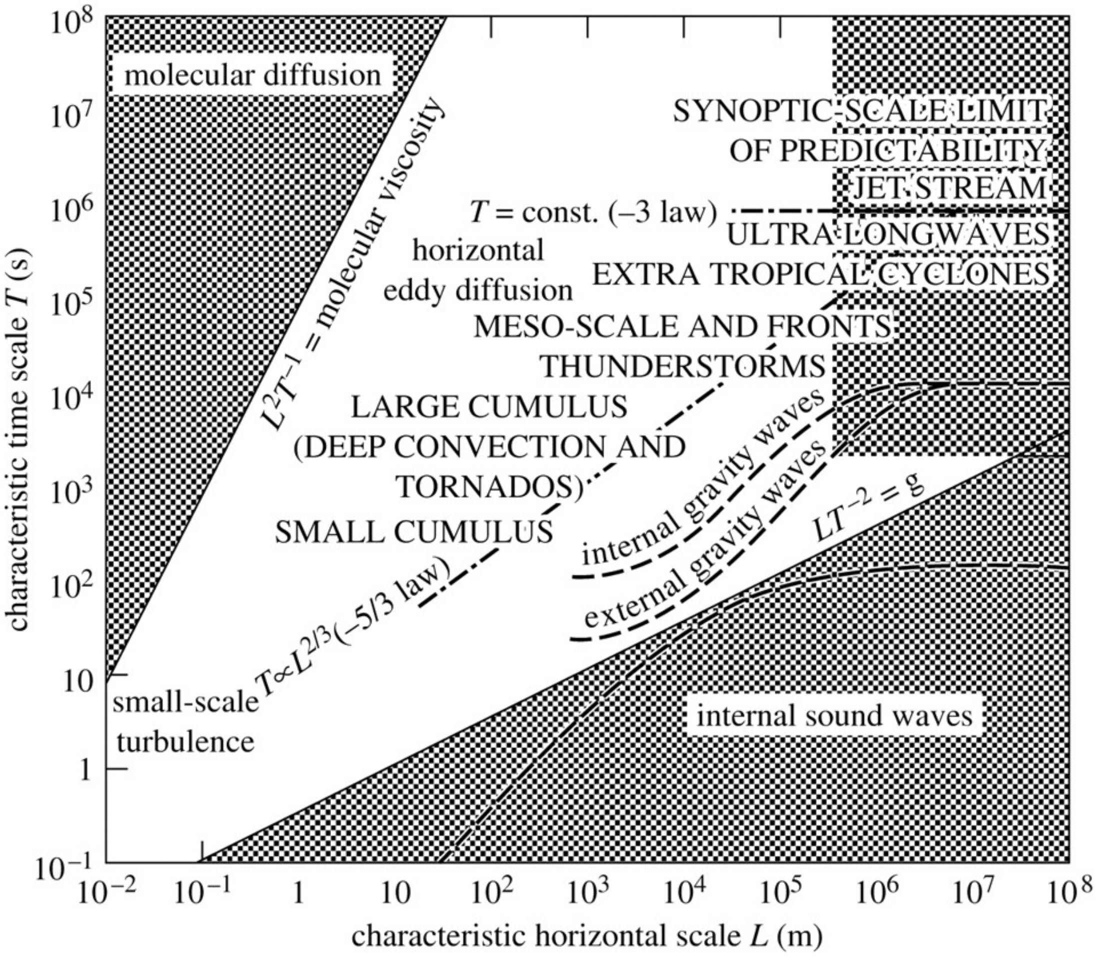


FIGURE 1.4: Typical space and time scales of atmospheric phenomena (Cullen and Brown, 2009).

to internal friction. Turbulent flows must be supplied by a continuous transfer of energy to compensate for the loss due to viscosity. In the atmosphere the parcels of air, mechanically or thermally forced, dissipate the energy supplied by large scales to move with respect to other air parcels, the surface or any inhomogeneity. This generates fast and short-lived turbulent eddies which are typically not solved - thus resulting unpredictable - for climate and weather models but are nevertheless very important in the energy budget of the atmosphere - and the entire climate, since this phenomenon occurs even in the hydrosphere.

A complete and thorough simulation would therefore require a vast amount of equations - e.g. for the computation of three dimensional moment, humidity, specific entropy - to be simultaneously computed on an extremely fine grid. Lacking infinite computational power, the standard solution to this conundrum is to run models in a reduced phase space representing only variables of interest, physical quantities which have particular importance related to the problem to solve. Furthermore, these processes must be simulated on a course grained grid, that is a grid where the points in which the calculations are performed are so far away from each other that a whole set of physical phenomena, which occur on a smaller spatial scale, is left unresolved. For example, a weather prediction model would require hourly precipitation on a space scale of a few kilometers, while climate models necessitate of daily, or even monthly, average amounts of precipitation. These models run on a particular grid, which scale imposes the minimum scale of the phenomena solved by the equations. Fig.1.4 shows the scales at which atmospheric phenomena occurs. A model set to solve cyclonic activity is likely built considering convection as an unresolved process, while a weather prediction model, which is concocted to solve thunderstorms - and therefore convection -, does not solve every small scale turbulent eddy, which effects on convection itself must be represented somehow. Therefore, even though the unresolved processes to approximate are different with respect to the used model, the problem of parametrization is a conundrum that must be faced on any application, thus a sistematic methodology is required.

Everything occurring on a smaller space scale than the ones resolved is called *subgrid phenomenon*; due to the chaotic feature of the climate system, this kind of phenomena strongly impacts on the variables of interest through the transport of energy, momentum or various substances (like water vapour) and, therefore, must be somewhat represented. The methodologies employed to deal with this problem, providing a mathematical representation to reproduce the influence of the unresolved scales on the resolved ones, are collectively called *parametrizations* (Peixoto and Oort, 1992). Furthermore, the development of parametrizations allows to gain insight related to the complex dynamics interacting among multiscale systems like the atmosphere-ocean system (Uboldi and Trevisan, 2015; Vannitsem and Lucarini, 2016), where the parametrization procedure can be applied to represent the effect of oceanic slow time scale dynamics on atmospheric fast time scales.

Since the problem is basically that the number of equations to be solved is smaller than the number of variables, the unresolved physical quantities must be substituted by an expression relating them to computed variables; this is called *closure approximation* or *parametrization*.

Parametrizations are usually divided into two categories:

- Bottom-up, or data-driven: this kind of parametrization is calculated through samples of time series of the unresolved system, the resolved system or observations.

- Top-down: a type of parametrization which necessitates the mathematical expressions for the evolution of the variables of interest and the unresolved variables.

A somewhat peculiar point of view is provided by the so-called super-parametrizations, mostly used to represent convection (Majda, 2007; Li et al., 2012). The idea is to have low dimensional (and so computationally much cheaper) models run in parallel with the main code for resolving at high resolution the dynamics inside each atmospheric column.

1.2 A practical example

The aim of this section is to show a typical case where it is common to apply a parametrization, the atmospheric convection. Indeed, the Navier-Stokes equation, which describes the evolution of the fluid momentum, must be numerically integrated on a grid that usually does not allow to solve small scale phenomena like turbulence, depicted by the eddy transport in the Reynolds Averaged Navier Stokes equations (hereafter RANS) (Pope, 2002; Holton, 2004; Plant and Yano, 2016).

The momentum equation for an incompressible fluid, which describes the velocity variation of a parcel of fluid, can be written as

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\nabla \cdot \mathbf{u} = -\frac{1}{\rho}\nabla p + \nu \nabla^2 \mathbf{u}, \tag{1.1}$$

where $\mathbf{u} = (u, v, w)$ is the velocity of the parcel, ρ its density, p the pressure acting on it and ν the kinematic viscosity. Eq.(1.1), written in the case of absence of external forcings such as gravity, is called *Navier-Stokes equation* (Chandrasekhar, 1961).

When this equation has to be numerically simulated, a widely used method is to split the terms in two components:

$$\mathbf{u} = \overline{\mathbf{u}} + \mathbf{u}'. \tag{1.2}$$

The first component $\overline{\mathbf{u}}$ is the large scale, slowly changing, velocity field and is the variable which is going to be computed in the numerical model, while the

second component $\mathbf{u}'$, i.e. the small scale, fast turbulent component, is one of the unresolved variables. This latter variable is, by definition, a deviation from the mean field and is therefore null when averaged in time.
The same decomposition can be applied also to the pressure field, yielding to $\mathbf{p} = \overline{\mathbf{p}} + \mathbf{p}'$.

Applying this decomposition, employing the continuity equation

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{1.3}$$

and averaging in time leads to

$$\frac{\partial \overline{\mathbf{u}}}{\partial t} + \overline{\mathbf{u}}\nabla \cdot \overline{\mathbf{u}} = -\frac{1}{\rho}\nabla \overline{p} + \nu \nabla^2 \overline{\mathbf{u}} - \sum_{i,j} \frac{\partial \overline{u'_i u'_j}}{\partial x_i} \tag{1.4}$$

Note that this procedure cancels out several mixed terms since the product of a deviation with a mean is null when time averaged:

$$\overline{uv} = \overline{(\overline{u} + u')(\overline{v} + v')} = \overline{u}\overline{v} + \overline{u'v'}. \tag{1.5}$$

The term $\sum_{i,j} \frac{\partial \overline{u'_i u'_j}}{\partial x_i}$ in Eq.(1.4) is called eddy transport and represents the correction due to unresolved subgrid variables.

The parametrization of eddy transport is a particularly difficult challenge which has long deserved scientists effort (Pope, 2002; Holton, 2004; Arakawa, 2004; Plant and Yano, 2016). Since the equation for the evolution of fluctuation fields is rather complicated (see for example Garratt, 1992), not only providing a top-down parametrization results an extremely difficult task, but also a surrogate time series is often unavailable. This problem, along with the fact that the eddy transport is such a volatile quantity that it is practically impossible to measure, makes also data-driven approaches challenging.

1.3 Mass flux parametrization

One of the main attempts to solve the convection parametrization problem was proposed by Riehl and Malkus, 1958 and is called *mass flux parametrization* (Plant and Yano, 2016). This hypothesis is based on the idea that convection happens in quasi-isolated areas where a large mass of air rises forming updrafts without almost any interaction with the sorrounding environment. This upward mass flux, a term that indicates the vertical transport rate of air, can be written in broad terms as

$$M = \rho \sigma w, \tag{1.6}$$

with ρ, σ and w defining, respectively, the density of the air, the fractional area of the updraft and the vertical velocity.

Applying the Reynolds Averaging (see Section 1.2) to the budget equations for energy and moisture, Arakawa, 1969 expressed the problem in terms of finding the vertical fluxes of sensible heat and moisture, $\rho\overline{w's'}$ and $\rho\overline{w'q'}$. Expressing these quantities relatively to the convective updrafts and taking into account the relative magnitude between updrafts and downdrafts (see Smith, 1997 for a complete dissertation about this topic), he discovered that the relation between the vertical flux of a quantity ϕ and the mass flux can be written as

$$(\rho\overline{w'\phi'})_i \approx M_i(\phi'_i - \overline{\phi}), \tag{1.7}$$

where the subscript i indicates the subensebles.

Thus, the problem of parametrization shifts to finding a closure that allows to express satisfactorily the mass flux, from which it is straightforward to find heat and moisture fluxes.

Several attempts have been tried to formulate an expression for the mass flux at the cloud base M_{cb}, a quantity necessary to calculate the latent heat release to be used in the microphysical model to determine the physical phenomena occurring in the clouds. A typical approach is the CAPE closure (Fritsch and Chappell, 1980), named after the Convective Available Potential Energy - the amount of energy within an air column which can be transformed in kinetic energy:

$$M_{cb} \sim \frac{CAPE}{\tau}, \tag{1.8}$$

where τ is the time scale needed to reduce the CAPE to zero. The flaw of this approach is that assumes the convection in quasi-equilibrium with the large scale flow, in contrast to what happens in nature.
Another example is the so-called boundary layer closure (Mapes, 2000; Fletcher and Bretherton, 2010), where the leading parameter is CIN - Convective INhibition, the energy a parcel requires in order to reach the Level of Free Convection:

$$M_{cb} \sim w e^{-\frac{CIN}{w^2}}, \tag{1.9}$$

where w is the typical vertical velocity scale and, therefore, w^2 is proportional to the total kynetic energy. This model is good for representing shallow convection, but the CIN value is usually very low and this makes difficult to set a threshold value for the onset of convection.

These two methods, among with several others, provide representations of reality with severe physical flaws and/or rather limited applications. Therefore, it is understandable how a great deal of resources has been spent trying to solve the problem.

The search for an effective parametrization is one of the main intents of the ARCS (Advancing the Representation of Convection across Scales) a project which works within the Hans Ertel Centre for Weather Research to deepen the comprehension of convection phenomenon.

This thesis, which belongs to the ARCS project, aims to give a shifted perspective to the problem of parametrization, proposing a mathematical-based approach to solve this long lasted debate and a proof of concept of its application to simple - yet meaningful and preparatory - models.

1.4 Parametrizations so far

Traditionally, the development of parametrizations consisted in deriving deterministic empirical laws able to describe the effect of the small scale dynamical processes on the large scale phenomena, which is explicitly solved in numerical simulations. These parametrizations are often optimized - generally without a sistematic procedure - to have skill on specific variables of interest for weather or climate studies - e.g. temperature, humidity, pressure. More recently, it has become apparent the need to include stochastic terms able to provide a theoretically more coherent representation of such effects including the fluctuations and, at practical level, an improved skill.

The pursuit of stochastic parametrizations - that is, a parametrization which uses random terms to represent the fluctuations of unresolved processes - for weather and climate models has then become an extremely active area of research, see e.g. the recent contributions by Palmer and Williams, 2008; Franzke et al., 2015; Berner et al., 2017 and the now classical collection of results in Imkeller and Storch, 2001. The construction of stochastic parametrizations in geophysical fluid dynamical models is also usually approached using a pragmatic method: one tries to construct empirical functions able to represent well the effect of mean state and of the fluctuations of the unresolved variables, see e.g. the illustrative examples of Orrell, 2003 and Wilks, 2005.

Mathematical arguments do indeed support the idea of going towards stochastic parametrizations. A first way to derive or at least justify the need for them comes from homogenization theory (Pavliotis and Stuart, 2008), which leads to constructing an approximate representation of the impact of the fast scales on the slow variables as the sum of two terms, a mean field term and a white noise term. Such an approach suffers from the fact that one has to take the rather nonphysical hypothesis that an infinite time scale separation exists between the fast and the slow scales. As the climate is a multiscale system, such a methodology is a bit problematic to adopt. Yet, this point of view has been crucial in the development of methods aimed at deriving reduced order models for systems of geophysical interest (Majda, Timofeyev, and Vanden-Eijnden, 1999; Majda, Timofeyev, and Vanden-Eijnden, 2001; Majda, Timofeyev, and Vanden-Eijnden, 2003; Franzke, Majda, and Vanden-Eijnden, 2005).

A different point of view focuses, instead, on constructing effective dynamics comprising deterministic as well as stochastic terms purely from data (Chekroun, Liu, and Wang, 2015a; Chekroun, Liu, and Wang, 2015b). The idea proposed by Kravtsov, Kondrashov, and Ghil, 2005 has been to extend the multilevel linear regressive method, which is suitable for linear problems, to the nonlinear case, allowing for dealing with the possibility of representing quadratic nonlinearities in the evolution equations, which are in fact typical of (geophysical) fluid dynamical processes. The method allows for constructing an optimal representation of the deterministic, linear and nonlinear, dynamics as well as of the stochastic forcing, so that its correlation properties are suitably recovered without making any assumption on the existence of time scale separation between resolved and neglected variables. Afterwards, Kondrashov, Chekroun, and Ghil, 2015 showed how non-markovian data-driven parametrizations emerge naturally when considering partial observations from a large-dimensional system.

Mori, Fujisaka, and Shigematsu, 1974; Zwanzig, 1960; Zwanzig, 1961 analyzed, in the context of statistical mechanics, the related problem of studying how one can project out the effect of a group of variables, with the goal of constructing effective evolution equations for a subset of variables of interest. They reformulated the dynamics of such variables expressing them as a sum of three terms, a deterministic term, a stochastic forcing and a memory term. The memory term defines a non-markovian contribution where the past states of the variables of interest enter the evolution equation. In the limit of infinite time scale separation such last term tends to zero, whilst the random forcing approaches the form of a (in general, multiplicative) white noise, in agreement with what predicted by the homogenization theory.

In a few recent papers, Wouters and Lucarini (Wouters and Lucarini, 2012; Wouters and Lucarini, 2013; Wouters and Lucarini, 2016) have provided explicit formulas for constructing parametrizations able to incorporate the deterministic, stochastic, and non-markovian components. The formulas have been obtained independently using two rather different approaches, namely a second order expansion of the Mori-Zwanzig projection operator, and a reworking of the Ruelle (Ruelle, 1998; Ruelle, 2009) response theory, which allows under suitable conditions to compute the change in the expectation value of any smooth observable of a system resulting from perturbations of the dynamics in terms of the statistical properties of the unperturbed flow.

The idea followed by Wouters and Lucarini has been to treat the coupling between the slow and fast variables as the weak forcing added on top of the uncoupled dynamics, and then evaluate the impact of the forcings on the statistical properties of a generic observable of the slow variables. Finally, the last step has been to retro-engineer explicit formulas for terms that, added to the uncoupled dynamics of the slow variables, provide up to second order the same results as the actual coupling.

It is important to note that since explicit formulas are provided, one can indeed construct the parametrizations *ab-initio*, and not empirically. Additionally, the parametrization is automatically optimized for all possible observables of the system. Such an approach seems especially promising in all systems, as in the extremely relevant case of the climate, where there is no spectral gap in the scales of motions that justifies the assumption of infinite time scale separation between fast and slow scales. It seems then in general relevant to be able to retain and check the relevance of the memory term and to construct a suitable model for the stochastic forcing, going beyond the approximation of using white noise or simple empirical autoregressive processes.

Note that the approach discussed here is not *per se* constructed to deal with multiscale systems only. In fact, the explicit expressions for the terms responsible for the parametrization (see Section 2.1) are constructed by performing an asymptotic expansion controlled by a parameter determining the degree of coupling between the set of variables of interest and those to parametrize. Clearly, if such a condition is satisfied, it is possible to apply this method also to multiscale systems. Recently, a parametrization constructed according to such a statistical mechanical point of view has been tested successfully in a

simple low dimensional model (Wouters et al., 2016) and in a more complex yet simple coupled model (Demaeyer and Vannitsem, 2017).
A further degree of flexibility of this approach has been explored in another recent publication (Lucarini and Wouters, 2017), which provided explicit formulas for modifying the parametrization when the parameters controlling the dynamics of the full system are altered.

1.5 Objectives of this thesis

A fundamental problem in the construction of parametrizations is that, even when they are efficient enough to represent unresolved phenomena with the desired precision, they are typically tuned for being accurate for a specific configuration of a model in terms of numerical resolution, and the operation of retuning needed when a new model version at higher resolution is available can be extremely tedious and costly in computational terms. The need of achieving scale-adaptive parametrizations has been recently emphasized in the scientific literature, see e.g. Arakawa, Jung, and Wu, 2011; Park, 2014; Sakradzija, Seifert, and Dipankar, 2016. Additionally, parametrizations are typically tested against specific observables of interest and tuned in order to better represent those observables, but it is not always clear whether optimizing the skill for such observables might come at the price of reducing the skill on other climatic properties that might prove crucial for, e.g., modulating the climatic response to forcings.

The first task of this dissertation is to stress the possibility of having automatically scale adaptive formulations of Wouters-Lucarini methodology, testing at the same time its skills in reproducing the statistics of the model used as benchmark system to work with, (a modified version of) the Lorenz 96 model (Lorenz, 1996), which provides a prototypical yet convincing representation of a two-scale system where large scale, synoptic variables are coupled to small scale, convective variables. The Lorenz 96 model has quickly become the testbed for evaluating new methods of data assimilation (Trevisan and Uboldi, 2004; Trevisan, Isidoro, and Talagrand, 2010) and is receiving a lot of attention also in the community of statistical physics (Abramov and Majda, 2008; Hallerberg et al., 2010; Lucarini and Sarno, 2011; Gallavotti and Lucarini, 2014). More importantly for this specific case, the Lorenz 96 model has been used in the papers of Orrell, 2003 and Wilks, 2005 to construct explicit models of stochastic parametrization, so there are previous results to compare to.

In the second part of the dissertation the WL parametrization has been applied to a simple dynamical system introduced by Bódai, Károlyi, and Tél, 2011 and constructed by coupling the Lorenz 84 model (Lorenz, 1984) with the Lorenz 63 model (Lorenz, 1963), introducing the Wasserstein distance (Villani, 2009) as a new statistical methodology to provide a quantitative estimate of the skill of a parametrization. The aim is to parametrize the dynamical effect of the variables corresponding to the Lorenz 63 system on the variables corresponding to the Lorenz 84 system, changing time scale separation to switch the roles of slow and fast scale systems between the two models, which evolve on the same grid-scale (no subgrid phenomenon involved). In order to extend what investigated in the first part, the study focuses on doing a systematic comparison of the properties of the projected measure of the original coupled system on the subspace spanned by the variables of the Lorenz 84 model with the actual measure of the parametrized model. In particular, the Wasserstein distance (see Section 2.3) between the coarse-grained estimates of the two invariant measures - which are measures μ for which, with respect to a transformation M, $\mu(A) = \mu(M^{-1}A)$ (see Ott, 1993 for a more detailed discussion about measures) - has been computed. Roughly speaking, the invariant measure is the time fraction that a trajectory on an attractor spends in a particular n-dimensional cube in the phase space, and its analysis allows to determine the long term statistical properties of the systems studied. Here it is proposed to test the skill of a parametrization to reproduce those properties through the computation of the distances between the measures. An additional analysis of the Wasserstein distance of the measures obtained by projecting on two of the three variables of interest allows for a comprehensive evaluation of how different the one-time statistical properties of the two systems are. The Wasserstein distance has been proposed by Ghil, 2015 as a tool for studying the climate variability and response to forcings, and applied by Robin, Yiou, and Naveau, 2017 in a simplified setting.

The results in this thesis are adapted from Vissio and Lucarini, 2018a; Vissio and Lucarini, 2018b.

1.6 Structure of the thesis

The dissertation is structured as follows.

- Chapter 2 is divided in three main sections. Section 2.1 provides the main ingredients of the method for constructing general parametrizations introduced by Wouters and Lucarini, 2012; Wouters and Lucarini, 2013; Wouters and Lucarini, 2016. Section 2.2 describes the Lorenz 96, Lorenz 84 and Lorenz 63 models, highlighting the modifications applied in the present work. Section 2.3 briefly introduces the Wasserstein distance, explaining how it is employed in this study.

- Chapter 3 describes the results of the tests performed on WL parametrization. In particular, Section 3.1 shows the application on Lorenz 96 model, discussing the scale adaptive properties, comparing the performance with Wilks approach and looking at the statistical properties of the obtained distributions. Section 3.2 is based on the parametrization of the Lorenz 63 model as forcing of the Lorenz 84 model, with a discussion on qualitative and quantitative performances and the computation of the Wasserstein distance between the projected measures of the WL parametrization and the full model.
- Chapter 4 draws the conclusions of the analysis performed.

Chapter 2

Mathematical Background

2.1 Wouters-Lucarini's Parametrization

Wouters and Lucarini (Wouters and Lucarini, 2012; Wouters and Lucarini, 2013; Wouters and Lucarini, 2016) introduced a methodology for constructing parametrizations for dynamical systems of the form:

$$\frac{dX}{dt} = F_X(X) + \epsilon\Psi_X(X,Y), \tag{2.1}$$

$$\frac{dY}{dt} = F_Y(Y) + \epsilon\Psi_Y(X,Y), \tag{2.2}$$

where the X variables correspond to the dynamics of interest and the Y variables correspond to the dynamics to parametrize, i.e. the unresolved dynamics. The F vector field on the right hand side of Eqs.(2.1)-(2.2) corresponds to the uncoupled dynamics of the X and Y variables respectively, while the Ψ field describes the coupling, with ϵ being a bookkeeping variable describing the coupling strength. Note that Eqs.(2.1)-(2.2) do not describe, in general, a multiscale dynamical system, where the X (slow) and the Y (fast) variables are essentially characterized by different scales of motion. Nonetheless, it is possible to bring it to the standard form elucidating multiscale behaviour by considering the following form for Eqs.(2.1)-(2.2):

$$\frac{dX}{dt} = F_X(X) + \epsilon\Psi_X(X,Y) \tag{2.3}$$

$$\frac{dY}{dt} = \gamma\tilde{F}_Y(Y) + \epsilon\Psi_Y(X,Y) \tag{2.4}$$

where $\gamma \gg 1$ and $F_Y(Y) = \gamma\tilde{F}_Y(Y)$. As clear from the later discussion, it is not important in the cases undertaken in this dissertation to include the factor γ also for the coupling term affecting the Y variables in Eq. (2.4), because the interest lies in separating the time scales of the decoupled ($\epsilon = 0$) X- and Y-systems. Following the discussion presented in the introduction, the goal is to find an approximate equation of the form

$$\frac{dX}{dt} = F_X(X) + \chi\{X\} \tag{2.5}$$

able to provide a good approximation of the statistical properties of the X variables, where $\chi\{X\}$ can in general correspond to an integro-differential contribution with also a stochastic component. It seems relevant aiming at being able to specify in advance the accuracy of the approximation in terms of the properties of the coupling and, in particular, of the coupling strength ϵ. Clearly, if $\epsilon = 0$, the resulting $\chi\{X\} = 0$ provides a (trivial) solution to the problem. The approach can be seamlessly followed also in the presence of a functional form for the equations where the parameter γ explicitly controls the scale separation between the X and Y variables. Note that Abramov, 2016 introduced an extension of the homogenization method able to deal with a problem formulated as in Eqs.(2.1)-(2.2).

