Universitat Autònoma de Barcelona

Escola d'Enginyeria
Departament d'Arquitectura de Computadors
i Sistemes Operatius

Genetic Ensemble (G-Ensemble): An Evolutionary Computing Technique for Numerical Weather Prediction Enhancement

Thesis submitted by **Hisham W. Y. Ihshaish** in fulfillment of the requirements for the doctoral degree from Universitat Autònoma de Barcelona, advised by Dra. Ana Cortés Fité and Dr. Miquel A. Senar Rosell.

Barcelona, June 2012

Genetic Ensemble (G-Ensemble): An Evolutionary Computing Technique for Numerical Weather Prediction Enhancement

Thesis submitted by **Hisham W. Y. Ihshaish** in fulfillment of the requirements for the doctoral degree from Universitat Autònoma de Barcelona. This work has been developed under RD 1393/2007 in the High Performance Computing doctoral program and presented to the Computer Architecture & Operating Systems Department at the Escola d Enginyeria of Universitat Autònoma de Barcelona. This thesis was advised by Dra. Ana Cortés Fité and Dr. Miquel A. Senar Rosell.

Barcelona (Spain), June 2012

Ph.D. candidate:

Hisham W.Y. Ihshaish

Thesis advisors:

Dra. Ana Cortés Fité   Dr. Miquel A. Senar Rosell

to those who lost their lives in that long way to freedom;
in Egypt, Tunis, Libya, Syria, Palestine and everywhere,

to ***Lama****, the small niece who is waiting to see her uncle,*

to my parents, ***Waleed*** *and* ***Samira****,*
and to ***Adriana****.*

Acknowledgements

This dissertation would not have been possible without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

Above all, I would like to express my gratitude to my advisors Dra. Ana Cortés (Anna), and Prof. Miquel A. Senar, for their excellent guidance, caring, and providing me with an excellent *atmosphere* for doing research. My deepest thanks to Anna for her daily practice of patience with me.

My grateful thanks also go to my professors for their continuos help and advise. I would specially mention Dr. Bahjat Qazzaz, the one who pushed me to apply to this opportunity in order to continue my studies at this department. Great deals appreciated go as well to Prof. Tomàs and Prof. Emilio, those who were available always, the source of fast and efficient solutions.

I would thank every body in our Computer Architecture and Operating Systems department, who in one way or another contributed by their valuable assistance and friendship, in the completion of this study. I would never forget how people like Gemma, Dani, or Javi offer their valuable help continuously, with a wide smile.

When I came to Barcelona, every thing for me was new, different, and thus, discoverable. In spite of the difficulties I faced before starting comfortably to use the Metro, I have to admit, however, that I never suffered being in Barcelona, I was always welcome, people are helpful and so kind. And a part of that rigid stuff of informatics, I enjoyed and, I learned a lot of valuable things, language/s, cultural and social nroms, and life style.

I would also mention those who were always there, those true friends; Alexandre for his sincere and special friendship, muito obrigado, Ganso. Grazie Matteo, goes to that brother whom I discovered here in Barcelona, It was always special to tell him that Palestine won one football match against Italy (I hope It was true), cultura ultras, Gaspa.

I would also like to thank my parents, sisters, and brothers. They were always supporting me and encouraging me with their best wishes. To Victoria, Juan, and Claudia for making me feel one more in their special family. Last, but by no means least, I would like to thank my wife, Adriana. She was always there cheering me up and stood by me through the good times and bad.

For any errors or inadequacies that may remain in this work, of course, the responsibility is entirely my own.

Contents

Chapter 1

Introduction

The trouble with weather forecasting is that it s right too often for us to ignore it and wrong too often for us to rely on it.

Patrick Young

1.1 Weather Prediction - An Ongoing Demand

While the first and foremost aim of weather forecasting was to render timely advice to farmers:- on the actual and expected weather, and its likely impact on the various day-to-day farming operations (e.g., Frost can damage young plants or blooming trees and lead to economic lost). However, it s very generally known, and true, the weather has such a widespread impact on people s personal and social lives, including their jobs, their recreation, their safety, and their property. When the weather is bad, many activities become more difficult to perform. Commercial transportation slows down on the roads, on the waterways, and in the air. Businesses of all kinds are interrupted by bad weather. Power plants and energy traders rely on knowledge of the weather to operate their equipment and to deliver power to consumers, government and business.