2.1.1 The method

The basic idea is to consider the dynamical system (2.1)-(2.2) as resulting from an $\epsilon-$perturbation of the following dynamical system:

$$\frac{dX}{dt} = F_X(X), \tag{2.6}$$

$$\frac{dY}{dt} = F_Y(Y), \tag{2.7}$$

where the coupling plays the role of the perturbation. The focus is now on the X variables, considering a general observable $A = A(X)$, i.e. a smooth function of the X variables only. Making suitable hypotheses on the mathematical properties of the unperturbed system and taking advantage of the Ruelle response theory (Ruelle, 1998; Ruelle, 2009), Wouters and Lucarini have been able to find a useful expression for the expectation value $\rho_\epsilon(A)$ of the observable A taken accordingly to the invariant measure $\rho_\epsilon(dXdY)$ of the coupled dynamical system (2.1)-(2.2):

$$\rho_\epsilon(A) = \int \rho_\epsilon(dXdY)A(X). \tag{2.8}$$

In what follows, it is assumed that all invariant measures considered are of the Sinai-Ruelle-Bowen kind (Eckmann and Ruelle, 1985; Young, 2002). This assumption can be physically motivated by taking into account the chaotic hypothesis which, in the context of the theory of turbulence (Ruelle, 1978), allows to consider the attracting set of a dynamical system as smooth surfaces (e.g. Gallavotti, 2014).

The projected measure is introduced by

$$\rho^*_\epsilon(dX) = \int dY \rho_\epsilon(dXdY) \tag{2.9}$$

such that $\rho_\epsilon(A) = \rho^*_\epsilon(A)$. In case of smooth and hyperbolic systems, the chaotic hypothesis implies the ergodic hypothesis. This property subsequently implies that a natural measure average is replaceable with a time average and Eq.(2.9) can be calculated as

$$\rho^*_\epsilon(A) = \lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau A(x(t)), \tag{2.10}$$

where $x(t) = \widetilde{f}^t x_0$, with $\widetilde{f}^t$ defining the flow determined by the dynamical system (2.1)-(2.2). It is possible to find a perturbative expansion of the expectation value of A taken accordingly to the invariant measure of the coupled system. One can in fact write

$$\rho^*_\epsilon(A) = \rho_{0,X}(A) + \epsilon\delta^{(1)}_{\Psi}\rho(A) + \epsilon^2\delta^{(2)}_{\Psi,\Psi}\rho(A) + \mathcal{O}(\epsilon^3), \tag{2.11}$$

where the first term $\rho_{0,X}(A)$ is the expectation value of A taken accordingly to the invariant measure of the X-component of the unperturbed system (2.6):

$$\rho_{0,X}(A) = \int \rho_{0,X}(dX) = \lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau A(f^\tau(x_0)), \tag{2.12}$$

where it was again used ergodicity and defined f^t as the flow of the X variables part of the dynamical system (2.6). The second term $\epsilon\delta^{(1)}_{\Psi}\rho(A)$ and the third term $\epsilon^2\delta^{(2)}_{\Psi,\Psi}\rho(A)$ correspond to the first and second order corrections, and can be also expressed as expectation values on $\rho_{0,X}(dX)$ of explicitly determined observables, which are constructed non-trivially from A and the vector field Ψ. All the terms can be computed from the statistical properties of the uncoupled dynamics of the Y variables given in Eq.(2.7). The explicit expressions can be found in Wouters and Lucarini, 2012.

While the previous result allows for computing the impact of the coupling on the statistics of any given A observable, it is not useful *per se* for constructing a parametrization. Nonetheless, it is possible to retro-engineer an educated guess for the term $\chi\{X\}$ introduced in Eq.(2.5), such that *up to second order in* ϵ the expectation value of A according to the invariant measure $\rho'_\epsilon(dX)$ of the system:

$$\frac{dX}{dt} = F_X(X) + \epsilon D(X) + \epsilon S\{X\} + \epsilon^2 M\{X\} \tag{2.13}$$

is the same as the expectation value of A according to ρ_ϵ, or, more explicitly:

$$\rho_\epsilon(A) = \rho'_\epsilon(A) + O(\epsilon^3). \tag{2.14}$$

Therefore, Eq.(2.14) provides a useful basis for defining a parametrization where it is possible to control the error on the statistics of the surrogate dynamics with respect to the full dynamics as a function of ϵ, and where this applies *for all possible observables* A.

The three perturbation vector fields D, S and M correspond to, respectively, a mean field term, a stochastic forcing and a non-markovian memory term. Note that the stochastic term has a second order effect on the measure even if its intensity is proportional to ϵ; see Lucarini, 2012. As shown in Wouters and Lucarini, 2012; Wouters and Lucarini, 2013; Wouters and Lucarini, 2016, the explicit expression for these three terms can be obtained also by performing a second order expansion of the Mori-Zwanzig projector operator, which constructs the effective projected dynamics for the X variables only describing the instantaneous evolution of the trajectories in the phase space. This suggests that the proposed parametrization, obtained by means of Ruelle response theory and therefore based on the description of long term averages, might have skill also in terms of prediction on the short time scale (in the sense of weather forecast). Nonetheless, this hypothesis is not analyzed in this dissertation and should be investigated elsewhere. In what follows, this approach will be referred to as the WL parametrization.

The explicit expressions for the three terms providing the parametrization shown in Eq.(2.13) are given below in Eqs.(2.15), (2.16) and (2.20). Therefore, once D, S, and M are derived, they can be used to construct parametrizations for all values of ϵ within the radius of convergence of the expansion. Additionally, if the coupled model given in Eqs.(2.1)-(2.2) is multiscale, this approach allows for constructing parametrizations integrating the single scale equation (2.7). This can significantly ease the computational burden of the problem.

Deterministic, stochastic, and non-markovian terms

Assuming the coupling terms $\Psi_X(X,Y)$ and $\Psi_Y(X,Y)$ separable in the X and Y variables, they can be written as $\Psi_X(X,Y) = \Psi_{X,1}(X)\Psi_{X,2}(Y)$ and $\Psi_Y(X,Y) = \Psi_{Y,1}(X)\Psi_{Y,2}(Y)$. As explained in Wouters and Lucarini, 2012; Wouters and Lucarini, 2013; Wouters and Lucarini, 2016, such an assumption does not really impact the generality of the results presented here.

$D(X)$ is a deterministic term that accounts for the average impact that the coupling has on the X variables and it is given by:

$$D(X) = \Psi_{X,1}(X)\rho_{0,Y}(\Psi_{X,2}(Y)). \tag{2.15}$$

The second order contribution is composed of two parts. $S\{X\}$ represents a stochastic forcing due to the temporal correlation of the fluctuations of the forcing exerted by the Y-variables onto the X variables. It can be written

$$S\{X\} = \Psi_{X,1}(X)\sigma(t), \tag{2.16}$$

where $\sigma(t)$ is a stochastic term and is constructed in such a way to reproduce the lagged correlation of the fluctuations of the forcing. The statistical properties of the noise $\sigma(t)$ can be expressed as:

$$\begin{aligned} R(t) &= \langle \sigma(t), \sigma(0) \rangle \\ &= \rho_{0,Y}\left((\Psi_{X,2}(Y) - \rho_{0,Y}(\Psi_{X,2}(Y)))(\Psi_{X,2}(f^t(Y)) - \rho_{0,Y}(\Psi_{X,2}(Y))) \right), \end{aligned} \tag{2.17}$$

$$\langle \sigma(t) \rangle = 0. \tag{2.18}$$

where the brackets indicate the expectation value of the stochastic process and $R(t)$ is the lagged correlation of the (stationary) noise.

Finally, $M\{X\}$ is a memory term that describes the effects of the history of the X variables on their present value through the influence of the Y variables. This term is essential for capturing the effect of the hidden (Y) variables on the (X) variables of interest, as clarified by Chekroun, Liu, and Wang, 2015a; Chekroun, Liu, and Wang, 2015b. It is expressed as:

$$M\{X\} = \int_0^\infty h(\tau, X(t-\tau))d\tau, \tag{2.19}$$

where the integral kernel is given by:

$$h(\tau, \tilde{X}) = \Psi_{Y,1}(\tilde{X})\Psi_{X,1}(f^{\tau}(\tilde{X}))\rho_{0,Y}(\Psi_{Y,2}(Y)\partial_Y \Psi_{X,2}(f^{\tau}(Y))). \tag{2.20}$$

Such an average resembles a cross-correlation between the actual state of the two fields X, Y and the deviation of the trajectory of the same fields evolved at $t = \tau$.

A remarkable property of this parametrization is its universality, as shown by Eq.(2.15) through (2.20), because the three factors D, S and M can be computed for any given expression of the coupling terms or of the uncoupled dynamics.

Another positive aspect of these equations is the scale adaptivity of the parametrization terms. Considering the case where the equations of motions can be written as in Eqs.(2.6)-(2.7), one can check that the expectation values are computed according to the invariant measure of the uncoupled equation $\frac{dY}{dt} = \gamma \tilde{F}_Y(Y)$, which can be rewritten as

$$\frac{dY}{d\tau} = \tilde{F}_Y(Y) \tag{2.21}$$

where $\tau = \gamma t$.

The constant D in Eq.(2.15) is clearly not affected by the choice of the time scale. Instead, the correlation function in Eq.(2.17) and the memory kernel in Eq.(2.20) are affected by the rescaling in the time and only the rescaled time τ will appear in their arguments. By substituting $\tau = \gamma t$ one then obtains the actual parametrization for every choice of γ. In particular, large values of γ will lead to a compression of the time axis for the correlation function and the memory kernel. In the limit of $\gamma \to \infty$ (i.e. infinite time scale separation), the stochastic forcing tends to a white noise and the non-Markovian term vanishes, accordingly with the predictions of the homogenization and Mori-Zwanzing theories (see Section 1.4).

Independent coupling case

The special case where the two coupling terms are independent from the variable they are affecting, namely $\Psi_X(X, Y) = \Psi_X(Y)$ and $\Psi_Y(X, Y) = \Psi_Y(X)$, is particularly important for the scopes of this thesis. The three terms discussed above take the following simpler form:

$$D(X) = \rho_{0,Y}(\Psi_X(Y)), \tag{2.22}$$

$$S\{X\} = \sigma(t), \tag{2.23}$$

where

$$\begin{aligned} R(t) = \langle \sigma(0), \sigma(t) \rangle = \rho_{0,Y}((\Psi_X(Y) - D)(\Psi_X(f^t(Y)) - D)), \\ \langle \sigma(t) \rangle = 0, \end{aligned} \tag{2.24}$$

and

$$M\{X\} = \int_0^\infty h(t_2, X(t - t_2))dt_2, \tag{2.25}$$

where

$$h(t_2, \tilde{X}) = \Psi_Y(\tilde{X})\rho_{0,Y}(\partial_Y \Psi_X(f^{t_2}(Y))). \tag{2.26}$$

In this special case, the stochastic contribution reduces to a simple additive noise term - compare Eqs.(2.16) and (2.23) - while the evaluation of the memory kernel h is significantly easier as a simpler expression appears in the ensemble average - compare Eqs.(2.20) and (2.26).

Forcing in the fast dynamics

As discussed in Wouters and Lucarini, 2012; Wouters and Lucarini, 2013; Wouters and Lucarini, 2016, a basic requirement for the proposed approach to allow for the construction of a parametrization for the Y variables is to have that the uncoupled dynamics of the Y variables given in Eq.(2.7) features a non-trivial invariant measure and fast decay of correlations due to the presence of chaos. Physically, this requires presence of an external forcing, leading to the injection of energy for the Y variables; this is achieved in the system studied here by choosing a sufficiently large value for the constant F_2. Another way to address such a problem is shown in Wouters et al., 2016, where a stochastic forcing, corresponding to the presence of energy injection coming from even smaller, unresolved scales, is considered.

In order to extent the method to physical situations where energy is injected only in the X variables, a simple mathematical operation, that amounts to changing the background state around which the perturbation induced by the presence of coupling is considered, will be applied.

The idea is to rewrite Eq.(2.2) (here for simplicity the case with $\Psi_X(X,Y) = \Psi_X(Y)$ and $\Psi_Y(X,Y) = \Psi_Y(X)$ is dealt with) as follows:

$$\frac{dY}{dt} = F_Y(Y) + \epsilon G + \epsilon \Psi_Y(X) - \epsilon G, \tag{2.27}$$

such that the vector flows defining the uncoupled dynamics and the coupling are defined as follows:

$$\widetilde{F}_Y(Y) = F_Y(Y) + \epsilon G, \tag{2.28}$$

$$\widetilde{\epsilon\Psi_Y}(X) = \epsilon\Psi_Y(X) - \epsilon G. \tag{2.29}$$

The choice of the *artificial* forcing G gives a degree of flexibility and must obey only the requirement that $\dot{Y} = \tilde{F}_Y(Y)$ is chaotic. Note that, within the radius of expansion ensuring the validity of the perturbative approach, the specific choice of G affects only weakly the final result.

An obvious choice is to choose $G = \rho_{0,X}(\Psi_Y(X))$, which makes sure that, at zero order, the uncoupled system has nontrivial dynamics.

2.2 Lorenz Models

In this Section the models used in the thesis, either the original and the modified versions, are exposed.

2.2.1 Lorenz 96

The Lorenz 96 model (Lorenz, 1996), represented schematically in Fig.2.1, provides a conceptually meaningful yet extremely simplified representation of the atmosphere; there are two sets of variables, one describing the dynamics on large scale (so-called synoptic variables), and one characterizing the dynamics on small scale (so-called convective variables). The convective variables are divided in as many subgroups of equal size as the number of synoptic variables, each subgroup being coupled to a different synoptic scale variable. The system is then characterized by coupling within and across scales of motions. The Lorenz 96 model has quickly established itself as one of the reference models in nonlinear dynamics for testing e.g. data assimilation (Trevisan and Uboldi, 2004; Trevisan, Isidoro, and Talagrand, 2010), schemes and properties of Lyapunov exponents and covariant Lyapunov vectors and is becoming increasingly popular also within the community of statistical mechanics (Abramov and Majda, 2008; Hallerberg et al., 2010; Lucarini and Sarno, 2011; Gallavotti and Lucarini, 2014).

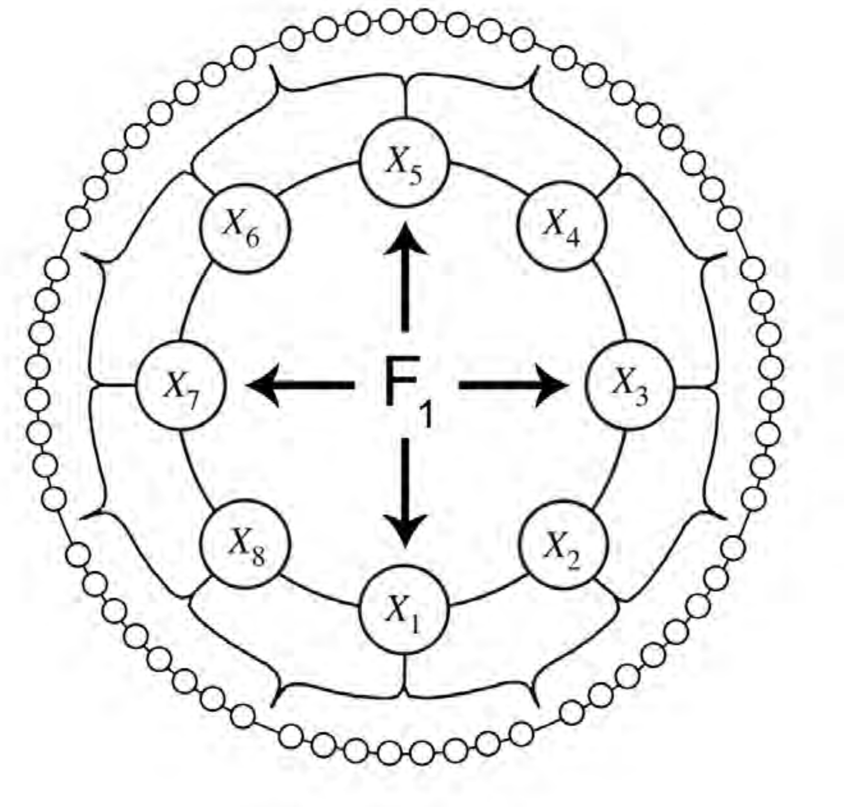


FIGURE 2.1: A schematic representation of Lorenz 96 model in the case $K = 8$ and $J = 8$, with the inner circle composed by the slow X variables while the outer circle represents the fast J variables (adapted from Wilks, 2005).

The evolution equations of the two-layers, coupled Lorenz 96 model can be written as:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F_1 - \frac{hc}{b}\sum_{j=1}^{J} Y_{j,k}, \tag{2.30}$$

$$\frac{dY_{j,k}}{dt} = cbY_{j+1,k}(Y_{j-1,k} - Y_{j+2,k}) - cY_{j,k} + \frac{hc}{b}X_k, \tag{2.31}$$

with $k = 1, ..., K$; $j = 1, ..., J$. The boundary conditions are defined as

$$\begin{aligned} X_{k-K} &= X_{k+K} = X_k, \\ Y_{j,k-K} &= Y_{j,k+K} = Y_{j,k}, \\ Y_{j-J,k} &= Y_{j,k-1}, \\ Y_{j+J,k} &= Y_{j,k+1}. \end{aligned} \tag{2.32}$$

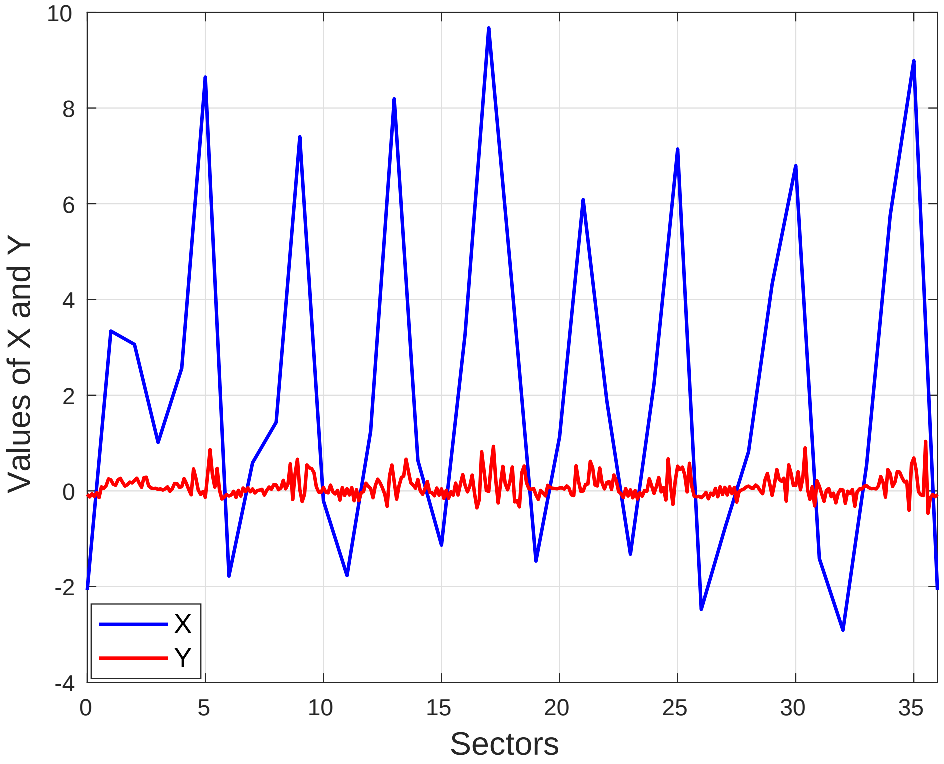


FIGURE 2.2: Typical longitudinal profiles of X_k and $Y_{j,k}$.

The latitudinal circle is divided into K sectors, each one corresponding to one *synoptic* slow X variable. Each X variable is coupled to J *convective* fast Y variables. The constant c defines the time scale separation between the fast and slow variables (see also the general form of a multiscale system as given in Eqs (2.3)-(2.4)), while the amplitude of the fluctuations is determined by b, while h controls the strength of the coupling.

In absence of forcing and dissipation, the sum of the squares of the variables (the *energy* of the system) is conserved. For a detailed description of the statistical mechanical and conservation properties of the system (yet in a simplified version), the reader is encouraged to look into Lucarini and Sarno, 2011; Blender and Lucarini, 2013; Gallavotti and Lucarini, 2014.

The coupling between the X and the Y terms has the simplified form discussed in the previous Section (the so-called independent coupling), and is linear. This simplifies the application of the Wouters-Lucarini parametrization (see Section 2.1.1), which is nonetheless possible also for more complex forms of coupling.

The choice of the parameters defining the strength of the external forcing, the number of sectors and subsectors, the strength of the coupling, the relative

amplitude of the fluctuations and the time scale separation between the two systems determines the properties of the dynamical system.
The original parameters chosen by Lorenz are $c = 10.0$, $b = 10.0$, $h = 1.0$, $K = 36$ and $J = 10$ - providing therefore a total of 36 X variables and 360 Y variables. Following the original derivation of the model, 1 unit of time is equivalent to 5 days, while the usual integration time step is 0.005, corresponding to 36 minutes. In Fig.2.2 a typical profile both of the X_k's and of the $Y_{j,k}$'s is shown.

When one is well within the chaotic regime (e.g. F_1 is sufficiently large) and considers a sufficiently large number of sectors (and subsectors), it is reasonable to expect to be able to define intensive properties that are stable with respect to the specific choice of K and J, see discussion in Gallavotti and Lucarini, 2014 for a simpler version of the model.

Another version of Lorenz 96 is the one-layer model:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F_1, \tag{2.33}$$

where the effect of subgrid scale is neglected. This uncoupled Lorenz 96 model works as a useful benchmark to assess the influence of fast scales and, consequently, parametrizations based on their dynamics.

Modifications to the model

Two modifications to the standard Lorenz 96 model have been implemented (see Fig.2.3):

- The introduction of a forcing term also in the equations describing the dynamics of the Y variables (see Eq.(2.35)), in order to represent the direct effect of forcings at small scales (mimicking, e.g., the impact of direct solar forcing on convective motions). This has the effect of making the fast variables an active component of the system: they can also pump energy into the X variables and are not exclusively dissipating energy coming from larger scales.