Reliable weather forecasting outcomes are crucial and very necessary for our (almost daily) decision making processes, for example, to reduce economical loss facing a predicted strong storm, to redirect and control traffic in places where heavy snowing is to occur, to optimize and schedule the trade and distribution that depend on flights or ships, and among a wide number of examples, the most important, those decisions depending on weather forecasting, may even help keep us out of danger. Moreover, the evolution of some very dangerous natural disasters like hurricanes and floods, when predicted, could help in avoiding large damages of properties, and save lives.

A good example to illustrate the effects of the atmospheric changes on our lives are a class of strong storms on ocean surfaces, *Tropical Cyclones* and *Hurricanes*, which are dangerous meteorological phenomenon with the potential to cause damage, serious social disruption, and loss of human life. Every year, they cause considerable loss of life and do immense damage to property. To have an idea, a list of some notable *Tropical Cyclones* in the last five decades is shown below (from [1]):

- Katrina - 2005 - Katrina was the most costly hurricane on record causing an estimated $75 billion in damage in Louisiana and Mississippi.
- Andrew - 1992 - Andrew was a Category 5 hurricane which hit south-east Florida and south-east Louisiana causing $44,878 million of damage.
- The most deadly tropical cyclone ever recorded hit Bangladesh in 1970 killing approximately 300,000 people as a result of the storm surge.
- Camille - 1969 - Camille was a Category 5 hurricane with winds of 190 m.p.h. It hit Mississippi, south-east Louisiana and Virginia, causing damage of around $14,870 .

Furthermore, accurate predicted weather variables are critically needed for other environmental modelling systems. For instance, wind direction and velocity variables are needed as precise as possible to predict the expansion direction and velocity of a fire propagation disaster predicted by wildfire models.

Consequently, the pre-knowledge of the atmosphere future state has been continuously demanded for thousands of years, and correspondingly, efforts to predict weather phenomena began very early, around 650 B.C., when the Babylonians observed cloud patterns to predict the weather. Later, the Greek philosopher *Aristotle* described weather patterns in *Meteorologica* (ca. 350 B.C.) and later, his student *Theophrastus* compiled the *Book of Signs* on weather forecasting. In China and India, weather prediction can be traced as far back as 300 B.C. [2].

However, it was recently in the last decades, when modern weather forecasting began to be more reliable and thus, more useful. That is, weather forecasting started to be a sophisticated process that involves a combination of computer models, observations, and knowledge of weather trends and patterns. Using these methods, reasonably accurate forecasts can be made up to few days in advance. Beyond that, detailed forecasts are less useful, since atmospheric conditions such as temperature and wind direction are very complex.

Because there is no analytical solution for the equations that describe the atmospheric flow, therefore numerical methods are needed. Precisely, it was in the early 1950s when the USA National Weather Service (NWS)[3] began to utilize some of the early versions of computers to make large-scale weather forecasts, running Numerical Weather Prediction models (NWP models) (described in detail in *Chapter 2*).

Since that time, computers have become faster and more sophisticated being able to provide the scientific community (particularly to the weather forecasting community) with High Performance Computing (HPC) platforms, which allow the execution of highly computing demanding weather forecast simulations. The origins of computer weather prediction to up-to-date is described in details in [4].

Nowadays, forecasts, both for the next couple of hours and for the next couple of days, are issued daily. Apart from helping people decide when they should invite their neighbors for a barbecue, weather forecasts provide vital information for a wide range of occupational categories such as farmers, pilots, sailors and soldiers. However, as most scientific applications (including those numerical models for weather prediction) continue to be more complex while research is getting more sophisticated as a result of the natural human growth of requirements. In the context of weather predictions, higher accuracy, larger time scales, more complex processing of enormous data amounts and less waiting time constitute some of the new demands that should be considered.

1.2 Weather Prediction Quality Problem

Numerical models are used for forecasting across a wide range of environmental applications. The aim is to have a model that characterizes the behavior of the system of interest as accurately as possible, whilst satisfying known physical properties. Unfortunately, a mathematical model can never completely describe the complex physical processes underlying a real world dynamical system.

Everyone knows that weather forecasts go wrong sometimes. Reasons for this vary, but the chaotic nature of the atmosphere [5] means there is always a limit to what we can predict accurately. NWP models as well as the atmosphere itself can be viewed as nonlinear dynamical systems in which the evolution depends sensitively on the initial conditions. Moreover, weather prediction is, by its very nature, a process that has to deal with uncertainties. The initial conditions of a NWP model can be estimated only within a certain accuracy. During a forecast, some of these initial errors can amplify and result in significant forecast errors.

Advances in knowledge and computing technology mean that environmental forecasting models are becoming increasingly sophisticated, and particularly, NWP models have been strongly developed in the last decades and their performance constantly increases with the computational power [4], but in practice these models suffer from uncertainty in their initial conditions and parameters. The initial conditions are not known with exactitude, that is, the meteorological observational network is heterogeneously distributed around the world. Moreover, the observations are punctual and do not allow the monitoring of the current state of the atmosphere in the three dimensions. Additionally, there are measurements errors. Even with perfect initial data, inaccurate representation of model parameters will lead to the growth of model error and therefore, affect the ability of a model to accurately predict the true system state.

Besides initial-condition error, weather and climate prediction models are also sensitive to errors associated with the model itself. In particular, the uncertainty due to the parameterizations of sub-grid-scale physical processes is known to play a crucial role in prediction quality (e.g., [6]). Model parameters are intrinsic to environmental modelling. Parameterizations (details in *Chapter 2*) are typically used in applications where the underlying physics of a process are not fully known or understood, or to model sub-grid scale effects that cannot be captured within a particular model resolution. Prediction errors caused by the uncertainty in physical parameterizations is commonly referred to as model errors. Being that said, weather predictability errors are normally subject to two kinds of errors, initial condition errors and model errors.

By figuring out the main sources of error in predictability of NWP models, many efforts had been focusing on enhancing prediction quality, mainly on developing sophisticated and skillful next-generation NWP models (e.g., [7] and [8]), addressing the uncertainty of initial conditions by better estimation techniques, and also on developing physical parametrization models or schemes which are nowadays coupled with NWP models and lead to improved predictive skill.

Over the past 20 years or so, stochastic or ensemble forecasting [9] became a practical and successful way of addressing the predictability problem associated with the uncertainty in initial conditions. Ensemble forecasting is conducted by better estimations of the atmospheric initial state (initial conditions) which is produced by data assimilation (DA)[10] techniques, and then, initial state perturbations are computed and launched in different forecasts, each is initiated by a perturbed initial state. Early on moreover, several weather prediction centers have addressed this problem by developing operational ensemble prediction systems (EPS) (e.g., [11]). The Ensemble spread finally, is used to indicate forecast uncertainty. However, and although it has been realized that there is a stochastic nature of physical parameterizations in ensemble prediction (predictability is sensitive to variations in physical parameters), it has not been straightforward to develop theoretically sound, and also practical, formulations for how to insert parameterization uncertainty into ensemble development [12, 13].

On the other hand, and in contrast to the dynamics of NWP models, which are based on fundamental physical concepts, physical parameterizations, although partly are based on fundamental concepts of physics, involve empirical functions and tunable parameters, which usually referred to as model closure parameters. Practically, all physical parametrization schemes contain closure parameters and typically, expert knowledge and manual techniques are used to define the optimal parameter values, based on observations, process studies, large eddy simulations, etc. Therefore, some parameter value combinations score better than others, but it is very demanding to manually specify the optimal combination.

1.3 Computational science and High Performance Computing

Understanding various phenomena and processes from science, nature, and engineering is, today, no longer merely based on theory and experiment, but more on computations, as well.

Computational Science - sometimes referred to as *Computational Science and Engineering* (CS&E) or *Scientific Computing* - allows us to supplement experiments by simulations in order to study some technical systems and natural phenomena, that would be too time-consuming, expensive, or dangerous (if possible at all) to study by experiments alone. It is generally defined as an interdisciplinary field situated in the intersection of three major scientific domains: mathematics, computer science, and (natural and social) sciences and engineering (see figure 1.1), to study systems of real-world scientific or societal interest, usually through computer simulation and modelling [14].

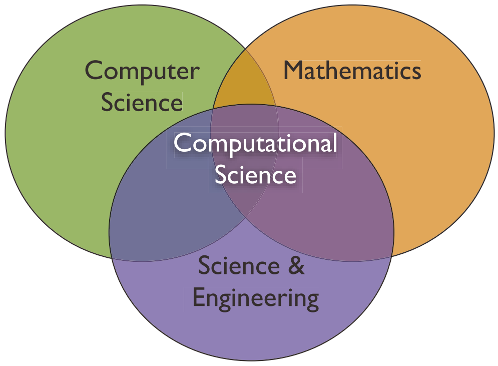


Figure 1.1: Computational Science: an interdisciplinary field which is correctly positioned in the intersection between three scientific domains: Mathematics, Computer science, and Science and Engineering.