- The modification of the boundary conditions on the Y variables in such a way that the fast variables of different sectors do not interact with each other, in the spirit of having the fast variables representing sub-grid scale phenomena (see Eqs.(2.36)). Note that if $J \gg 1$ and in chaotic dynamics regime, it is reasonable to think that this change has negligible impact on

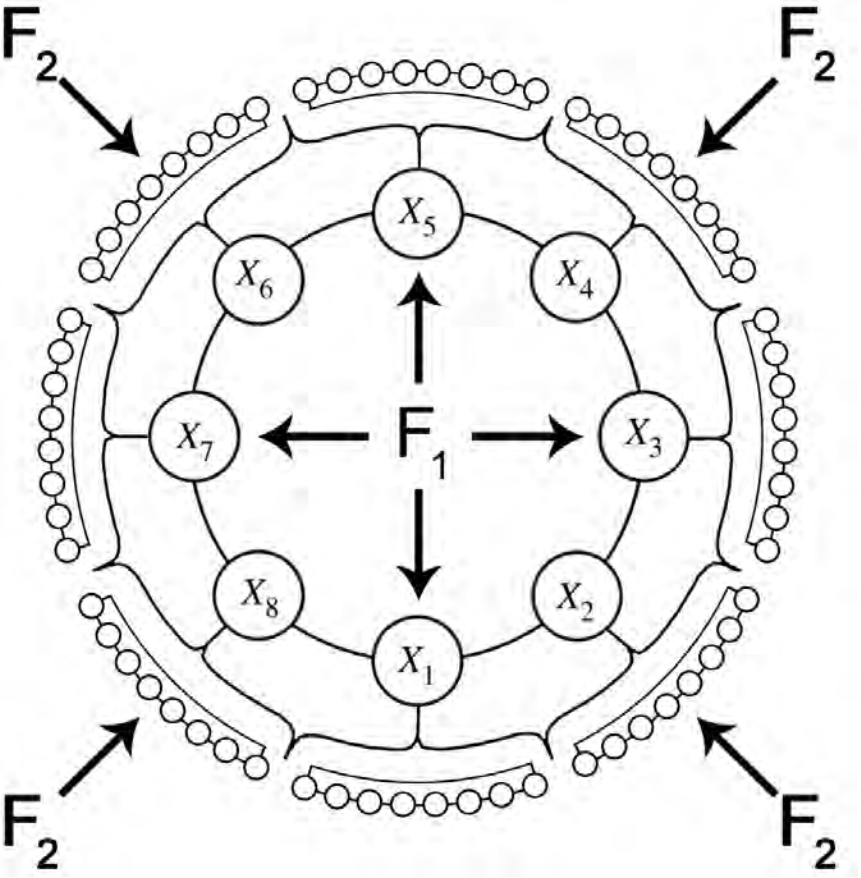


FIGURE 2.3: A schematic representation of the modified Lorenz 96 model used in this thesis (adapted from Wilks, 2005). Note the presence of the external forcing F_2 and the division of the Y variables in non interacting subsections.

the statistics of the system, as information does not propagate efficiently between convective variables belonging to neighbouring sectors. Additionally, the parametrization becomes easier to implement, because, following the basic idea behind super-parametrization, subgrid variables belonging to different X sectors are independent and equivalent in the uncoupled case (see Eqs.(2.6)-(2.7)).

Therefore, the evolution equations (2.30)-(2.31) are modified as follows:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F_1 - \frac{hc}{b}\sum_{j=1}^{J} Y_{j,k}, \tag{2.34}$$

$$\frac{dY_{j,k}}{dt} = cbY_{j+1,k}(Y_{j-1,k} - Y_{j+2,k}) - cY_{j,k} + \frac{c}{b}F_2 + \frac{hc}{b}X_k, \tag{2.35}$$

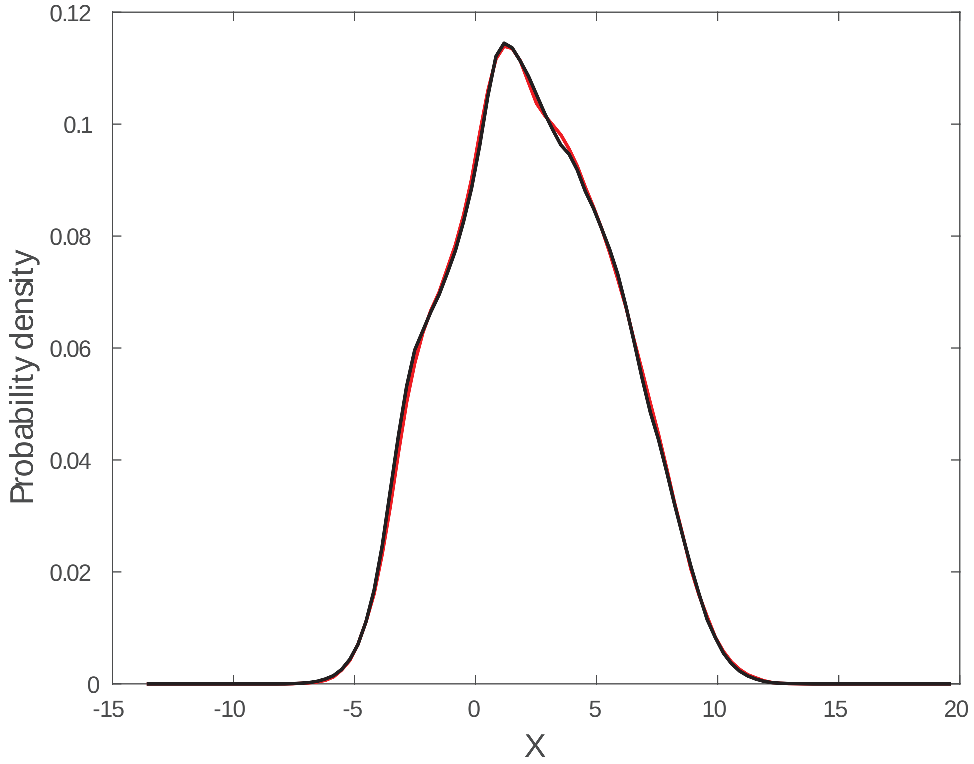


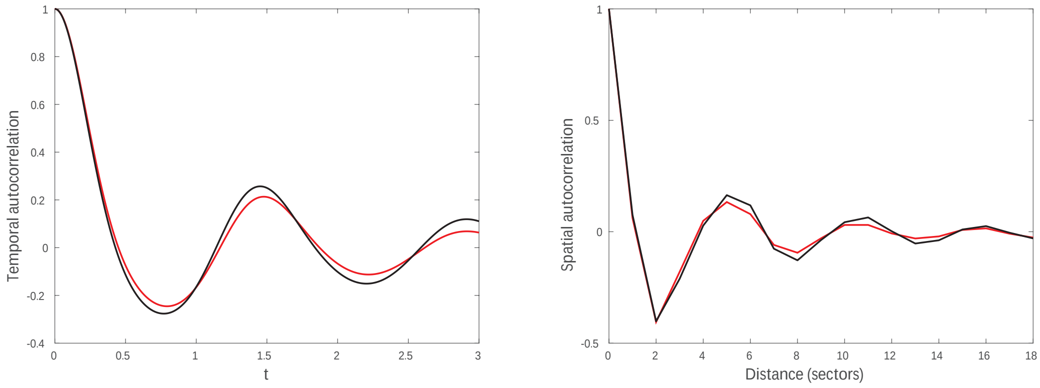
FIGURE 2.4: Probability density of the X variable for the original (black line) and the modified (red line) Lorenz 96 model.

with modified boundary conditions

$$\begin{aligned} X_{k-K} &= X_{k+K} = X_k, \\ Y_{j-J,k} &= Y_{j+J,k} = Y_{j,k}. \end{aligned} \tag{2.36}$$

The parameter ϵ in Eqs.(2.1)-(2.2) is $\frac{hc}{b}$ and the coupling terms are $\Psi_X = -\epsilon \sum_{j=1}^{J} Y_{j,k}$ and $\Psi_Y = \epsilon X_k$, b defines the ratio between the typical size of the X and Y variables, while the parameter γ controlling the scale separation is given by c. The external forcings are set as $F_1 = 10.0$ and $F_2 = 6.0$, so that chaos is realized in the uncoupled version of the system (obtained from Eqs.(2.34)-(2.35) by setting $h = 0$) for both the large and small scale variables of the system separately. For h, b, c, K, and J the standard values mentioned above are chosen.

Changes in the boundary conditions for the Y variables have a negligible effect on the statistical properties of the X variables: the probability distribution of each X variable (Fig.2.4), its time correlation (Fig.2.5a), and the spatial correlation of the X variables at zero time lag (Fig.2.5b) are virtually identical



(A)

(B)

FIGURE 2.5: a) Time autocorrelation; b) Spatial autocorrelation of the X variable for the original (black line) and the modified (red line) Lorenz 96 model.

for the original and the modified Lorenz 96 model. The presence of chaos and of a corresponding nontrivial invariant measure for the Y variables are necessary for being able to construct the WL parametrization.

Wilks Parametrization

Wilks proposed an empirical, data-driven parametrization (Wilks, 2005) of the fast dynamical variables specifically for the Lorenz 96 model deriving a fourth order polynomial and a stochastic term to replace the coupling in the X variables equation.

The idea is to fit the unresolved tendencies of the X variables (i.e. the coupling terms $\frac{hc}{b} \sum_{j=1}^{J} Y_{j,k}$ in Eq.(2.30)) using a polynomial regression in the form

$$g_U(X_k) = b_0 + b_1 X_k + b_2 X_k^2 + b_3 X_k^3 + b_4 X_k^4 + e_k, \tag{2.37}$$

where the bs are the regression coefficients, while e_k is a stochastic function constructed according to the following AR(1) process:

$$e_k(t) = \phi e_k(t - \Delta) + \sigma_e (1 - \phi^2)^{1/2} z_k(t), \tag{2.38}$$

written in term of the fitting parameters ϕ (lag-1 autocorrelation of e_k), σ_e (standard deviation of the process e_k), where z_k is a Gaussian uncorrelated process with zero mean and unitary variance.

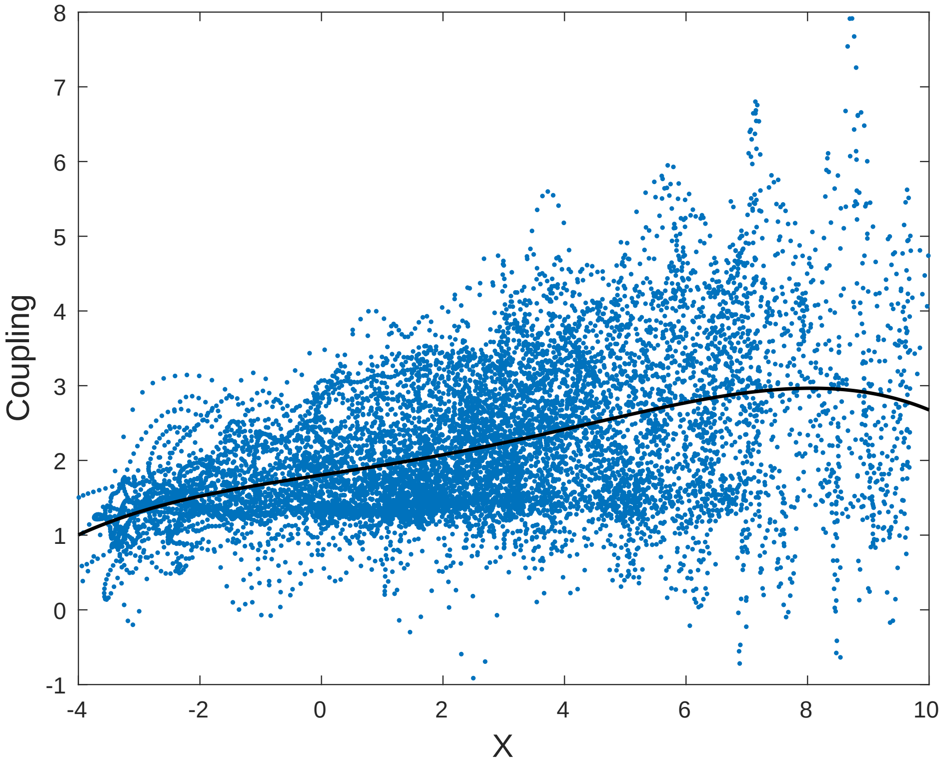


FIGURE 2.6: Example of scatterplot with related fit in the case of Lorenz 96 model.

The polynomial regression in Eq.(2.37) is computed from the scatterplot obtained with the values of the coupling from a time series of the coupled model plotted with respect to the corresponding values of the X's. Then, fourth order fit reveals the desired coefficients (see Fig.2.6), while the parameters ϕ and σ_e in Eq.(2.38) must be calculated using the same time series.

The parametrized model is then written as follows:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F_1 - g_U(X_k). \tag{2.39}$$

Eq.(2.39) represents the general form of a data-driven parametrization, with $g_U(X_k)$ depending on the case studied. Note that for the example studied by Wilks all terms are markovian and there is no clear justification of why the stochastic residual is captured by an $AR(1)$ process, nor of why a fourth order polinomial is chosen. On the other side, the WL parametrization provides a simple constant as deterministic term $D(X)$ (see Eq.(3.4)), which seems an oversimplification compared to the fourth order polynomial used by Wilks.

As previously stated, Wilks developed this approach for the Lorenz 96

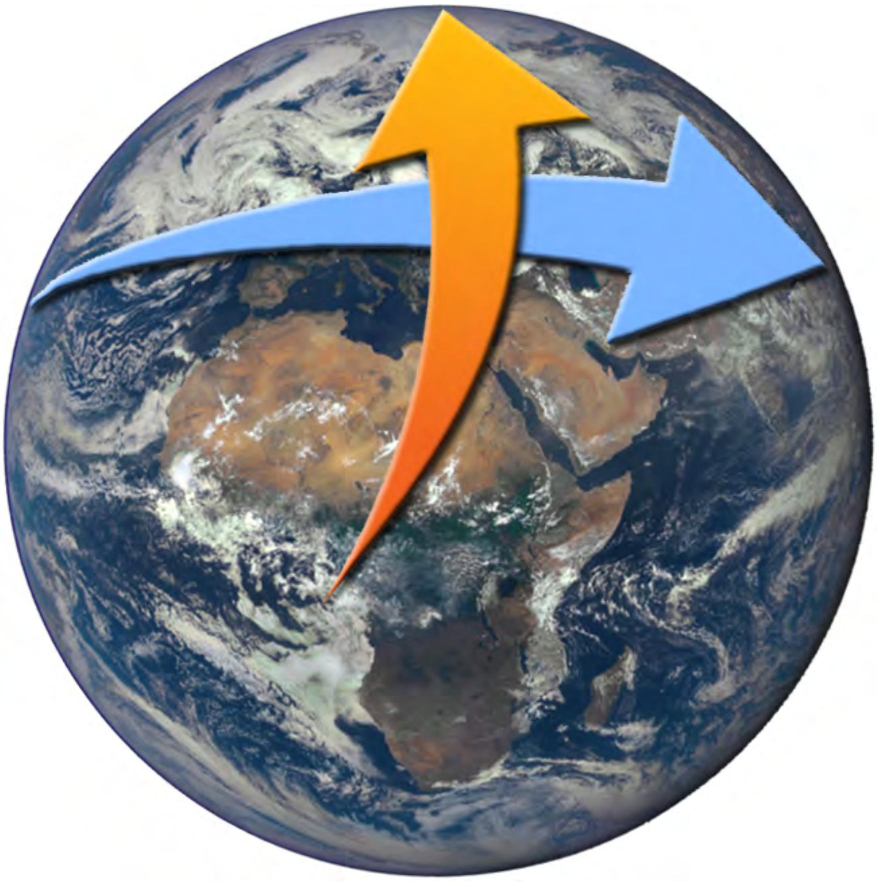


FIGURE 2.7: Schematic representation of the Lorenz 84 model. The blue arrow represents the westerlies, therefore the X variable, while the orange arrow represents the poleward heat transport and is therefore related to the Y and the Z variables.

model; one of the aims of this study is to check the performance of this parametrization against the WL methodology.

2.2.2 Lorenz 84

The Lorenz 84 model (Lorenz, 1984) provides an extremely simplified representation of the large scale atmospheric circulation:

$$\frac{dX}{dt} = -Y^2 - Z^2 - aX + aF_0, \tag{2.40}$$

$$\frac{dY}{dt} = XY - bXZ - Y + G, \tag{2.41}$$

$$\frac{dZ}{dt} = XZ + bXY - Z. \tag{2.42}$$

where the variable X describes the intensity of the westerlies, while the variables Y and Z correspond to the two phases of the planetary waves responsible

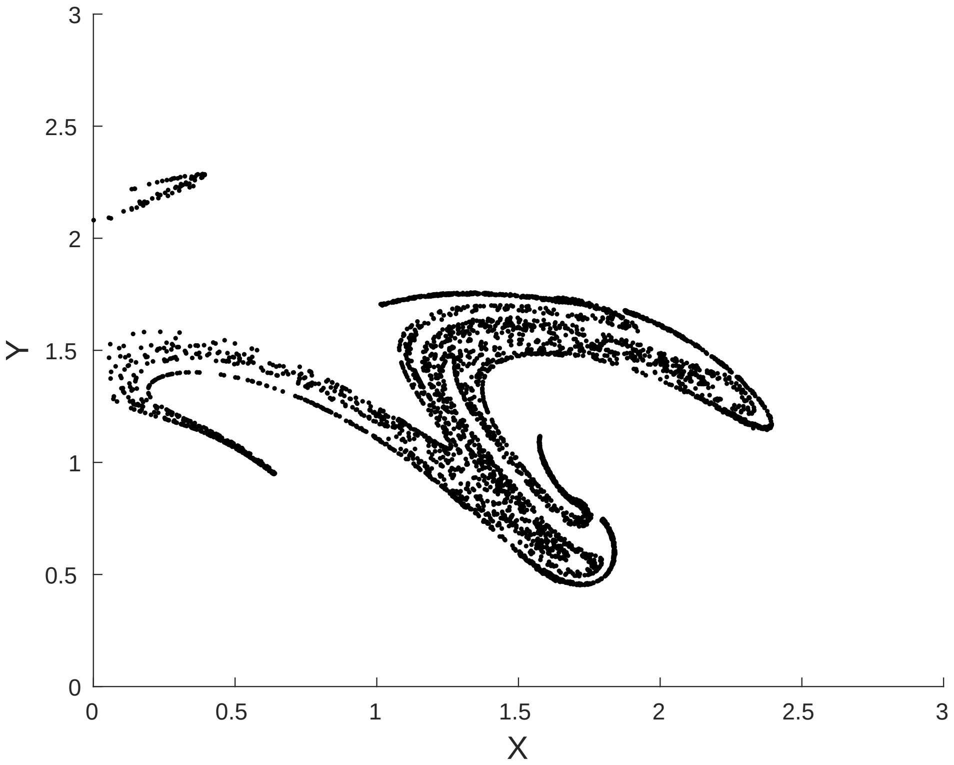


FIGURE 2.8: Poincaré section in $Z = 0$ of Lorenz 84 model.

for the meridional heat transport (see Fig.2.7 and Section 1.1). Thus, Eq.(2.40) describes the evolution of the westerlies, subject to the external forcing F_0, dampened both by the linear term $-aX$ and by nonlinear interaction with the eddies $-Y^2$ and $-Z^2$. This interaction amplifies the eddies through the terms XY and XZ in Eqs.(2.41)-(2.42). Furthermore, the eddies are affected by the westerlies through the terms $-bXZ$ and bXY. The constant b regulates the relative time scale between diplacements and amplifications. Eqs.(2.41)-(2.42), as Eq.(2.40), show a linear dissipation, whilst the symmetry between the two equations is broken by the external forcing G.

A commonly used method to visualize Lorenz 84 dynamics consists in plotting its Poincaré section in $Z = 0$ (Fig.2.8).

2.2.3 Lorenz 63

The Lorenz 63 model is probably the most iconic chaotic dynamical system (Saltzman, 1962; Lorenz, 1963; Ott, 1993) and was developed through a severe truncation of the partial differential equations describing the two-dimensional

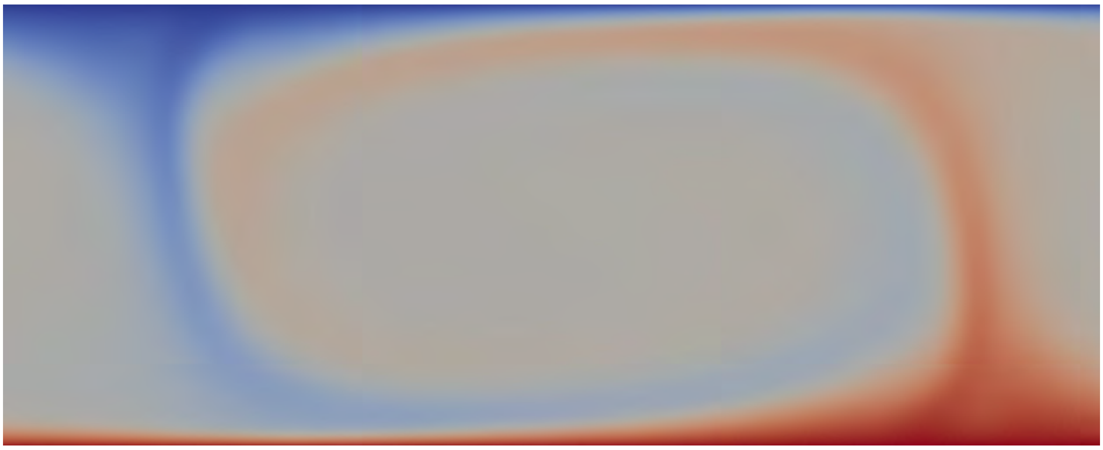


FIGURE 2.9: Convective roll in Rayleigh Bénard convection model. The flux of rising fluid in the right convective plume is balanced by a downward flux of fluid in the descending plume on the left.

Rayleigh-Benard convection (see e.g. Hilborn, 2000 for a complete, yet simple, derivation of the model) and describe the evolution of three modes corresponding to large scale motions and temperature modulations in the Rayleigh-Bénard problem.

Rayleigh Bénard convection is a simple type of natural convection where a heated plate displaced horizontally warms up the fluid immediately above it, forcing it to raise due to the difference in density with respect to the colder air above. In atmospheric sciences the hot plate can be seen as the Earth surface, which is heated by short waves radiation coming from the Sun and in turn heats the air by long waves radiation emission. In the example treated by Lorenz the upper boundary layer is another plate at a lower temperature with respect to the bottom one. This setting creates convective plumes of raising fluid which reach the upper plate pushing down the cold fluid, where it will be in turn heated by the ground. These masses of fluid, spinning around a horizontal axis, are called *convective rolls* (see Fig.2.9).

The three equations of the Lorenz 63 model are:

$$\frac{d\widetilde{x}}{dt} = \sigma(\widetilde{y} - \widetilde{x}), \tag{2.43}$$

$$\frac{d\widetilde{y}}{dt} = \rho\widetilde{x} - \widetilde{y} - \widetilde{x}\widetilde{z}, \tag{2.44}$$

$$\frac{d\widetilde{z}}{dt} = -\beta\widetilde{z} + \widetilde{x}\widetilde{y}, \tag{2.45}$$

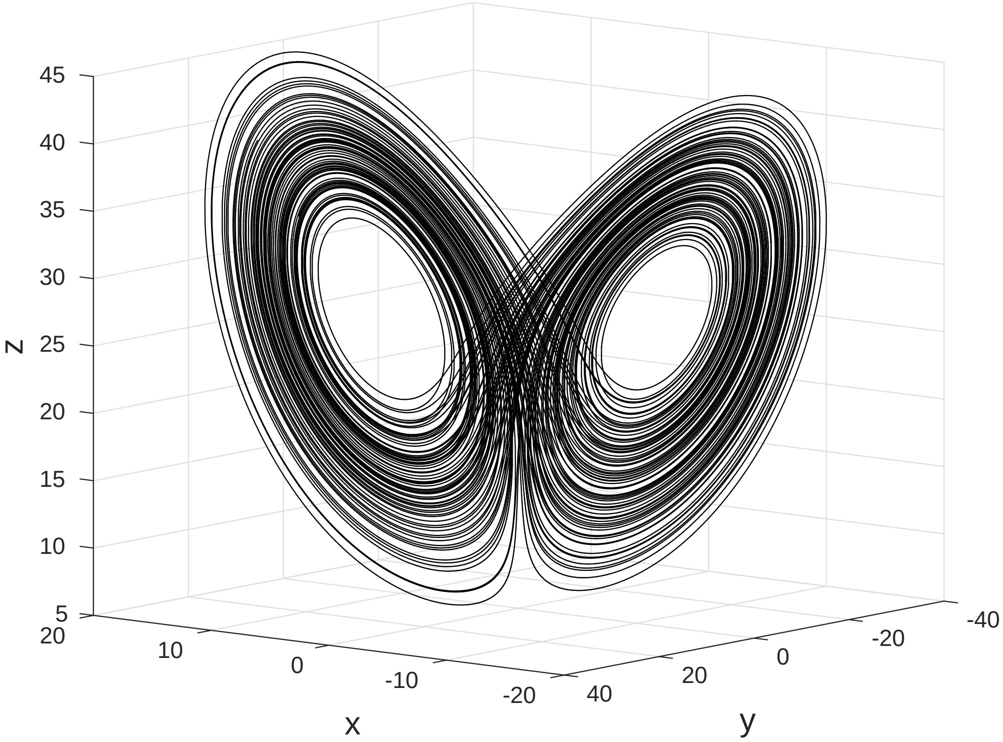


FIGURE 2.10: Lorenz attractor drawn by Lorenz 63 model.

where $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ are proportional, respectively, to the intensity of the convective motions, to the difference between temperatures of upward and downward fluid flows and to the difference of the temperature in the center of a convective cell with respect to a linear profile (since Eqs.(2.44)-(2.45) derive from thermal diffusion equation). The constants σ, ρ and β depend on kinematic viscosity, thermal conductivity, depth of the fluid, gravity acceleration, thermal expansion coefficient; specifically, σ is also known as the *Prandtl Number*.

The three dimensional attractor which, together with Bradbury's short story *A Sound of Thunder*, forged the idiomatic expression "butterfly effect" is the most emblematic representation of Lorenz 63 (Fig.2.10).

2.2.4 Lorenz 84 forced by Lorenz 63

It is possible to forge a multi scale model coupling Lorenz 84 and Lorenz 63 (Bódai, Károlyi, and Tél, 2011), where the latter acts unidirectionally as a forcing on the former, which represents the dynamics of interest. The dynamics of the two systems has a time scale separation given by the factor τ and can be written as follows:

$$\frac{dX}{dt} = -Y^2 - Z^2 - aX + a(F_0 + h\tilde{x}), \tag{2.46}$$

$$\frac{dY}{dt} = XY - bXZ - Y + G, \tag{2.47}$$

$$\frac{dZ}{dt} = XZ + bXY - Z, \tag{2.48}$$

$$\frac{d\tilde{x}}{dt} = \tau\sigma(\tilde{y} - \tilde{x}), \tag{2.49}$$

$$\frac{d\tilde{y}}{dt} = \tau(\rho\tilde{x} - \tilde{y} - \tilde{x}\tilde{z}), \tag{2.50}$$

$$\frac{d\tilde{z}}{dt} = \tau(-\beta\tilde{z} + \tilde{x}\tilde{y}). \tag{2.51}$$

It is important to underline that the coupling between the Lorenz 84 and the Lorenz 63 is unidirectional: the latter model affects the former acting as an external forcing, with no feedback acting the other way around.

In what follows, the parameters are chosen from the common setting: $a = 0.25$, $b = 4$, $\sigma = 10$, $\rho = 28$, $\beta = 8/3$; the two forcings are set as $F_0 = 8$ (corresponding to the so-called winter conditions) and $G = 1$. The parameter h is a modulation coefficient that defines the coupling strength and it is chosen as $h = 0.25$ in order to provide a stochastic forcing between two and four orders of magnitude smaller (on average) than the tendencies of the X variable.
The parameter τ defines the ratio between the internal time scale of the two systems: in case of $\tau > 1$, the Lorenz 63 provides a forcing that is typically on time scales shorter than those of the system of interest, while if $\tau < 1$ the forcings can be interpreted as a modulating factor of the dynamics of the Lorenz 84 model. In the first case, in particular, the Lorenz 63 can be interpreted as the forcing exerted by convective motions in the synoptic scale dynamics described by the Lorenz 84 model.

Fig.2.11 shows the effect of forcing the Lorenz 84 model with the Lorenz 63 model on the Poincaré section in $Z = 0$ (compare with Fig.2.8) in the case $\tau = 5$.

The numerical integration scheme used is a Runge-Kutta 4 with a time step of 0.005 (Bódai, Károlyi, and Tél, 2011).

In the tests related to these models, the standard Lorenz 84 will be called *uncoupled model*, whilst the Lorenz 84 subject to the coupling with the Lorenz 63 will take the denomination of *coupled model*.

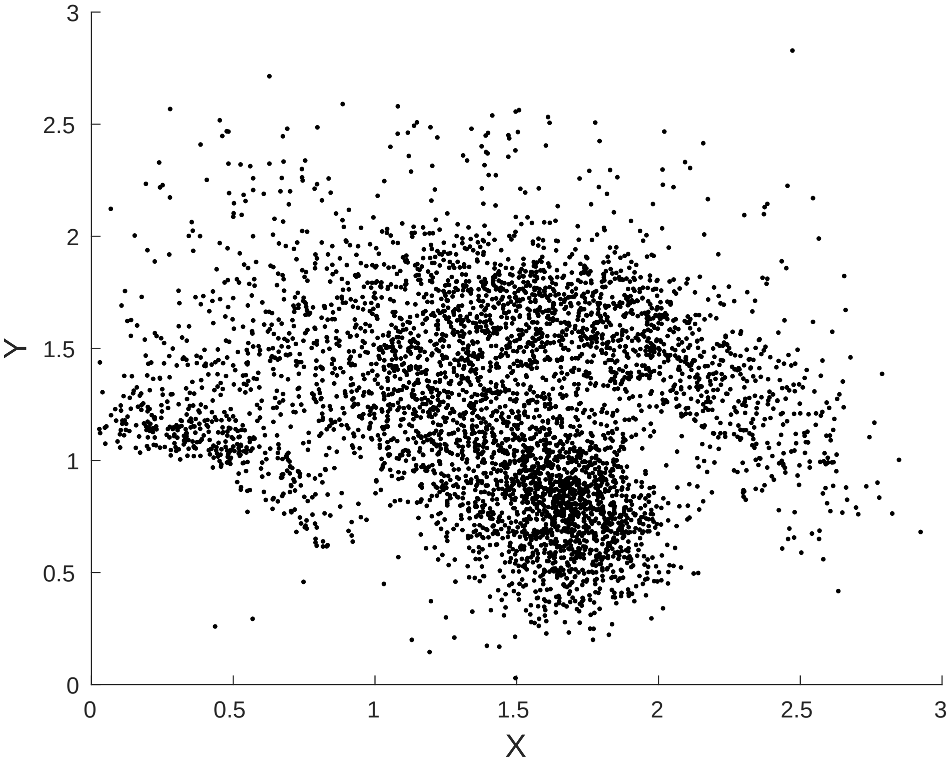


FIGURE 2.11: Poincaré section in $Z = 0$ of Lorenz 84 model coupled with Lorenz 63 model.

2.3 Wasserstein Distance

One of the aims of this dissertation is to assess how well a parametrization allows to reproduce the statistical properties of the full coupled system. At this regard, it seems relevant to quantify to what extent the projected invariant measure of the full coupled model on the variables of interest differs from the invariant measures of the surrogate models containing the parametrization. In order to evaluate how much such measures differ, their Wasserstein distance (Villani, 2009) is computed. Such a distance quantifies the minimum "effort" in morphing one measure into the other, and was originally introduced by Monge, 1781 to study problems of military relevance, and later improved by Kantorovich, 1942.

Starting from two distinct spatial distribution of points, described by the measures μ and ν, the optimal transport cost (Villani, 2009) is defined as the minimum cost to move the set of points corresponding to μ into the set of points corresponding to ν:

$$C(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \int c(x,y)d\pi(x,y), \tag{2.52}$$

where $c(x,y)$ is the cost for transporting one unit of mass from x to y and $\Pi(\mu,\nu)$ is the set of all joint probability measures whose marginals are μ and ν. The function $C(\mu,\nu)$ in Eq.(2.52) is called Kantorovich-Rubinstein distance. In the rest of the thesis, Wasserstein distance of order 2 will be considered:

$$W_2(\mu,\nu) = \left\{ \inf_{\pi\in\Pi(\mu,\nu)} \int [d(x,y)]^2 d\pi(x,y) \right\}^{\frac{1}{2}}. \tag{2.53}$$

The Wasserstein distance can be defined also in the case of two discrete distributions

$$\mu = \sum_{i=1}^{n} \mu_i \delta_{x_i}, \tag{2.54}$$

$$\nu = \sum_{i=1}^{n} \nu_i \delta_{y_i}, \tag{2.55}$$

where x_i and y_i represent the location of the different points, which mass is given, respectively, by μ_i and ν_i. Recalling the definition of Euclidean distance

$$d(\mu,\nu) = \left[\sum_{i=1}^{n} (x_i - y_i)^2 \right]^{\frac{1}{2}}, \tag{2.56}$$

the order 2 Wasserstein distance for discrete distributions can be constructed as follows:

$$W_2(\mu,\nu) = \left\{ \inf_{\gamma_{ij}} \sum_{i,j} \gamma_{ij} [d(x_i,y_j)]^2 \right\}^{\frac{1}{2}}. \tag{2.57}$$

where γ_{ij} is the fraction of mass transported from x_i to y_j.

This latter definition of Wasserstein distance has already been proven effective (Robin, Yiou, and Naveau, 2017) for providing a quantitative measurement of the difference between the snapshot attractors of the Lorenz 84 system in the instance of summer and winter forcings.

Hereby it is proposed to further assess the reliability of WL stochastic parametrization by studying the Wasserstein distance between the projected invariant measure of the original system on the first three variables (X, Y, Z) and the invariant measures obtained using the surrogate dynamics corresponding to the first and second order parametrizations. Nevertheless, since the numerical computations for optimal transport through linear programming theory are not cheap, a new approach is required. In order to accomplish it, a standard Ulam discretization (Ulam, 1964; Tantet, Lucarini, and Dijkstra, 2018) of the measure supported on the attractor is performed by coarse-graining on a set of cubes with constant sides across the phase space.

The coordinates of the cubes will then be equal to the location x_i, while the corresponding densities of the points are used to define γ_{ij}; finally, all the grid boxes containing no points at all are excluded by the computation and the distance is rescaled.

The calculations are performed using a modified version of the software for Matlab written by Gabriel Peyré and made available at `http://www.numerical-tours.com/matlab/optimaltransp_1_linprog/`, conveniently modified to include the subdivision of the phase space in cubes and the assignment of the corresponding density to those cubes.

Chapter 3

Results

3.1 Wouters-Lucarini's parametrization on Lorenz 96

3.1.1 Constructing the Parametrization

Here it is shown how to apply the parametrization introduced in Section 2.1 to the specific case of the Lorenz 96 model. The scale-adaptive parametrization shown here provides us with a great deal of flexibility and extremely parsimonious numerical costs. The uncoupled evolution equation for the Y variables (Eq.(2.7)) can be written in a universal form. In fact, it is easy to check that, operating on Lorenz 96 model the substitutions

$$\tau = ct \tag{3.1}$$

and

$$Z_{j,k} = bY_{j,k}, \tag{3.2}$$

the uncoupled evolution equation for the rescaled Y variables becomes:

$$\frac{dZ_{j,k}}{d\tau} = Z_{j+1,k}(Z_{j-1,k} - Z_{j+2,k}) - Z_{j,k} + F_2. \tag{3.3}$$

Therefore, for all values of h, b, and c it is feasible to construct the parametrization just by resorting to the invariant measure of Eq.(3.3) and adopting the suitable rescaling. Note that in the case of this specific system it is possible to rescale also the size of the Y variable and achieve a higher degree of flexibility than in the general case discussed above. This emphasizes the scale-adaptivity of the approach proposed here, and makes sure that only modest computation effort is needed to deal with the problem of parametrization.

Note that, compared to the general case of multiscale system discussed before, in this case the additional problem that changing the value of $c = \gamma$ leads also to an increase in the value of ϵ arises, so that large values of c might break the weak coupling hypothesis. The problem can be circumvented by increasing at the same time the value of b or considering smaller values of h.

The first order term in the parametrization is recovered using ergodicity and averaging $D(X)$ in Eq.(2.22). By symmetry, the coupling is the same for all the X variables:

$$D_k(X_k) = D(X) = D = -\frac{1}{b} \lim_{T\to\infty} \frac{1}{T} \int_0^T \sum_{j=1}^{J} Z_{j,k}(\tau) d\tau, \tag{3.4}$$

where $k = 1, \ldots, K$ and the average is performed by integrating Eq.(3.3).
The value of this term is $\frac{-20.12}{b}$ for all k's; choosing $b = 10$ leads to $D_k(X_k) = -2.012$. Therefore, the coupling between fast and slow scales leads on the average to a reduction in the effective forcing applied to the slow variables. In other terms, this indicates a net energy flux from slow to fast variables. Despite the simplicity of the model considered here and of the coupling between the X and Y variables, this corresponds to the effect of introducing eddy viscosity in more complex fluid dynamical models.

The k^{th} component of the stochastic term $S\{X\}$ in Eq.(2.23) is constructed as an additive noise $\sigma(t)$ featuring the following lagged covariance:

$$R_k(\tau) = R_k(ct) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \Big(-\sum_{j=1}^{J} \frac{Z_{j,k}(\tau_1)}{b} - D\Big)\Big(-\sum_{j=1}^{J} \frac{Z_{j,k}(\tau+\tau_1)}{b} - D\Big) d\tau_1, \tag{3.5}$$

where the evolution of the Z variables is given by Eq.(3.3) and the covariance is reported in Fig.3.1. The production of time series of σ to be used for the parametrized simulation is achievable either from properly resampling time series of the fluctuation term $-\sum_{j=1}^{J} \frac{Z_{j,k}}{b} - D$ or by reproducing them using simple stochastic models like those belonging to the $AR(n)$ family. This latter route was chosen for this application, taking advantage of the software package ARFIT (Neumaier and Schneider, 2001; Schneider and Neumaier, 2001). This term describes the backscatter of energy from the small towards the large scales.

Note that, as the argument of the function is ct, in the limit of $c \to \infty$ the autocovariance tends to zero for all $t > 0$, because the function R tends to zero for large values of its argument, while one has for all values of c that $R(0)$ is finite. As a result, one obtains as limit a white noise of vanishing amplitude for any fixed value of ϵ.

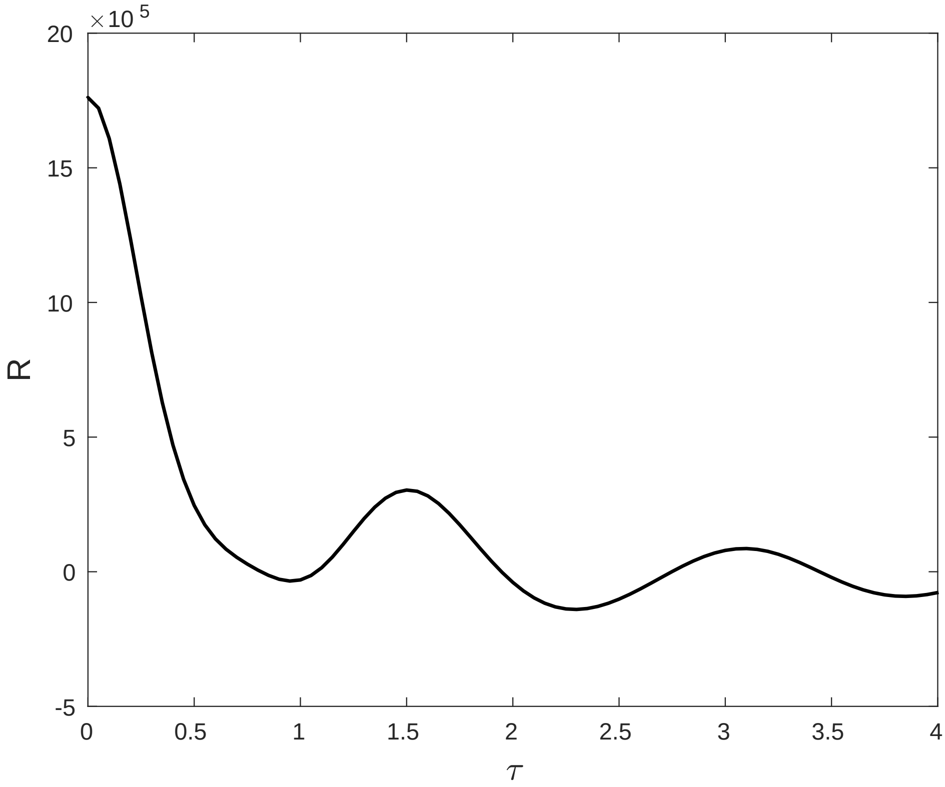


FIGURE 3.1: Time lagged autocovariance of the noise term $\sigma(t)$ with $b = 10$ and $h = 1$.

It is useful to provide an explicit expression of the k^{th} component of the non-markovian term $M\{X\}$ given in Eq.(2.25). The memory kernel $h_k(\tau, \tilde{X}_k)$ (where $\tau = ct$) can be expressed as follows:

$$h_k(\tau_1, \tilde{X}_k) = -\frac{1}{b}\tilde{X}_k H(\tau_1), \tag{3.6}$$

where

$$H(\tau_1) = lim_{\Omega\to\infty}\frac{1}{\Omega}\int_0^{\Omega}\sum_{j=1}^{J}\frac{\partial}{\partial Z_{j,k}(\omega)}Z_{j,k}(\tau_1+\omega)d\omega. \tag{3.7}$$

The factor H on the right hand side of Eq.(3.6) is plotted in Fig.3.2; this clarifies that the kernel weighs less states of the X variables with larger time separation, as expected. Increasing the value of c leads to a compression of the time axis by a factor of c. Since $H(\tau) \to 0$ in the limit of $c \to \infty$, h vanishes for all values of $t > 0$. Hence, memory effects disappear in this limit.

For a given value of ϵ, the larger the value of c, the more dominant is the contribution to the parametrization coming from the deterministic first order term.

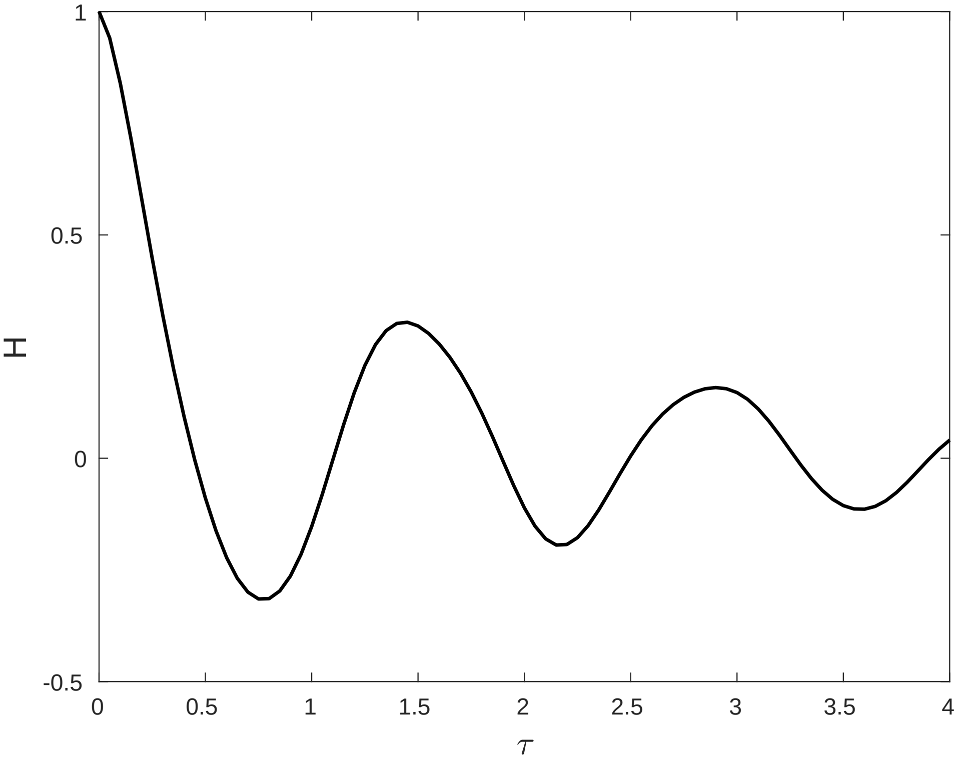


FIGURE 3.2: Memory effects as measured by the factor H, see Eq. (3.7), with $b = 10$ and $h = 1$.

Note that Abramov, 2016 addresses the problem of parametrizing a modified version of the Lorenz 96 system similar to the one presented here by a modified version of the homogenization method. The derived parametrization is different from what obtained here as the homogenization method assumes infinite separation of scales between the fast and slow variables. Abramov obtains a stochastic contributions that is always white (yet its variance depends on the time scale separation), and an extra deterministic linear term that, from construction, might point at a surrogate way to implicitly deal with memory effects.

3.1.2 Performance on Lorenz 96

In this section a series of statistical tests is performed in order to check the skill of the WL parametrization applied on Lorenz 96 model. In what follows, Eqs.(2.34)-(2.35) indicate the *coupled model*. The *uncoupled model* is, instead, given by the evolution equations of Eq.(2.33) or, equivalently, Eq.(2.34) where

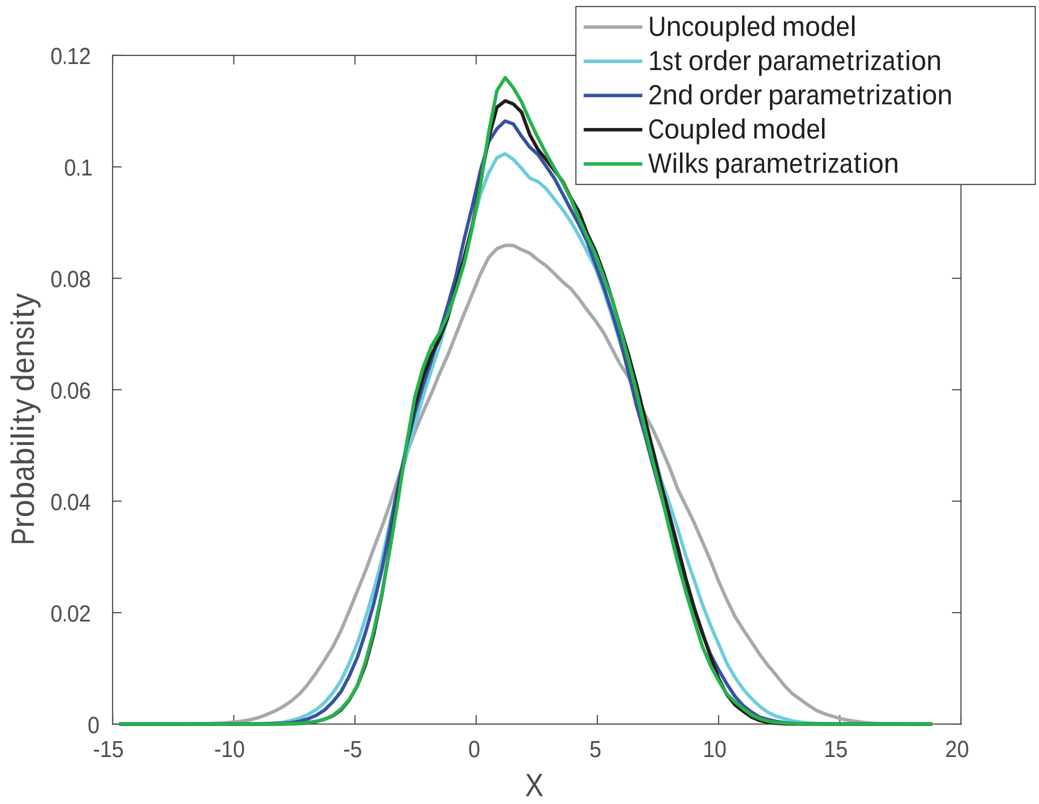


FIGURE 3.3: Probability density of the X variable for the different models considered.

the last term is excluded (or h is set to 0). The model with *first order parametrization* is obtained by inserting expression (3.4) in Eq.(2.13) and disregarding the other terms. The model with *second order parametrization* is obtained by inserting in Eq.(2.13) both the first and second order terms. The skill of the parametrization is tested reproducing the statistical properties of the coupled model and comparing it to the performance of the parametrization constructed according to the method proposed by Wilks, 2005. For this test the standard values of the parameters $c = 10$, $b = 10$, $h = 1$ are chosen; every other possible choice for these factors can be covered through a proper rescaling of the values for D, S and M.

The first tests consist in checking the ability of the parametrizations in reproducing the probability density function of the variable X_k, the lagged temporal correlation $Corr(t) = \langle X_k X_k(t) \rangle$, and the spatial correlations at zero time lag $Sp(l) = \langle X_k X_{k+l} \rangle$.

Fig.3.3 shows the probability density of the X_k variables for all considered models. It is clear how the second order parametrization offers a better result with respect to the first order, which is in turn a clear improvement of the

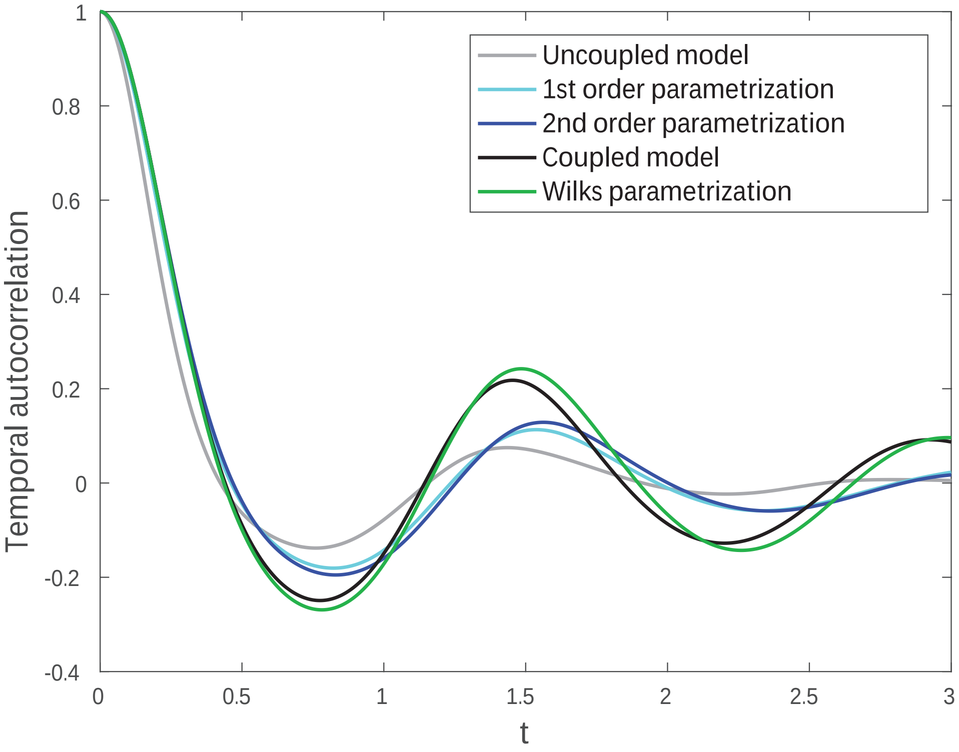


FIGURE 3.4: Temporal autocorrelation of the X variable for the different models considered.