In other words, it could be described as a science that focuses on the appropriate use of the computational architectures applying some mathematical algorithms to solve large and complex scientific problems.

In the last 30 years or so, computational science in various areas has emerged and revolutionized significantly enabling the scientific community to better study, understand, and predict some complex scientific problems and phenomena of considerable importance. By integrating knowledge and expertise from different application fields, like chemistry, biology, astronomy, climatology and many other fields, and to understand related complex problems in each domain, scientific computing puts this knowledge together with mathematical modelling, numerical analysis, algorithm development, software implementation, execution, and furthermore, validation and visualization of results. By doing so, many of principal issues and questions in the science are being possible to be solved computationally.

Many examples could be mentioned in this context, for instance, solving complex chemical equations born from the mathematicians and theorists would not be possible without tools and simulations provided by computational scientists. The same happens in areas of environmental sciences between others, by the introduction of environmental modelling, many crucial environmental issues and complex problems are being able to be studied and analyzed, furthermore, the evolution of different natural phenomena with huge effect on our life are being predicted by methods of scientific computing. One of these very important natural phenomena is climate change and weather variability. Numerical weather prediction models along with climate models are major instruments which today, rationally talking, are making our lives easier, more effective, and sometimes safer.

It should be mentioned, however, that the emergence of computational science would not have been possible without the advances in supercomputing and the availability of high performance tools and architectures. These tools and technologies made possible the execution to solve and process massive amount of calculations and data. But beyond all advances in computer architectures, it is of crucial importance to provide computational methods and algorithms to efficiently use the available infrastructures in order to solve large-scale applications. Thus, many tools, algorithms, and parallel computing paradigms are being developed for the optimization of parallel large-scale applications in order to maximize the benefits from the advances in computer architecture and technology, and even more, to enable the execution of larger scientific applications treating more scientific details.

In this thesis particularly, a scientific methodology is introduced based on computational science concepts and the knowledge in high performance computing tools and platforms, in order to enhance prediction quality in weather forecasting models. That is, by the development of scientific computing methods and algorithms which are optimized and parallelized to be executed efficiently over high performance computing platforms which allows us to get better weather predictions in shorter waiting times.

More details can be found in the following section, which gives a more detailed description of the overall contribution of this work.

1.4 Contribution

The main goal of the presented work is to tackle the problem of accuracy and waiting time in weather forecasting. As mentioned before, NWP models have been strongly developed in the last decades and their performance constantly increases with the advances in computational power. However, in practice, many serious challenges in this field are still gaining considerable efforts by the scientific community in order to reduce what is widely known as weather limited predictability . Mainly, the major two challenges are the willingness to get more reliable weather predictions, and to do it faster.

As in many other areas of environmental modelling, more especially in NWP models, most simulation software works with well-founded and widely accepted models, the need for input parameter optimization to improve model output is a long- known and often-tackled problem. Particularly in such environments where correct and timely input parameters cannot be provided. Efficient computational parameter estimation and optimization strategies are required to minimize the deviation between the predicted scenario and the real phenomenon behaviour. With the continuously increasing availability of computing power, evolutionary optimization methods, especially Genetic Algorithms (GA), have become more popular and practicable to solve the parameter problem of environmental models.

Based on the before mentioned, this thesis intends to:

1. Provide a sensitivity study of the effect of NWP model input parameters on prediction quality.

2. Propose a valid framework, which allows to search for the most optimal values of model input parameters which, in our hypothesis, will provide better prediction quality.

3. Reduce the waiting time needed to get more reliable weather predictions.

4. Demonstrate that, by implementing evolutionary computing techniques and an efficient use of the available high performance computing platforms, we can enhance weather predictions and reduce at the same time, the waiting time needed for it, in comparison with other operational methods of weather prediction enhancement.

To accomplish the objectives of the presented proposal, a new weather prediction scheme is introduced. This new scheme implements an evolutionary computing algorithm which focuses on the calibration of input parameters in NWP models. In more detail the main contributions are:

1. **G-Ensemble (1-point observation)**: which is a two-phase prediction scheme, where a combination of model input parameters are optimized within an aggregated calibration phase. The calibration process is done using one single observation, available at the end of calibration interval. As such, all the possible combinations of model parameters are used in short forecasts, each of which is evaluated using the available observation, then, Genetic Algorithm operators are applied to reproduce a next generation of parameter combinations. This process is repeated iteratively until satisfying a predefined condition.