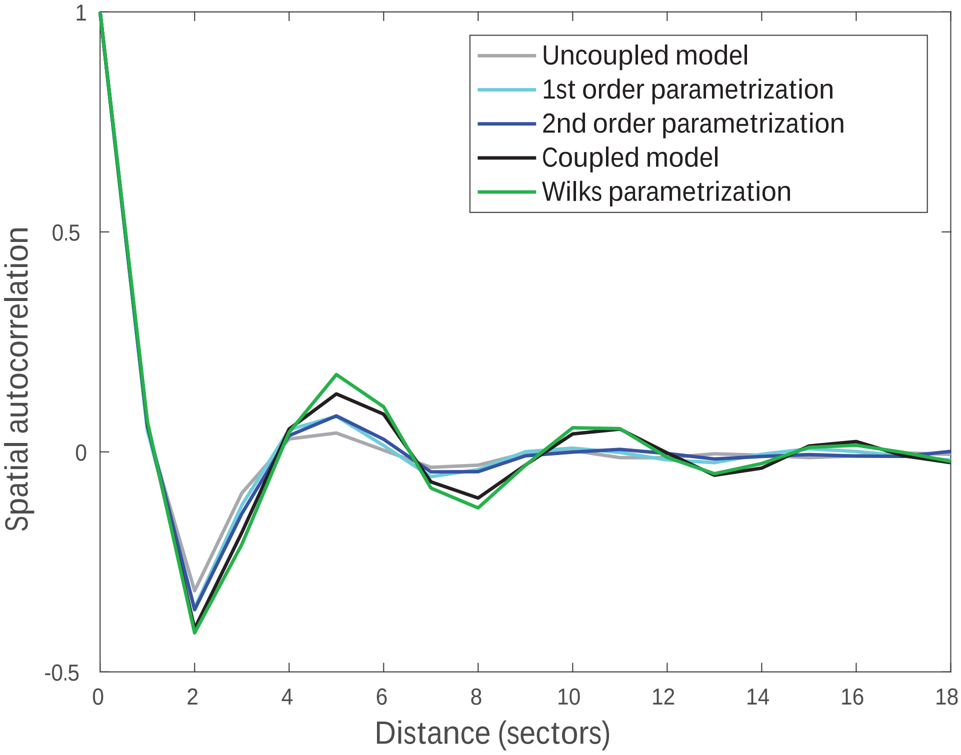


FIGURE 3.5: Spatial autocorrelation of the X variable for the different models considered.

basic uncoupled system. Both Wilks's approach and the WL parametrization provide rather good approximations of comparable quality for the distribution of the X variable of the original system.

Considering normalized second order properties for the X variables, the first look at the lagged time autocorrelation (see Fig.3.4) shows how higher order parametrizations lead to a better agreement with the coupled model, even if the improvement in the skill is most evident for small time lags. Nevertheless, the Wilks method provides very good results also for lags larger than 0.4 time units.

Fig.3.5 shows the performance of the parametrization in simulating the spatial correlation of the X_k variable. Considering higher order approximations in the parametrizations does not imply a substantial improvement of the results, even if the first and second order parametrizations lead to an improvement with respect to the uncoupled case. In this case Wilks's parametrization follows closely the full coupled model and overperforms the parametrizations constructed according to the method discussed here.

The analysis of the second and higher order moments is shown in the next section.

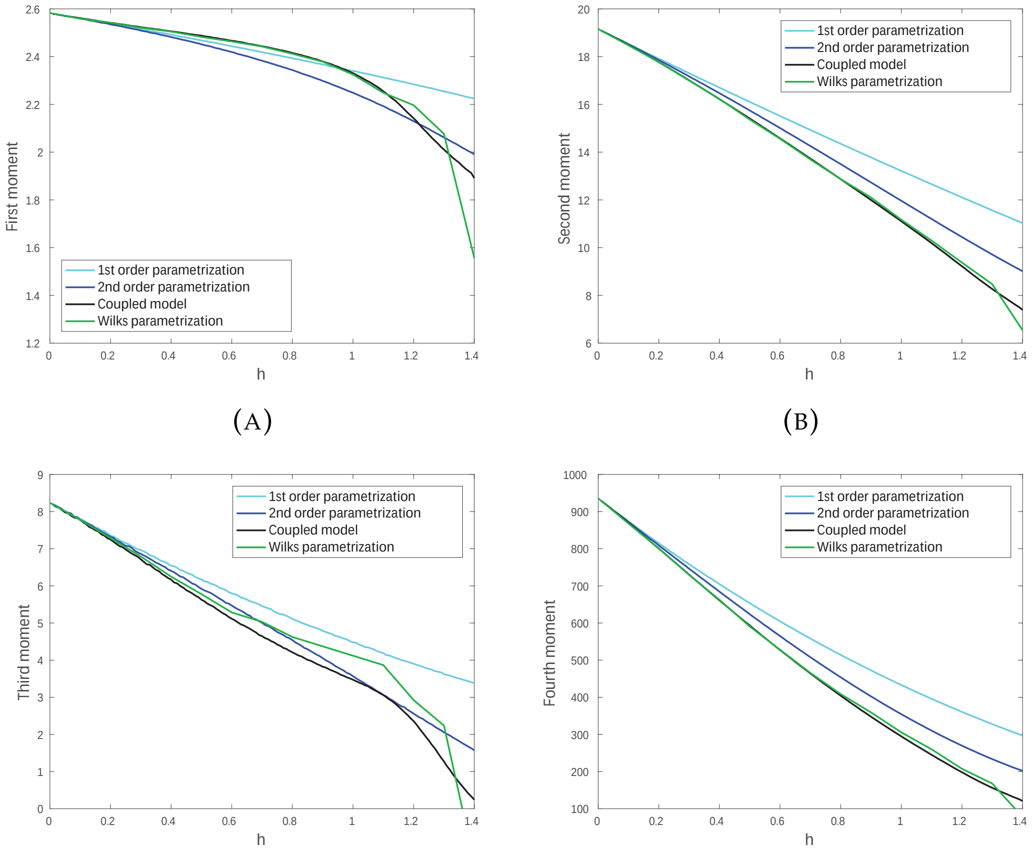
Sensitivity to the Strength of the Coupling

As Wouters-Lucarini expansion is based on assuming the presence of weak coupling between the slow and the fast variables, it is crucial to test its performance as the value of the coupling strength $h = \epsilon\frac{b}{c}$ (see Eqs.(2.34) and (2.35)) changes with respect to its standard value of 1. Note that in the case treated here $b = c = 10$ are held fixed, so that $h = \epsilon$ is changed in what follows. The focus will be on the first moment and on the second, third and fourth central moments of the variable X_k.

Fig.3.6a) shows that all parametrizations perform rather well in terms of representing the first moment of X_k for all considered values of $h < 1.4$. Larger values of h lead to a qualitative change in the properties of the system and fall outside the range of interest.

Surprisingly, the first order parametrization constructed using the WL method overperform the second order model for $h \lessapprox 1$, which hints at the importance (at least in this case) of possibly developing a theory for the third order scheme, beyond the WL parametrization.

Figs.3.6b)-d) portray the performance of the parametrizations in reproducing the values of the second, third and fourth centered moments, respectively.



(A) (B)

(C) (D)

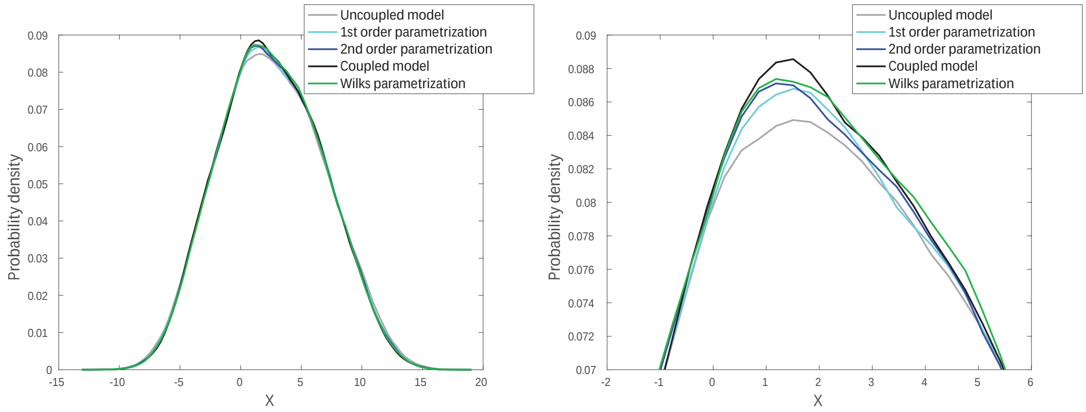
FIGURE 3.6: a) First moment; b) Second centered moment; c) Third centered moment; d) Fourth centered moment as functions of the coupling strength for the different models considered.

It is consistently found that, while all methods are quite successful, the Wilks parametrization provides the best results, with the second order model constructed with the WL method coming close second.

It is important to underline that the Wilks parametrization needs to be constructed from scratch for each different value of h (as well as of b and c). This marks a fundamental difference with the parametrization tested in this study, where just a linearly rescaling of the first order term and a quadratical rescaling of the two second order terms are needed. Another problem shown by Wilks's method is the lack of stability in case of high values of h; as a matter of fact, in order to obtain results for $h > 1.2$ the time step in the numerical integration had to be drastically reduced, thus having a much higher computational cost.

TABLE 3.1: Values of the constants determining the parametrization according to the Wilks method for various values of the model's parameter c.

c	b_0	b_1	b_2	b_3	b_4	σ_e	ϕ
1	1.694×10^{-1}	1.619×10^{-2}	7.507×10^{-4}	-7.868×10^{-5}	3.06×10^{-7}	1.04×10^{-1}	0.9995
5	9.122×10^{-1}	9.419×10^{-2}	-5.644×10^{-3}	4.025×10^{-5}	1.852×10^{-5}	4.659×10^{-1}	0.9996
10	1.81	1.467×10^{-1}	-1.357×10^{-3}	1.446×10^{-3}	-1.313×10^{-4}	8.965×10^{-1}	0.9997
20	3.721	4.861×10^{-1}	2.752×10^{-2}	-2.38×10^{-2}	2.427×10^{-3}	1.47	0.9997
100	1.317×10^{1}	1.059	2.969×10^{-1}	6.534×10^{-2}	3.421×10^{-3}	9.905×10^{-2}	0.9998



(A) (B)

FIGURE 3.7: a) Probability density of the X variable in the case of $c = 1$, $b = 10$ and $h = 1$. b) Zoom on the peak of the distribution.

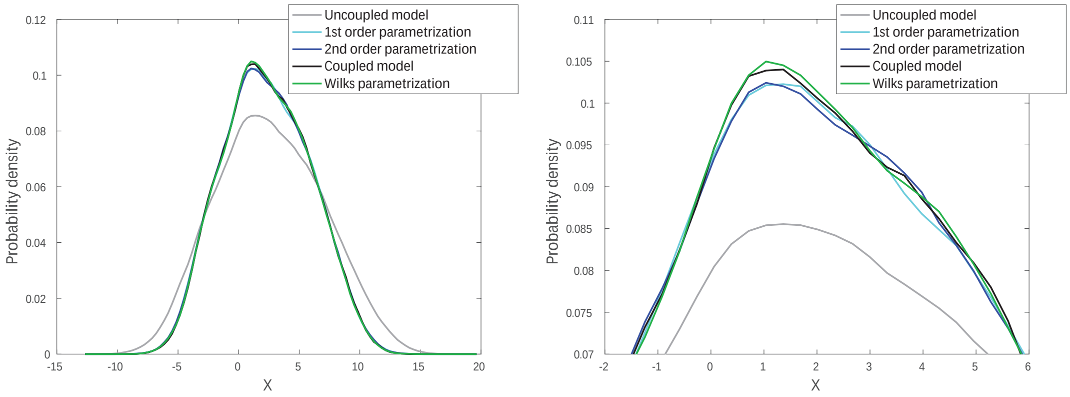
Scale adaptivity

The most relevant advantage of the WL approach proposed here is that it allows one to construct general parametrizations by suitably rescaling the three terms - deterministic, stochastic and non-markovian - after having estimated them through a single numerical simulation.

The method proposed by Wilks is more precise for each given choice of the system's parameters but lacks such a flexibility, which might be of crucial relevance when trying to develop self-adaptive parametrizations. The coefficients appearing in the Wilks parametrization (see Table 3.1) cannot be readily predicted with suitable expressions.

Using the general results for the first and second order terms of the WL parametrization and adopting the suitable rescaling for the amplitude and the time axis discussed in the previous section, it is possible to explore an infinite range of scenarios for the values of b, c, and h. Some examples are presented below.

In Fig.3.7 it is shown that the probability densities of X_k obtained through the different parametrizations in the case $c = 1$ are in good agreement with



(A) (B)

FIGURE 3.8: a) Probability density of the X variable in the case of $c = 100$, $b = 10$ and $h = 0.1$. b) Zoom on the peak of the distribution.

what shown by the coupled model. Note that choosing $c = 1$ implies also assuming that there is *no* scale separation between the X and the Y variables. In fact, as discussed before, the WL method does not actually require a scale separation between the parametrized and the resolved systems.

Since in the case treated here the coupling should not be too strong compared to the unperturbed vector flow (this is the condition under which it is possible to use the WL method), as said before, increasing the value of c can be problematic unless the value of h is reduced accordingly (or the value of b is increased). Fig.3.8 shows the probability density function of the X variable in the case $c = 100$, $b = 10$, $h = 0.1$, with a much stronger coupling than the previous case of Fig.3.7. In this case, it is clear that considering a parametrization is crucial for reproducing satisfactorily the statistics of the X variable, and the first order parametrization is already rather successful. Note that as c becomes larger, the memory term has a less and less relevant role and the stochastic contributions is rather similar to a white noise forcing.

As last test (Fig.3.9) the rescaling of the model is stressed applying the transformation to all the parameters at the same time, shifting from the $c = 10$, $b = 10$, $h = 1$ to the $c = 5$, $b = 8$, $h = 1.1$ scenario.

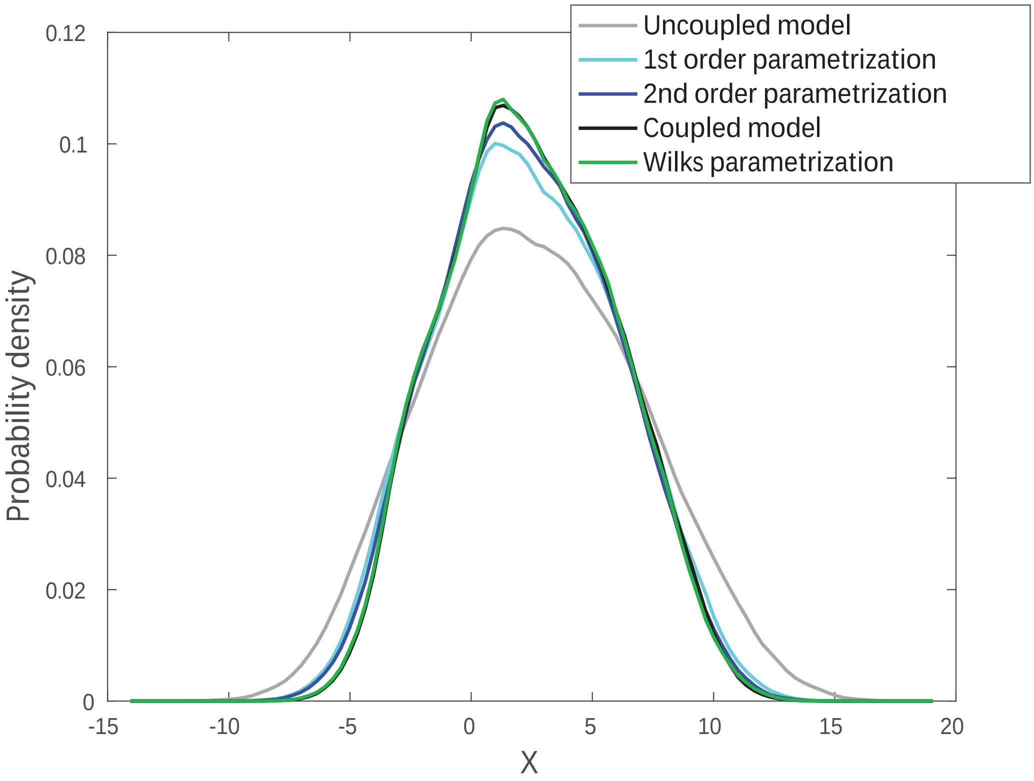


FIGURE 3.9: Probability density of the X variable in the case of $c = 5$, $b = 8$ and $h = 1.1$.

Forcing the fast scale dynamics

As stated in Section 2.1, the WL methodology can be applied also in the case where no forcing is acting on the fast scales. This procedure can be repeated also in the case, like the one analyzed here, where the requirement of shifting the background state is not strictly necessary, and the natural definition of the uncoupled dynamics of the Y variables given in Eq.(2.7) can be used. This hypothesis is tested here by using the framework given in Eqs.(2.28)-(2.29) and choosing the standard values for the system's parameters and $G = \rho_{0,X}(\Psi_Y(X)) = 2.57$. Figs.3.10 and 3.11 show that, at the second order, the results obtained are almost undistinguishable with respect to what shown in Fig.3.3 for probability density and Fig.3.4 for the time autocorrelation using $F_2 = 6$ and $\frac{c}{b}F_2 + G = 8.57$ as forcing of the uncoupled Y equation.

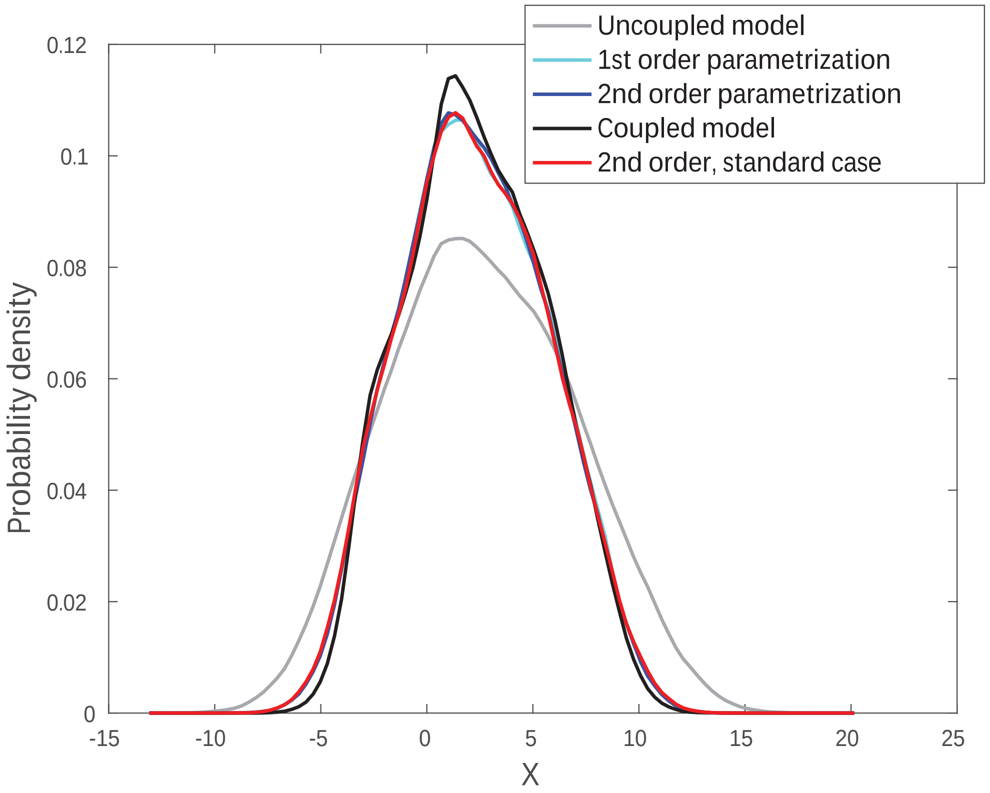


FIGURE 3.10: Probability density of the X variable calculated adding G to uncoupled Y equation. The standard case is the one shown in Fig.3.3.

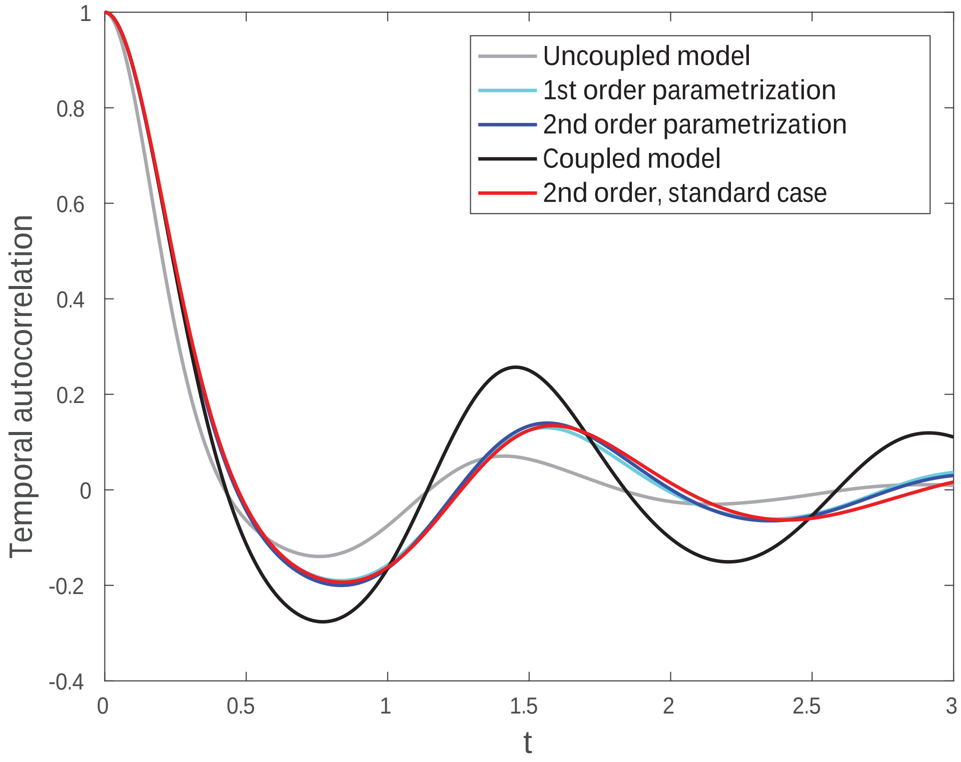


FIGURE 3.11: Time autocorrelation of the X variable calculated adding G to uncoupled Y equation. The standard case is the one shown in Fig.3.4.

3.2 Wouters-Lucarini's parametrization on Lorenz 84 forced by Lorenz 63

3.2.1 Constructing the Parametrization

For the sake of readability, Eqs.(2.1)-(2.2) are rewritten performing the substitution $(\mathbf{X}, \mathbf{Y}) \rightarrow (\mathbf{K}, \mathbf{J})$:

$$\frac{d\mathbf{K}}{dt} = \mathbf{F_K}(\mathbf{K}) + \epsilon \mathbf{\Psi_K}(\mathbf{K}, \mathbf{J}), \tag{3.8}$$

$$\frac{d\mathbf{J}}{dt} = \mathbf{F_J}(\mathbf{J}) + \epsilon \mathbf{\Psi_J}(\mathbf{K}, \mathbf{J}), \tag{3.9}$$

In the case of Lorenz 84 model forced by Lorenz 63 model, the coupling strength ϵ, shown in Eqs.(3.8)-(3.9) and in Eq.(2.13), assumes the value $\epsilon = ah$, while the couplings are, with respect to the vector (X, Y, Z) in Lorenz 84 phase space, $\mathbf{\Psi_K}(\mathbf{K}, \mathbf{J}) = \mathbf{\Psi_K}(\mathbf{J}) = (\tilde{x}, 0, 0)$ and $\mathbf{\Psi_J}(\mathbf{K}, \mathbf{J}) = \mathbf{\Psi_J}(\mathbf{K}) = (0, 0, 0)$. Note that this is another case of independent coupling - i.e. a coupling that depends only on the variable of the other equation -, for which the application of the methodology is simpler than the dependent coupling case (Wouters and Lucarini, 2012).

The deterministic term $\mathbf{D}$ in Eq.(2.13) is a measure of the average impact of the coupling on the $\mathbf{K}$ dynamics and can be written as:

$$\begin{aligned} \mathbf{D}(\mathbf{K}) = \rho_{0,\mathbf{J}}(\mathbf{\Psi_K}(\mathbf{J})) &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbf{\Psi_K}(\mathbf{J}) d\tau = \rho_{0,\mathbf{J}}((\tilde{x}), 0, 0) \\ &= \lim_{T\to\infty} \frac{1}{T} \int_0^T (\tilde{x}(\tau), 0, 0) d\tau = (D, 0, 0), \end{aligned} \tag{3.10}$$

where $\rho_{0,\mathbf{x}}(A)$ ($\mathbf{x} = \mathbf{K}, \mathbf{J}$) is the expectation value of A computed according to the invariant measure given by the uncoupled dynamics of the $\tilde{x}$ variables and ergodic averaging has been implemented. The expression of the coupling, requested to compute the ensemble average as time average on the ergodic measure of $\tilde{x}$, is given by Eq.(2.46). Since the measure of Lorenz 63 is symmetric for $\tilde{x} \rightarrow -\tilde{x}$, one could think of choosing $\mathbf{D}(\mathbf{K}) = (0, 0, 0)$. Nevertheless, this is the limit for a run of infinite time length, while in a numerical simulation a finite number of steps is required - in this case 146000, ten years in Lorenz models. Therefore, it seems appropriate to calculate $\mathbf{D}$ using the time series given by the uncoupled Lorenz 63 and Eq.(3.10), as done for the second order of the parametrization.

Since the coupling shown in Eq.(2.46) depends only on one of the variables (in this case the $\widetilde{x}$) of the system to parametrize, the stochastic term can be written as

$$\mathbf{S}\{\mathbf{K}\} = (\omega(t), 0, 0), \tag{3.11}$$

where the properties of $\omega(t)$, a stochastic noise, are defined by its correlation $R(t)$ and its time average $\langle \omega(t) \rangle$:

$$\begin{aligned}
\mathbf{R}(\mathbf{t}) =& \langle (\omega(0), 0, 0), (\omega(t), 0, 0) \rangle \\
&= \rho_{0,\mathbf{J}}((\mathbf{\Psi}_{\mathbf{K}}(\mathbf{J}) - \mathbf{D}(\mathbf{K}))(\mathbf{\Psi}_{\mathbf{K}}(\mathbf{f}^t(\mathbf{J})) - \mathbf{D}(\mathbf{K}))), \\
&= \rho_{0,\mathbf{J}}(((\widetilde{x}(0), 0, 0) - (D, 0, 0))((\widetilde{x}(t), 0, 0) - (D, 0, 0))), \qquad (3.12) \\
\langle \omega(t) \rangle &= 0. \qquad (3.13)
\end{aligned}$$

As discussed in Wouters and Lucarini, 2012; Wouters and Lucarini, 2013, for more complex couplings the stochastic term assumes the form of a multiplicative noise. To construct time series of noise with the desired properties - defined by Eq.(3.12) - the software package ARFIT (Neumaier and Schneider, 2001; Schneider and Neumaier, 2001) has been used.

The last term in Eq.(2.13) is the non-markovian contribution to the parametrization and can be written as follows:

$$\mathbf{M}\{\mathbf{K}\} = \int_0^\infty \mathbf{h}(t_2, \mathbf{K}(t - t_2)) dt_2, \tag{3.14}$$

where

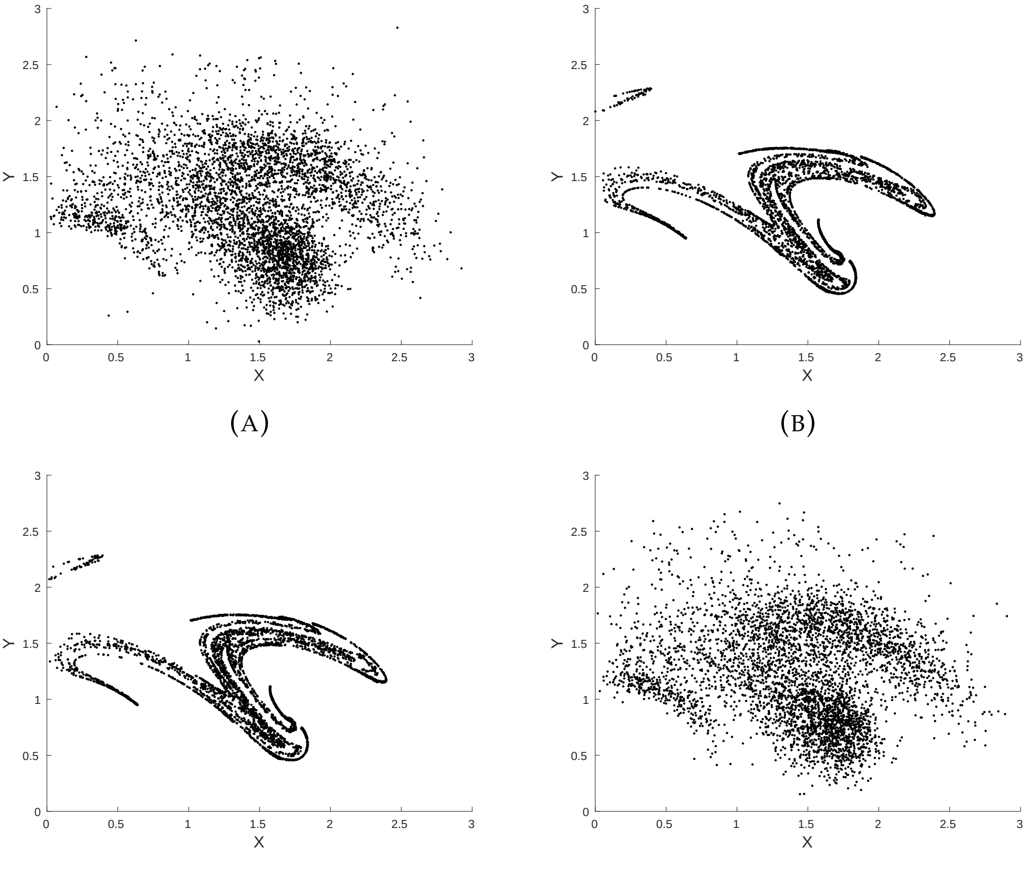
$$\begin{aligned}
\mathbf{h}(t_2, \mathbf{K}) =& \mathbf{\Psi}_{\mathbf{J}}(\mathbf{K}) \rho_{0,\mathbf{J}}(\partial_{\mathbf{J}} \mathbf{\Psi}_{\mathbf{K}}(\mathbf{f}^{t_2}(\mathbf{J}))) \\
=& (0, 0, 0) \cdot \rho_{0,\mathbf{J}}(\partial_{\mathbf{J}}(\widetilde{x}(f^{t_2}(\widetilde{x}, 0, 0)), 0, 0)). \qquad (3.15)
\end{aligned}$$

As discussed in Section 2.2.4, the evolution of the variables of the Lorenz 63 model - see Eqs.(2.49)-(2.51) - are independent of the state of the variables corresponding to the Lorenz 84 model. As a result, the first factor on the r.h.s. of Eq.(3.15) vanishes, so that the derived parametrization is fully markovian.

After the implementation of Wouters-Lucarini's procedure, Eq.(2.46) will be parametrized as

$$\frac{dX}{dt} = -Y^2 - Z^2 - aX + a[F_0 + h(D + S)]; \tag{3.16}$$

Eq.(3.16), together with Eqs.(2.47)-(2.48), will be henceforth indicated as the



(A) (B)

(C) (D)

FIGURE 3.12: Poincaré section in $Z = 0$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization.

system constructed with *second order parametrization*, whilst the same equations without the stochastic term (therefore comprehending the first order, deterministic term only), namely

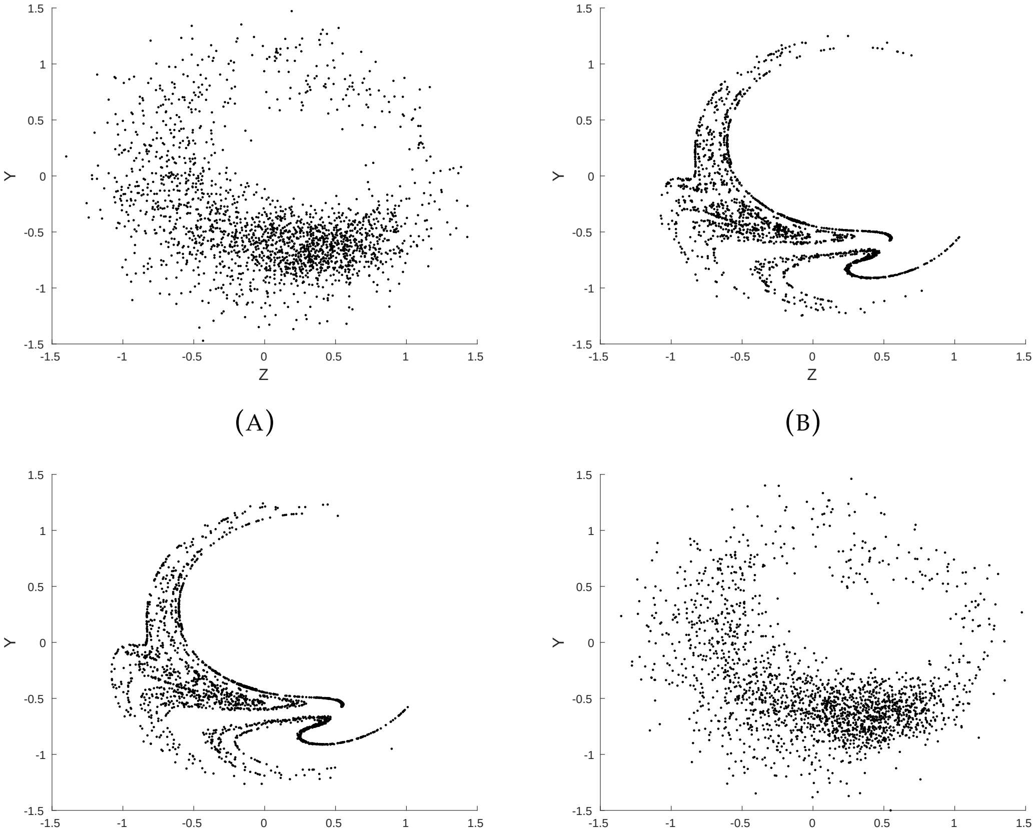
$$\frac{dX}{dt} = -Y^2 - Z^2 - aX + a[F_0 + hD], \tag{3.17}$$

will be called *first order parametrization*.

3.2.2 Performance on Lorenz 84 forced by Lorenz 63

Parametrizing the fast system

In the first part of this section the parametrization is applied to the Lorenz 84 model forced by a Lorenz 63 model with a time scale separation $\tau = 5$.



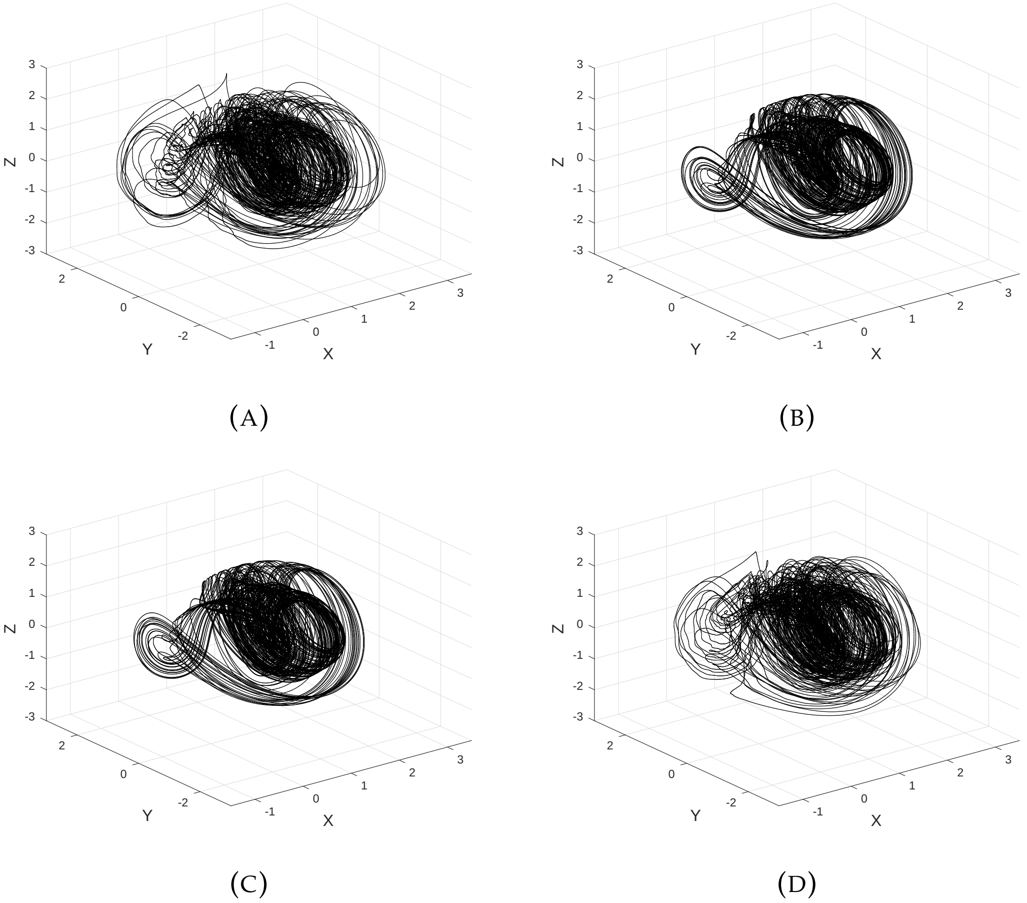
(A) (B)

(C) (D)

FIGURE 3.13: Poincaré section in $X = 1$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization.

Therefore, Lorenz 84 and Lorenz 63 are seen as, respectively, the slow and the fast dynamical systems.

A qualitative overview of the performance of the parametrization is given by investigating a few Poincaré sections, which provide a convenient and widely used method to visualize the dynamics of a system in a two-dimensional plot (Eckmann and Ruelle, 1985; Ott, 1993); typically, the plane chosen for the section of Lorenz 84 is $Z = 0$. Fig.3.12a) shows the Poincaré section at $Z = 0$ of the variables X, Y of the coupled model given in Eqs.(2.46)-(2.51). Panels b) of the same figure show the Poincaré section of the Lorenz 84 model obtained by removing the coupling with the Lorenz 63 model. Finally, Panels c) and d) show the Poincaré sections of the modified Lorenz 84 models obtained by adding the first and second order parametrizations, respectively. Visual inspection suggests that the second order parametrization does a good job in reproducing the properties of the full coupled model.



(A) (B) (C) (D)

FIGURE 3.14: $3D$ view of the attractor of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization.

Metaphorically, here WL parametrization aims at describing as accurately as possible the impact of "convection" on the "westerlies". It is insightful to look at how it affects the properties of the two variables - Y and Z - that are not directly impacted by it. This amounts to looking at the impact of the parametrization of "convection" on the "large scale planetary waves" and, consequently, on the "large scale heat transport". Therefore, a closer look into $X = constant$ Poincaré section allows to highlight the properties of Y and Z. The four panels in Fig.3.13 are structured as in Fig.3.12 and depict the Poncaré section of $X = 1$. Also in this case the second order parametrization provides a far better match to the coupled model, featuring a remarkable ability in reproducing the main features of the pattern of points.

In order to provide further qualitative evidence of these results, the four panels of Fig.3.14 show the trajectories in the phase space of the X, Y, and Z variables for the four considered models. For the sake of clarity, the plots are created using just 5 years (365 time units). In the case of the coupled model the attractor spans over more extreme values of the variables and the second order parametrization successfully imitates this feature, while the simple deterministic correction, once again, is completely inadequate.

Further to the qualitative inspection, quantitative comparisons to support this study have been performed. All the remaining simulations in this section are run for 100 years (7300 time units) with a time step of 0.005; thus, each attractor is constructed with 1460000 points. It has been tested that the results presented below are virtually unchanged when considering a smaller time step of 0.001.

The first look investigates the probability densities (PDFs) of the variables X, Y and Z, which describe, loosely speaking, our climate. Fig.3.15 shows the PDF of the X variable for the four considered models. As expected, the second order parametrization allows for reconstructing with great accuracy the statistics of the coupled model. The bimodality of the uncoupled Lorenz 84 model is reproduced by the model featuring the first order parametrization, while the second order model predicts accurately the unimodal distribution shown by the coupled model. The PDFs for Y and Z variables are shown in Figs.3.16-3.17, respectively. Also here, where the external forcing does not modify the bimodality of the distributions found in the uncoupled case, WL parametrization leads to a very good approximation of the properties of the coupled model. In particular, the tails of the distributions are represented with a high level of precision, making possible to seemingly reproduce with good accuracy the extreme values of the variables. Note that, since the WL parametrization is constructed to have skill for all observables, it is not so surprising that it can perform well also far away from the bulk of the statistics (see discussion in Lucarini et al., 2014).

Aside from the analysis of the PDF, a further statistical investigation can be provided by looking into the numerical results provided by first moments of the variables and their uncertainty, which is computed as the standard deviation derived from the analysis of an ensemble of runs. Due to the robustness of the results, ten runs were enough. The results for the statistics of the first two moments are reported in Table 3.2: all the quantities inspected clearly show that the second order parametrization allows for reproducing very accurately the moments statistics of the coupled model.

If the considered PDFs depart strongly from uni-modality, the analysis of the first moments can be of little utility, and it becomes hard to have a thorough

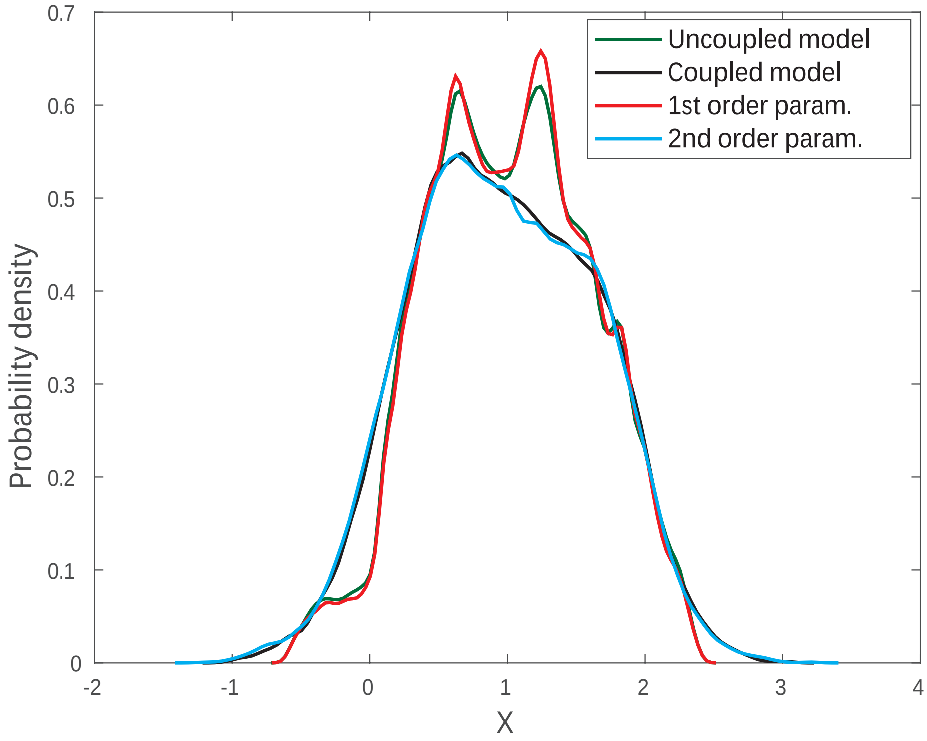


FIGURE 3.15: Probability density of the X variable.

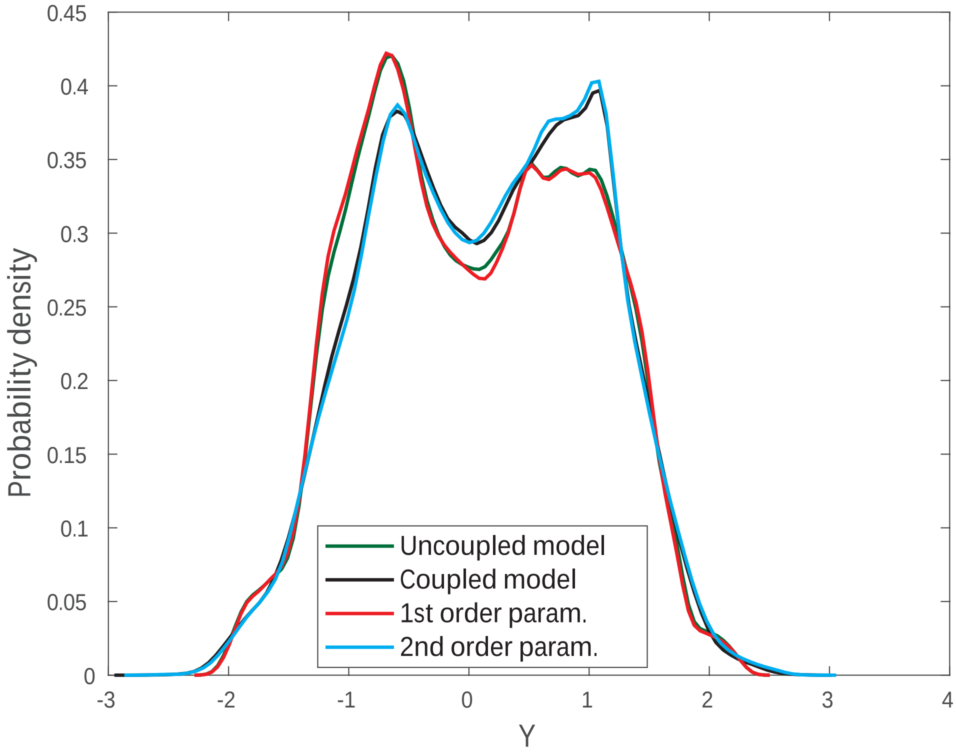


FIGURE 3.16: Probability density of the Y variable.

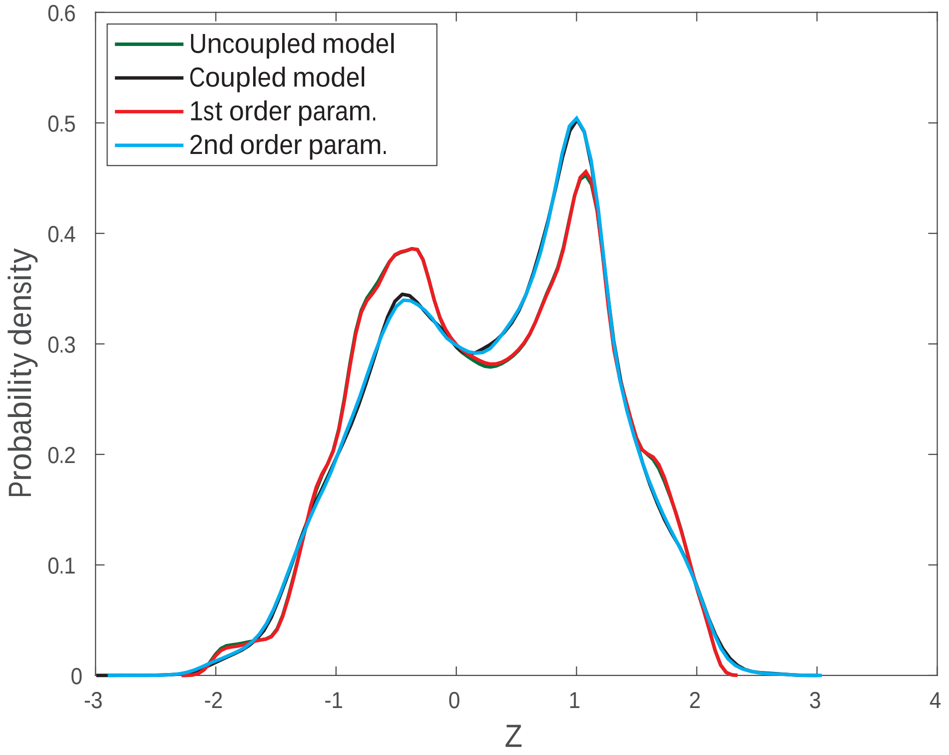


FIGURE 3.17: Probability density of the Z variable.

understanding of the statistics by adopting this point of view. As discussed above, this simple analysis is integrated with a more robust evaluation of the performance of the parametrizations by taking into account the Wasserstein distance. A first issue to deal with in order to compute the Wasserstein distance consists in carefully choosing the number of cubes used for the Ulam projection. Fig.3.18a shows the Wasserstein distance between the invariant measure of the coupled model projected on the XYZ space, used as reference measure, and the invariant measure of the uncoupled Lorenz and of the models obtained using the first and second order parametrization. For all choices of the coarse-graining the measure of the model with the second order parametrization is, by far, the closest to the projected measure of the coupled model. Instead, the deterministic parametrization provides a negligible improvement with respect to the trivial case of considering the uncoupled model, as expected given the discussion following Eq.(3.10). What shown here gives a quantitative evaluation of the improved performance resulting from adding a stochastic parametrization. The second piece of information is that the estimated Wasserstein distance has only a weak dependence on the degree of the coarse-graining and seems to approach its asymptotic value for the finest (yet still pretty coarse)

TABLE 3.2: Expectation values for the ensemble average of the first two moments of the variables X, Y, and Z. The uncertainty is indicated as standard deviation (std) σ over the ensemble of realizations. All the values are multiplied by 10^2.

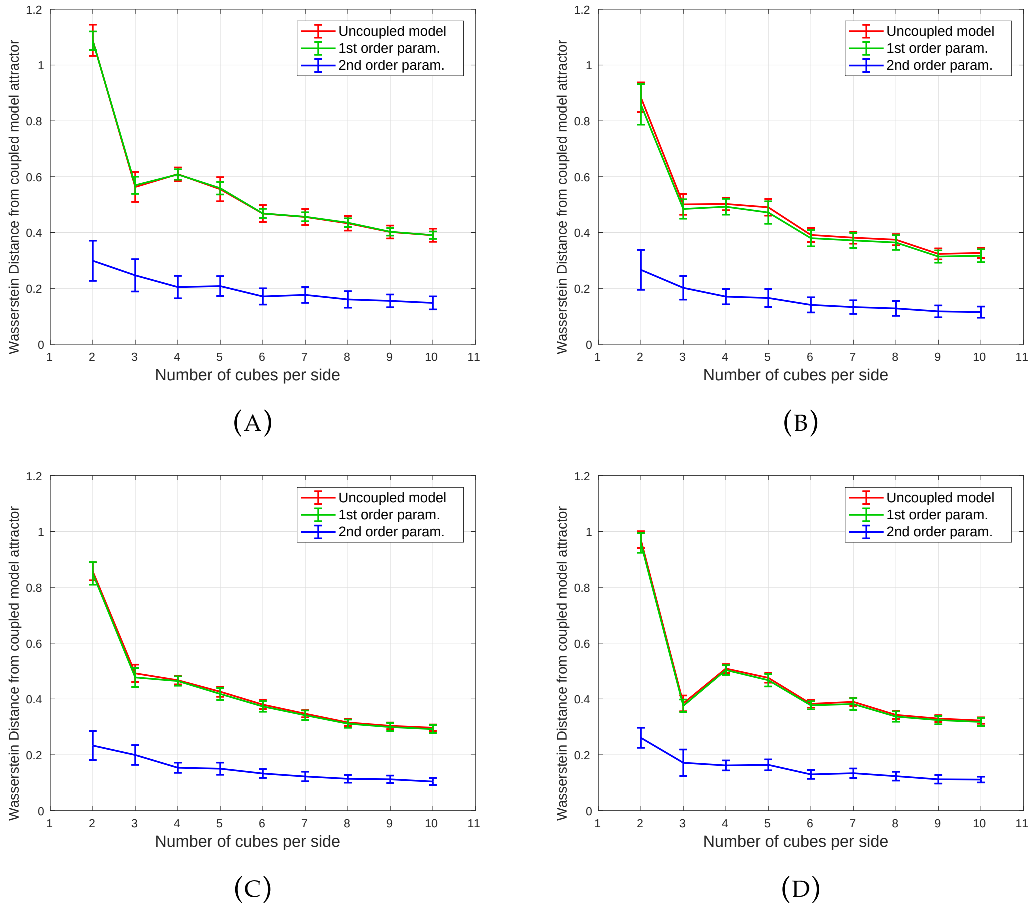
Observables	Uncoupled model ($\times 10^2$)	1st order parametrization ($\times 10^2$)	2nd order parametrization ($\times 10^2$)	Coupled model ($\times 10^2$)
$\overline{X} \pm \sigma_{\overline{X}}$	101.5 ± 0.4	101.3 ± 0.5	97.2 ± 0.3	97.1 ± 0.3
$\overline{Y} \pm \sigma_{\overline{Y}}$	6.1 ± 0.8	6.5 ± 1.2	13.7 ± 0.7	13.9 ± 0.4
$\overline{Z} \pm \sigma_{\overline{Z}}$	27.0 ± 0.2	26.9 ± 0.3	31.0 ± 0.2	31.3 ± 0.5
$var(X) \pm \sigma_{var(X)}$	34.9 ± 0.8	35.2 ± 1.0	43.6 ± 0.7	43.5 ± 0.3
$var(Y) \pm \sigma_{var(Y)}$	84.4 ± 0.1	84.4 ± 0.1	82.8 ± 0.4	82.6 ± 0.3
$var(Z) \pm \sigma_{var(Z)}$	82.6 ± 0.1	82.6 ± 0.2	81.5 ± 0.3	81.4 ± 0.3
$cov(XY) \pm \sigma_{cov(XY)}$	-5.4 ± 0.8	-5.7 ± 1.1	-11.1 ± 0.6	-11.2 ± 0.3
$cov(XZ) \pm \sigma_{cov(XZ)}$	-3.7 ± 0.1	-3.4 ± 0.2	-8.0 ± 0.2	-8.3 ± 0.4
$cov(YZ) \pm \sigma_{cov(YZ)}$	-7.7 ± 0.2	-7.7 ± 0.4	-1.6 ± 0.4	-1.3 ± 0.2

Ulam partitions considered here. This is encouraging as the findings one can obtain at low resolution seem to be already very meaningful and useful.

A well-known problem of Ulam's method is the fact that it can hardly be adapted to high dimensional spaces - this is the so-called curse of dimensionality. Additionally, evaluating the Wasserstein distance with a high number of dimensions becomes itself computationally extremely challenging. In order to partially address these problems the analysis shown in Fig.3.18a) is repeated for the measures projected on the XY, XZ and YZ planes, thus constructing the so-called sliced Wasserstein distances. Results are reported in panels b), c), and d) of Fig.3.18, respectively. Unsurprisingly, the distance of the projected measure is strictly lower than the distance in the full phase space, *ceteris paribus*. What is more interesting is that all the observations made for Fig.3.18a) apply for the other panels. Therefore, it seems reasonable to argue that studying the Wasserstein distance for projected spaces might provide useful information also on the full system.

Parametrizing the slow system

In order to extend the scope of this study the analysis described above has been repeated for the case $\tau = \frac{1}{6}$. Such a choice implies that the model responsible for the forcing has an internal time scale which is larger than the one of the model of interest. The WL parametrization, as discussed in Vissio and Lucarini, 2018a, is not based on any assumption of time scale separation between the variables of interest and the variables to parametrize.



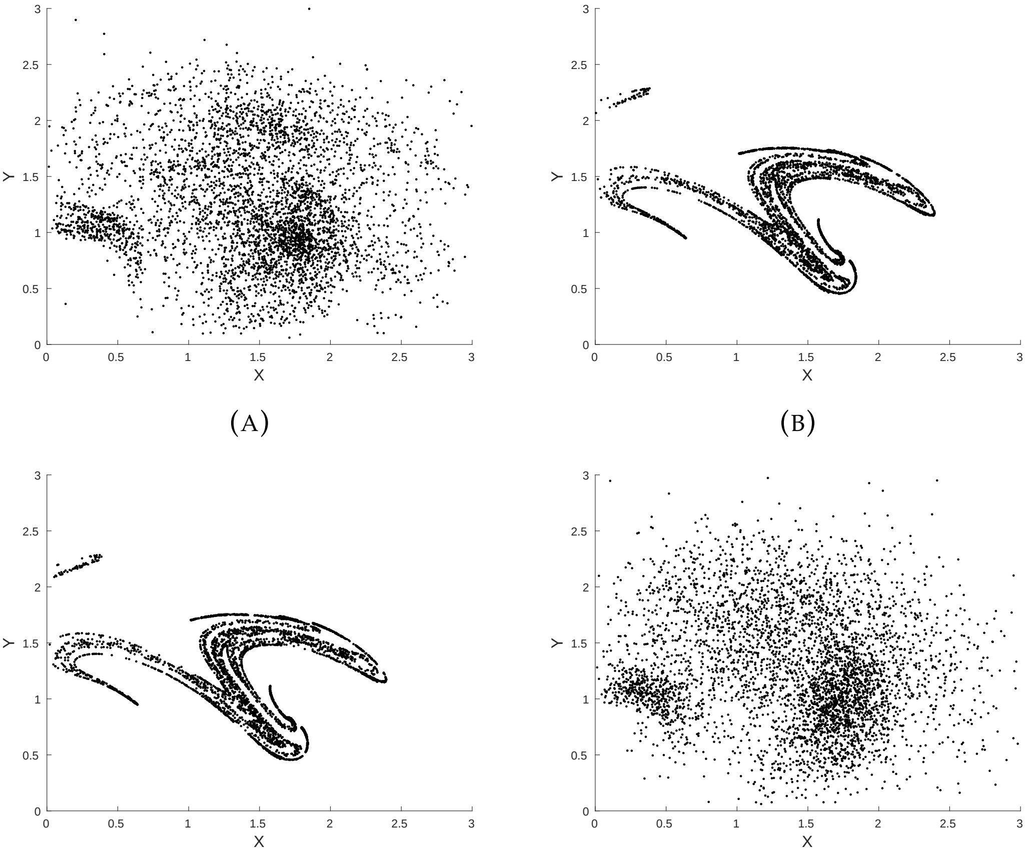
(A) (B)

(C) (D)

FIGURE 3.18: Wasserstein distances from the coupled model with respect to the number of cubes per side: a) $3D$ case; b) Projection on XY plane; c) Projection on XZ plane; d) Projection on YZ plane.

Figs.3.19a)-d) and Figs.3.20a)-d) show the Poincaré sections, respectively, in $Z = 0$ and $X = 1$ on all the considered models. In the case of the coupled system, most of the fine structure one finds in the uncoupled model is lost, and this emerges from a cloud of points with weaker features than what shown in Figs.3.12-3.13 for $\tau = 5$. Nonetheless, also in this case the model with the second order parametrization reproduces (visually) quite well what shown in Panel a), and, in particular, shows matching regions where the density of the points is higher. An inspection of Fig.3.21 allows to draw similar conclusions regarding the three dimensional trajectories in the phase space.

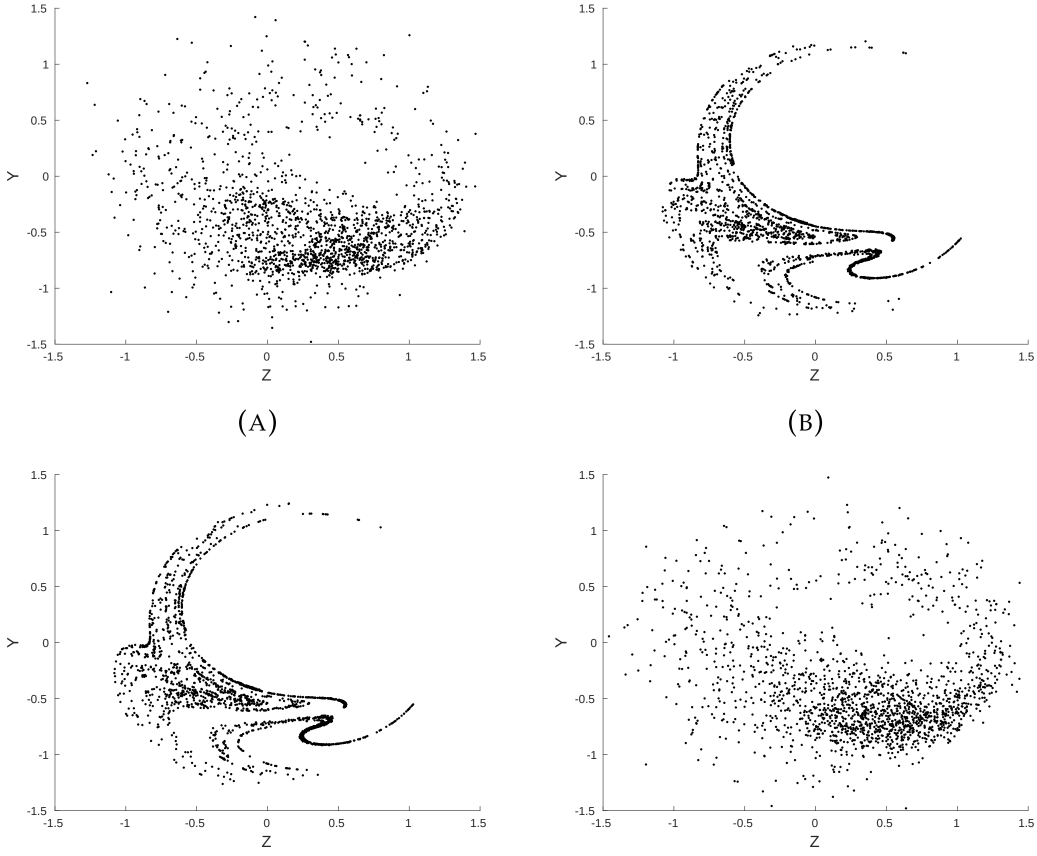
The PDFs shown in Figs.3.22-3.24, along with the moments written in Table 3.3, demonstrate that, notwithstanding the still very good approximation provided by the second order WL parametrization, the match between parametrization and full coupled model is not as good as in the previous case.



(A) (B) (C) (D)

FIGURE 3.19: Poincaré section in $Z = 0$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization.

The analysis performed considering the Wasserstein distance between the measures is shown in Fig.3.25. Without going into details, one finds that the same considerations made for $\tau = 5$ are still valid for $\tau = \frac{1}{6}$ regarding the performance of the parametrization schemes and the role of coarse graining. Additionally, for each choice of coarse-graining the distance between the measure of the parametrized models and the actual projected measure of the coupled model is larger for $\tau = \frac{1}{6}$, thus indicating the parametrization procedure performs worse in this case. This fits with the intuition one can have by checking out how well Panels b)-d) reproduce Panel a) in Fig. 3.19 versus the case of Fig. 3.12.



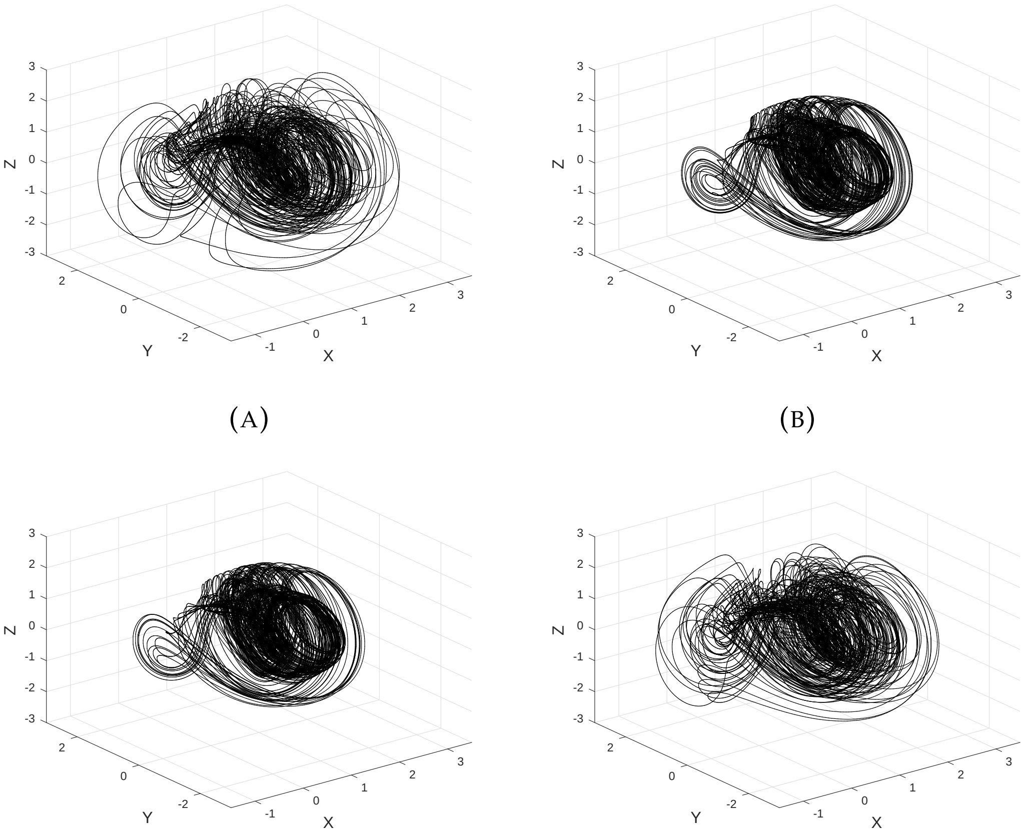
(A) (B)

(C) (D)

FIGURE 3.20: Poincaré section in $X = 1$ of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization.

TABLE 3.3: Expectation values for the ensemble average of the first two moments of the variables X, Y, and Z. The uncertainty is indicated as standard deviation (std) σ over the ensemble of realizations. All the values are multiplied by 10^2.

Observables	Uncoupled model ($\times 10^2$)	1st order parametrization ($\times 10^2$)	2nd order parametrization ($\times 10^2$)	Coupled model ($\times 10^2$)
$\overline{X} \pm \sigma_{\overline{X}}$	101.4 ± 0.3	101.4 ± 0.4	87.6 ± 0.5	90.7 ± 0.7
$\overline{Y} \pm \sigma_{\overline{Y}}$	6.3 ± 0.7	6.2 ± 0.8	26.8 ± 0.9	22.9 ± 1.1
$\overline{Z} \pm \sigma_{\overline{Z}}$	27.0 ± 0.2	27.0 ± 0.2	29.1 ± 0.4	32.6 ± 0.7
$var(X) \pm \sigma_{var(X)}$	35.0 ± 0.6	35.0 ± 0.8	50.9 ± 0.6	54.0 ± 0.6
$var(Y) \pm \sigma_{var(Y)}$	84.4 ± 0.1	84.4 ± 0.2	85.2 ± 0.3	82.0 ± 0.3
$var(Z) \pm \sigma_{var(Z)}$	82.6 ± 0.1	82.6 ± 0.2	78.0 ± 0.4	80.0 ± 0.2
$cov(XY) \pm \sigma_{cov(XY)}$	-5.5 ± 0.7	-5.5 ± 0.8	-21.0 ± 0.7	-17.6 ± 0.9
$cov(XZ) \pm \sigma_{cov(XZ)}$	-3.7 ± 0.1	-3.7 ± 0.1	-6.6 ± 0.3	-9.5 ± 0.6
$cov(YZ) \pm \sigma_{cov(YZ)}$	-7.6 ± 0.2	-7.7 ± 0.2	-0.1 ± 0.3	-0.4 ± 0.2



(A)

(B)

(C)

(D)

FIGURE 3.21: 3*D* view of the attractor of a) coupled model; b) uncoupled model; c) 1st order parametrization; d) 2nd order parametrization.

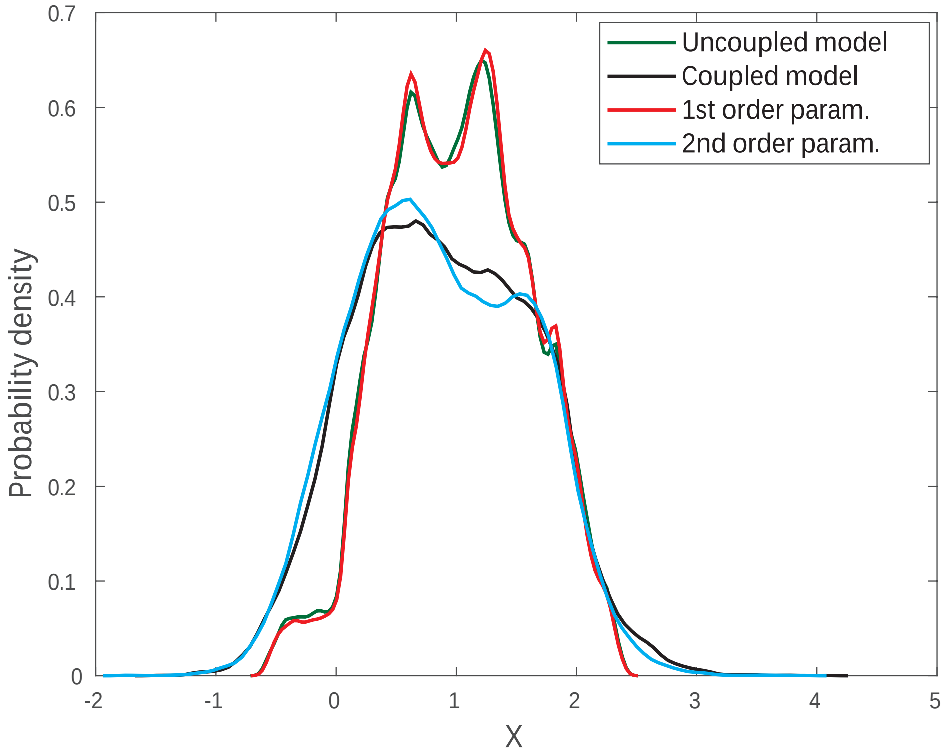


FIGURE 3.22: Probability density of the X variable.

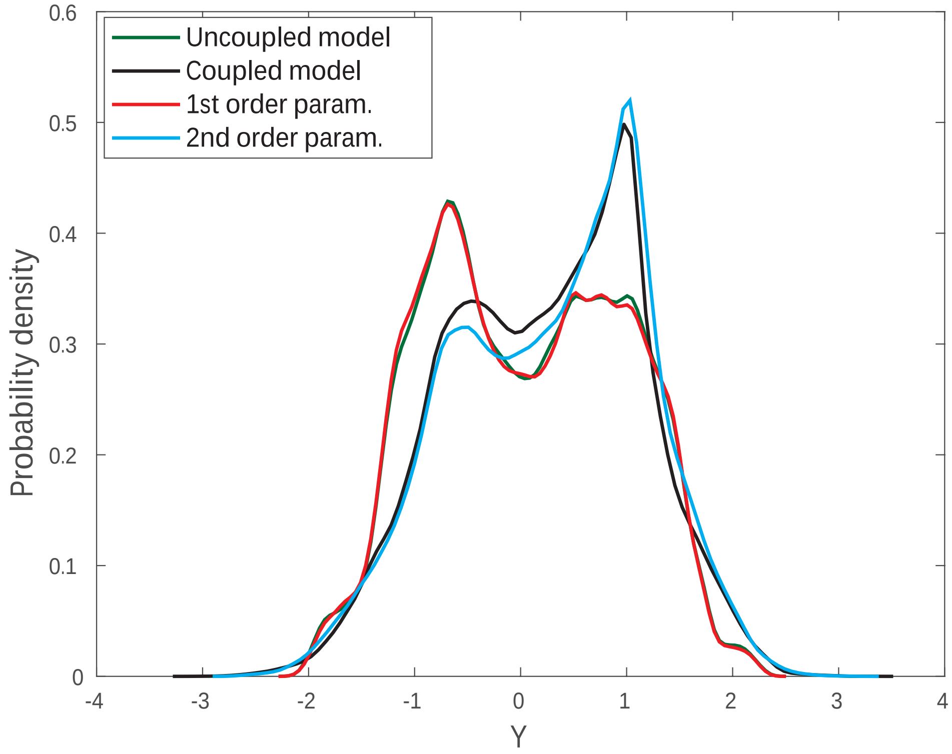


FIGURE 3.23: Probability density of the Y variable.

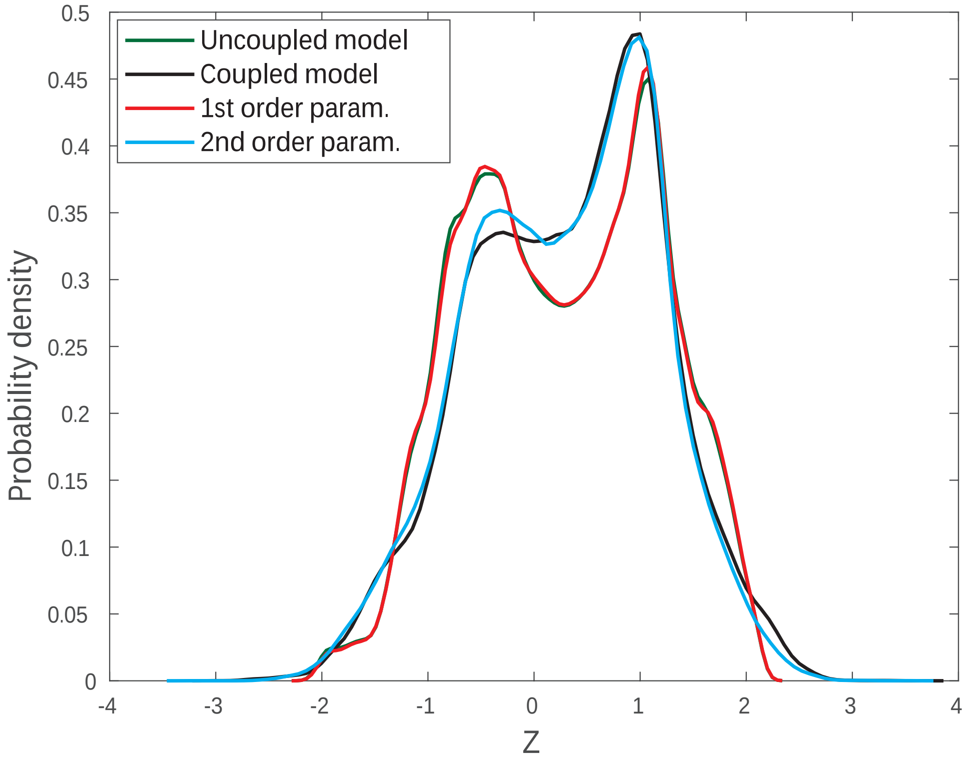
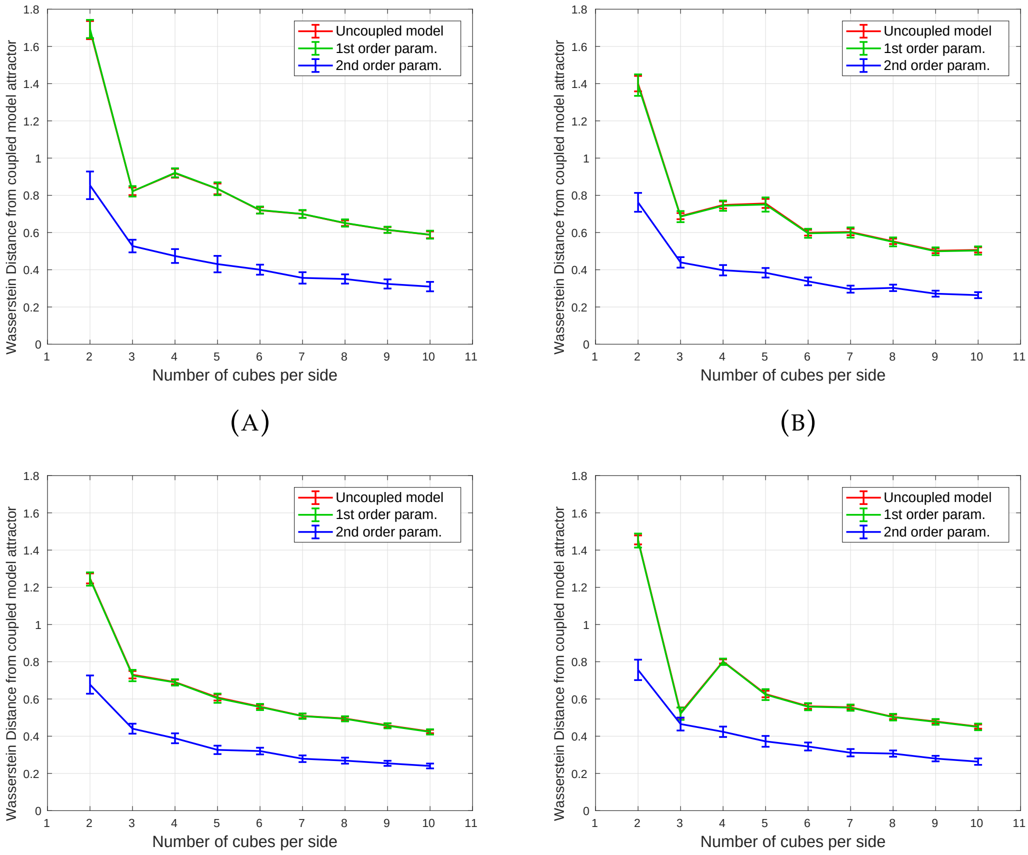


FIGURE 3.24: Probability density of the Z variable.



(A)

(B)

(C)

(D)

FIGURE 3.25: Wasserstein distances from the coupled model with respect to the number of cubes per side: a) $3D$ case; b) Projection on XY plane; c) Projection on XZ plane; d) Projection on YZ plane.

Chapter 4

Conclusions

The vast amount of scales on which the climate system evolves makes impossible to completely simulate it. Thus, the construction of reliable parametrizations that take into account the effect of the unresolved scales on the variables of interest is a fundamental task in climatological and meteorological sciences. Despite the extended effort and the ample available resources, a uniform and methodical approach to this issue has been unattainable so far. Indeed, these approximations are usually built to optimize some specific observable or property of interest, making the models ineffective on a whole range of physical phenomena. Furthermore, these parametrizations are attained in a way that does not normally allow for much flexibility, making necessary to compute them again in case of different settings.

As main objective, this thesis aims to apply the parametrization developed by Wouters and Lucarini, 2012; Wouters and Lucarini, 2013, in which the coupling between the variables of interest and the variables to parametrize is seen as a small perturbation to the uncoupled dynamics of the former ones, thus taking a weak coupling hypothesis. This parametrization, which describes the dynamical impact of the neglected variables, can be written as the sum of as a deterministic term (mean field effect), stochastic term (impact of fluctuations), and non-markovian term (role of memory). No hypotheses are taken on the scale separation between the systems. Furthermore, it is shown that the WL method can be used in general for constructing scale-adaptive parametrizations when multi-scale systems are considered.

The first application of this parametrization is on a slightly modified version of the two-layer Lorenz 96 model (Lorenz, 1996), which is a prototypical multi-scale system of great interest for nonlinear sciences in general. The coupling in the X variable equation of the Lorenz model is therefore substituted by a scale adaptive parametrization able to describe accurately the coupling between slow and fast scales, to describe conditions of finite scale separation and to reach the infinite time scale separation limit. In particular, the properties of the noise term responsible for the stochastic component of the parametrization and the memory kernel responsible for the non-markovian term are explicitly constructed.

The parametrization does a very good job in surrogating the effects of the fast variables, as tested by evaluating the expectation value and the correlation properties of the slow variables, and shows a great deal of flexibility when the coupling strength is set on different intensities.

The parametrization discussed is also tested against the heuristic approach previously proposed by Wilks, 2005. The Wilks method allows for constructing detailed parametrizations for each choice of the systems parameters, and outperforms the parametrizations constructed following Wouters and Lucarini. Nonetheless, the Wilks parametrization is not scale adaptive and needs to be retuned each time one or more parameters of the system are changed, whereas the WL parametrization is universal within the approximation defined by Eq.(2.14), except for the application of an algebraic rescaling, as proved by last tests. Indeed, the flexibility of this approach has been demonstrated by changing by two orders of magnitude the time scale separation and also in the most general case when all the parameters c, b and h are changed with respect to the original values.

Depending on the specific problem one needs to address, an accurate adhoc method or the flexible but less precise method proposed here might prove more advantageous.

In the second part of the thesis the WL parametrization is applied to a simple yet meaningful six-dimensional system constructed by coupling a Lorenz 84 model and a Lorenz 63 model, with the latter acting as forcing to the former, and the former being the subsystem of interest. A parameter controlling the time scale separation of the two systems and a parameter controlling the intensity of the coupling have been included. The second order scheme includes a stochastic term, which has proven to be essential for radically improving the quality of the parametrization with respect to the purely deterministic case (first order parametrization), as already qualitatively shown by looking at suitable Poincaré sections and at the three dimensional attractors and quantitatively by looking at probability density functions and first two moments. Remarkably, the WL approach flawlessly reproduces the change in modality of the distribution and, as expected, provides strong approximations for a wide range of observables.

One of the main novelties of the thesis lies in using the Wasserstein distance as a comprehensive tool for measuring how different the invariant measures ("the climates") of the uncoupled Lorenz 84 model and of its two versions with deterministic and stochastic parametrizations are from the projection of the measure of the coupled model on the variables of the Lorenz 84 model. The Wasserstein distance provides a robust tool for assessing the quality of the parametrization and, quite encouragingly, meaningful results can be obtained when considering very coarse grained representations of the phase space. A well-known issue of using a methodology like the Wasserstein distance is the so-called curse of dimensionality: the procedure itself becomes unfeasible when the system has a number of degree of freedom above few units. This issue has been (partially) addressed by looking at the Wasserstein distance of the projected measures on the three two-dimensional spaces spanned by two of the three variables of the Lorenz 84 model. The properties of the Wasserstein distance in the reduced spaces follow closely those found in the full space. As a result, it is found that diagnostics based on the Wasserstein distance in suitably defined reduced phase spaces could become standard in the analysis of the performance of parametrizations and in intercomparing models of any level of complexity.

The work done in this thesis is a first approach towards a new way of constructing parametrizations through statistical mechanical considerations. From the empirical, data-driven parametrizations, built ad-hoc for the particular systems studied, here the focus has been moved to top-down, mathematically sounding parametrizations. The aim is to reach a more sistematic and methodical way to represent the effects of the unresolved variables on the resolved ones, regardless of the system studied since, as seen in Chapter 1, the parametrization is a conundrum that crosses a wide range of different scales, both temporal and spatial.

The parametrization tested here has proven to be flexible and accurate, even though in the limited framework given by the simple Lorenz 96 model and the low order Lorenz 84 model forced by Lorenz 63 model. Nevertheless, useful evidence has been shown to support the feasibility of the application of the methodology to more complex and physical relevant models.

Bibliography

Abramov, R.V. (2016). “A Simple Stochastic Parameterization for Reduced Models of Multiscale Dynamics”. In: *Fluids* 1, pp. 1–18.

Abramov, R.V. and A.J. Majda (2008). “New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems”. In: *Journal of Nonlinear Science* 18.3, pp. 303–341.

Arakawa, A. (1969). “Parameterization of cumulus clouds”. In: *Proc. Symp. on Numerical Weather Prediction, Tokyo, Japan, WMO/International Union of Geodesy and Geophysics*, pp. 1–6.

— (2004). “The cumulus parameterization problem: Past, present, and future”. In: *Journal of Climate* 17, pp. 2493–2525.

Arakawa, A., J. Jung, and C. Wu (2011). “Toward unification of the multiscale modeling of the atmosphere”. In: *Atmospheric Chemistry and Physics* 11, pp. 3731–3742.

Berner, J. et al. (2017). “Stochastic Parameterization: Towards a new view of Weather and Climate Models”. In: *Bulletin of the American Meteorological Society* 98:3, pp. 565–588.

Blender, R. and V. Lucarini (2013). “Nambu representation of an extended Lorenz model with viscous heating”. In: *Physica D: Nonlinear Phenomena* 243, pp. 86–91.

Bódai, T., Gy. Károlyi, and T. Tél (2011). “A chaotically driven model climate: extreme events and snapshot attractors”. In: *Nonlinear Processes in Geophysics* 18, pp. 573–580.

Chandrasekhar, S. (1961). *Hydrodynamic and Hydromagnetic Stability*. Dover Publications.

Chekroun, M. D., H. Liu, and S. Wang (2015a). *Approximation of Stochastic Invariant Manifolds*. SpringerBriefs in Mathematics. Cham: Springer International Publishing.

— (2015b). *Stochastic Parameterizing Manifolds and Non-Markovian Reduced Equations*. SpringerBriefs in Mathematics. Cham: Springer International Publishing.

Cullen, M.J.P. and A.R. Brown (2009). "Large eddy simulation of the atmosphere on various scales". In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 367, 2947–2956.

Demaeyer, J. and S. Vannitsem (2017). "Stochastic parameterization of subgrid-scale processes in coupled ocean-atmosphere systems: Benefits and limitations of response theory". In: *Quarterly Journal of the Royal Meteorological Society* 143.703, pp. 881–896.

Eckmann, J. and D. Ruelle (1985). "Ergodic theory of chaos and strange attractors". In: *Reviews of Modern Physics* 57.4, pp. 617–656.

Fletcher, J.K. and C.S. Bretherton (2010). "Evaluating Boundary Layer–Based Mass Flux Closures Using Cloud-Resolving Model Simulations of Deep Convection". In: *Journal of the Atmospheric Sciences* 67, pp. 2212–2225.

Franzke, C., A.J. Majda, and E. Vanden-Eijnden (2005). "Low-order stochastic mode reduction for a realistic barotropic model climate". In: *Journal of the Atmospheric Sciences* 62, pp. 1722–1745.

Franzke, C.L.E. et al. (2015). "Stochastic climate theory and modeling". In: *Wiley Interdisciplinary Reviews: Climate Change* 6.1, pp. 63–78.

Fritsch, J.M. and C.F. Chappell (1980). "Numerical prediction of convectively driven mesoscale pressure systems. Part I. Convective parameterization." In: *Journal of the Atmospheric Sciences* 37, pp. 1722–1733.

Gallavotti, G. (2014). *Nonequilibrium and Irreversibility.* Springer International Publishing.

Gallavotti, G. and V. Lucarini (2014). "Equivalence of Non-equilibrium Ensembles and Representation of Friction in Turbulent Flows : The Lorenz 96 Model". In: *Journal of Statistical Physics* 156, pp. 1027–1065.

Garratt, J.R. (1992). *The Atmospheric Boundary Layer*. Cambridge, England: Cambridge University Press.

Ghil, M. (2015). "A Mathematical Theory of Climate Sensitivity or, How to Deal With Both Anthropogenic Forcing and Natural Variability?" In: *Climate Change: Multidecadal and Beyond*. Ed. by C.-P. Chang et al. Vol. 6. Singapore: World Scientific Publishing Co., pp. 31–52.

Hallerberg, S. et al. (2010). "Logarithmic bred vectors in spatiotemporal chaos: Structure and growth". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 81.6, pp. 1–8.

Hartmann, D.L. (1994). *Global Physical Climatology*. Cambridge, Massachusetts, USA: Academic Press.

Hilborn, R.C. (2000). *Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers*. Oxford, England: Oxford University Press.

Holton, J.R. (2004). *An introduction to dynamic meteorology.* Vol. 88. International Geophysics Series. Elsevier Academic Press.

Imkeller, P. and J.-S. von Storch (2001). *Stochastic Climate Models*. Basel, Switzerland: Springer International Publishing.

Kantorovich, L. (1942). "On the translocation of masses". In: *Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS* 37, pp. 199–201.

Kondrashov, D., M. D. Chekroun, and M. Ghil (2015). "Data-driven non-Markovian closure models". In: *Physica D: Nonlinear Phenomena* 297, pp. 33–55.

Kravtsov, S., D. Kondrashov, and M. Ghil (2005). "Multilevel Regression Modeling of Nonlinear Processes: Derivation and Applications to Climatic Variability". In: *Journal of Climate* 18, pp. 4404–4424.

Kundu, P.J. and I.M. Cohen (2010). *Fluid Mechanics - 4th Edition*. Elsevier Academic Press.

Li, F. et al. (2012). "Super-parameterization: A better way to simulate regional extreme precipitation". In: *Journal of Advances in Modeling Earth Systems* 4, pp. 1–10.

Lorenz, E.N. (1963). "Deterministic Nonperiodic Flow". In: *Journal of the Atmospheric Sciences* 20, pp. 130–141.

— (1984). "Irregularity: a Fundamental Property of the Atmosphere". In: *Tellus A: Dynamic Meteorology and Oceanography* 36, pp. 98–110.

— (1996). "Predictability - a problem partly solved". In: *Predictability of Weather and Climate*. Ed. by Tim Palmer and Renate Hagedorn. Cambridge University Press, pp. 40–58.

Lucarini, V. (2012). "Stochastic Perturbations to Dynamical Systems: A Response Theory Approach". In: *Journal of Statistical Physics* 146, pp. 774–786.

Lucarini, V. and S. Sarno (2011). "A statistical mechanical approach for the computation of the climatic response to general forcings". In: *Nonlinear Processes in Geophysics* 18, pp. 7–28.

Lucarini, V. and J. Wouters (2017). "Response formulae for n-point correlations in statistical mechanical systems and application to a problem of coarse graining". In: *Journal of Physics A: Mathematical and Theoretical* 50, p. 355003.

Lucarini, V. et al. (2014). "Towards a general theory of extremes for observables of chaotic dynamical systems". In: *Journal of statistical physics* 154, pp. 723–750.

Lutgens, F.K. and E.J. Tarbuck (2001). *The Atmosphere: 8th Edition*. New York City, USA: Pearson.

Majda, A.J. (2007). "Multiscale Models with Moisture and Systematic Strategies for Superparameterization". In: *Journal of the Atmospheric Sciences* 64, pp. 2726–2734.

Majda, A.J., I. Timofeyev, and E. Vanden-Eijnden (1999). "Models for stochastic climate prediction". In: *Proceedings of the National Academy of Sciences of the United States of America* 96, pp. 14687–14691.

— (2001). "A Mathematical Framework for Stochastic Climate Models". In: *Communications on Pure and Applied Mathematics* 54, pp. 891–974.

— (2003). "Systematic strategies for stochastic mode reduction in climate". In: *Journal of the Atmospheric Sciences* 60, pp. 1705–1722.

Mapes, B.E. (2000). "Convective inhibition, subgrid-scale triggering energy, and stratiform instability in a toy tropical wave model". In: *Journal of the Atmospheric Sciences* 57, pp. 1515–1535.

McGuffie, K. and A. Henderson-Sellers (2005). *A Climate Modelling Primer*. Chichester, UK: John Wiley & Sons, Ltd.

Monge, G. (1781). "Mémoire sur la théorie des déblais et des remblais". In: *Histoire de l'Académie Royale des Sciences*, pp. 666–704.

Mori, H., H. Fujisaka, and H. Shigematsu (1974). "A New Expansion of the Master Equation". In: *Progress of Theoretical Physics* 51, pp. 109–122.

Neumaier, A. and T. Schneider (2001). "Estimation of parameters and eigenmodes of multivariate autoregressive models". In: *ACM Transactions on Mathematical Software* 27.1, pp. 27–57.

Orrell, D. (2003). "Model Error and Predictability over Different Timescales in the Lorenz '96 Systems". In: *Journal of the Atmospheric Sciences* 60.17, pp. 2219–2228.

Ott, E. (1993). *Chaos in Dynamical Systems*. Cambridge, England: Cambridge University Press.

Palmer, T. and R. Hagedorn (2006). *The predictability of weather and climate*. Cambridge, England: Cambridge University Press.

Palmer, T.N. and P.D. Williams (2008). "Introduction. Stochastic physics and climate modelling." In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 366.1875, pp. 2421–2427.

Park, S. (2014). "A Unified Convection Scheme (UNICON). Part I : Formulation". In: *Journal of the Atmospheric Sciences* 71, pp. 3902–3930.

Pavliotis, G.A. and A.M. Stuart (2008). *Multiscale methods: averaging and homogenization*. Texts in applied mathematics : TAM, Springer.

Peixoto, J.P. and A.H. Oort (1992). *Physics of Climate*. New York, NY: American Institute of Physics.

Plant, R.S. and J.-I. Yano (2016). *Parameterization of Atmospheric Convection - Volume 1: Theoretical Background and Formulation*. London, England: Imperial College Press.

Pope, S.B. (2002). *Turbulent Flows*. Cambridge, England: Cambridge University Press.

Riehl, H. and J.S. Malkus (1958). "On the heat balance in the equatorial trough zone". In: *Geophysica* 6, pp. 503–537.

Robin, Y., P. Yiou, and P. Naveau (2017). "Detecting changes in forced climate attractors with Wasserstein distance". In: *Nonlinear Processes in Geophysics* 24, pp. 393–405.

Ruelle, D. (1978). "What Are the Measure Describing Turbulence?" In: *Progress of Theoretical Physics Supplement* 64, 339–345.

— (1998). "General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium". In: *Physics Letters A* 245(3-4), pp. 220–224.

— (2009). "A review of linear response theory for general differentiable dynamical systems". In: *Nonlinearity* 22(4), pp. 855–870.

Sakradzija, M., A. Seifert, and A. Dipankar (2016). "A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone". In: *Journal of Advances in Modeling Earth Systems* 8, pp. 786–812.

Saltzman, B. (1962). "Finite Amplitude Free Convection as an Initial Value Problem—I". In: *Journal of the Atmospheric Sciences* 19, pp. 329–341.

Schneider, T. and A. Neumaier (2001). "Algorithm 808: ARfit—a matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models". In: *ACM Transactions on Mathematical Software* 27.1, pp. 58–65.

Smith, R.K. (1997). *The Physics and Parameterization of Moist Atmospheric Convection*. Vol. 505. Nato Science Series C. Springer Netherlands.

Solomon, S. et al. (2007). *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (AR4). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Stocker, T.F. et al. (2013). *Climate Change 2013: The Physical Science Basis.* Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Tantet, A., V. Lucarini, and H.A. Dijkstra (2018). "Resonances in a Chaotic Attractor Crisis of the Lorenz Flow". In: *Journal of Statistical Physics* 170(3), pp. 584–616.

Trevisan, A., D. Isidoro, and O. Talagrand (2010). "Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension". In: *Quarterly Journal of the Royal Meteorological Society* 136, pp. 487–496.

Trevisan, A. and F. Uboldi (2004). "Assimilation of Standard and Targeted Observations within the Unstable Subspace of the Observation – Analysis – Forecast Cycle System". In: *Journal of the Atmospheric Sciences* 61, pp. 103–113.

Uboldi, F. and A. Trevisan (2015). "Multiple-scale error growth in a convection-resolving model". In: *Nonlinear Processes in Geophysics* 22, pp. 1–13.

Ulam, S.M. (1964). *Problems in Modern Mathematics.* A collection of mathematical problems. New York: Science Edition Wiley.

Vannitsem, S. and V. Lucarini (2016). "Statistical and dynamical properties of covariant Lyapunov vectors in a coupled atmosphere-ocean model - multiscale effects, geometric degeneracy, and error dynamics". In: *Journal of Physics A: Mathematical and Theoretical* 49, p. 224001.

Villani, C. (2009). *Optimal Transport: Old and New.* Berlin Heidelberg, Germany: Springer-Verlag.

Vissio, G. and V. Lucarini (2018a). "A proof of concept for scale-adaptive parametrizations: the case of the Lorenz '96 model". In: *Quarterly Journal of the Royal Meteorological Society* 144, pp. 63–75.

— (2018b). "Evaluating a stochastic parametrization for a fast–slow system using the Wasserstein distance". In: *Nonlinear Processes in Geophysics* 25, pp. 413–427.

Wilks, D.S. (2005). "Effects of stochastic parametrizations in the Lorenz '96 system". In: *Quarterly Journal of the Royal Meteorological Society* 131, pp. 389–407.

Wouters, J. and V. Lucarini (2012). "Disentangling multi-level systems: averaging, correlations and memory". In: *Journal of Statistical Mechanics: Theory and Experiment* 2012(03):P03003.

— (2013). “Multi-level Dynamical Systems: Connecting the Ruelle Response Theory and the Mori-Zwanzig Approach”. In: *Journal of Statistical Physics* 151(5), pp. 850–860.

— (2016). “Parametrization of Cross-scale Interaction in Multiscale Systems”. In: *Climate Change: Multidecadal and Beyond*. Ed. by C.-P. Chang et al. Vol. 15. Singapore: World Scientific Publishing Co., pp. 67–80.

Wouters, J. et al. (2016). “Parametrization of stochastic multiscale triads”. In: *Nonlinear Processes in Geophysics* 23, pp. 435–445.

Young, L.S. (2002). “What are SRB measures, and which dynamical systems have them?” In: *Journal of Statistical Physics* 108, pp. 733–754.

Zwanzig, R. (1960). “Ensemble Method in the Theory of Irreversibility”. In: *The Journal of Chemical Physics* 33, pp. 1338–1341.

— (1961). “Memory effects in irreversible thermodynamics”. In: *Physical Review* 124, p. 983.

Eidesstattliche Versicherung

Declaration of oath

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.
I hereby declare, on oath, that I have written the present dissertation by myself and have not used other than the acknowledged resources and aids.