2. **G-Ensemble (window observations)**: an extended version of the G-Ensemble, where a set of observations are used in the evaluation process during calibration phase, instead of one single point observation. This approach is designed in order to fairly guide the used

Genetic Algorithm in deciding which are the best combinations to be selected for the next iteration.

3. **M-Level G-Ensemble**: this approach is developed in a three-phase scheme, in order to consider more than one combination of model input parameters in the calibration process. This is accomplished by adding the parameter selection phase, which makes this approach capable of selecting automatically more than one level of model parameters to be calibrated, the selection process is done considering the particularity of the domain, that is, the selected parameters for one domain, would be totally different of the targeted parameters for another domain.

4. **BeGEM** and **G-Ensemble Set**: two different approaches for conducting a prediction process are introduced; the first is by BeGEM which refers to the Best Genetic Ensemble Member, considered as a deterministic forecast, i.e., one single calibrated forecast is to be conducted for prediction. The other approach (G-Ensemble), is an Ensemble Prediction approach, using a calibrated set of forecasts for prediction, rather than one single deterministic forecast.

5. **Parallel G-Ensemble**: all these strategies has been implemented in a parallel scheme under the Master/Worker programing paradigm in order to be executed on High Performance Computing platforms aiming at reducing the total execution time of the prediction process.

1.5 Outline

The next chapter discusses the concepts of numerical weather prediction modelling. The main principles of weather forecasting are explained, as well as NWP model components, input and output. Furthermore, the problem of predictability in these models is highlighted and analyzed along with the mainly used enhancement methods in operational forecasting.

In chapter three, the question of parameter estimation in environmental models in general, and in NWP particularly, is discussed, next, a brief study of the most popular methods for parameter optimization in such models is presented, with focus on Genetic Algorithms. Later, the proposed scheme Genetic Ensemble for Numerical Weather Prediction Enhancement is presented in detail. Basically, starting from describing its objectives and components, the framework of the proposal, and the targeted parameters for better estimation are also discussed.

Chapter four presents an experimental part, where the proposed scheme is evaluated by experimentations over a famous real weather prediction case;

Hurricane Katrina. In the same chapter, the objective of the experiments are listed along with the targeted weather variables for enhancement. Experimental results are presented finally and discussed duly.

Conclusions of the presented work is provided in chapter five. A short summery of the overall thesis is provided, furthermore, the chapter involves a list of scientific publications which were realized during the progress of the thesis. Finally, an overview of open research lines and future work is discussed.

Chapter 2

Numerical Weather Prediction

The need for local weather predictions increases. Various end-users need reliable forecasts to assure and optimize their activities. In fact, the level of detail in modern weather predictions allows for a wide variety of products and forecast fields to be delivered, for use not only in general meteorology, but also in specialized areas such as aviation and air quality. Moreover, in severe weather situations, weather forecasts and warnings can help protect property, and even more, the can help save lives. It is hence, certainly vital that weather forecasts be as accurate as possible.

Instead of simply looking at current conditions of the weather, and estimating based on past observances, forecasters or meteorologists, nowadays, use numerical weather prediction (NWP) models to do the job. These models, are complex computer programs, also known as forecast models, which run on supercomputers and provide predictions on multiple atmospheric variables such as, but not limited to, temperature, pressure, wind, and rainfall.

This thesis presents a proposal that intends to enhance the predictability of numerical weather prediction (NWP) models. As such, to increase the accuracy of the estimations produced by these models. Consequently, through this chapter, numerical weather models, their concepts, and principles are described. Moreover, the limitations in weather predictability, along with the sources of error in these models are highlighted. Later one, the commonly used statistical and visualization tools to evaluate and deliver weather predictions are discussed. Finally, existing weather prediction enhancement methods are introduced, including the mostly used approaches: Data Assimilation and Ensemble Prediction Systems.

2.1 Numerical Weather Prediction Modelling

The application of science and technology to predict the state of the atmosphere for a given location is generally known as Weather Forecasting [2]. Due to the complexity of the atmosphere as a very dynamic system, computer models are normally needed to produce forecasts.

Schulze in [15], described Weather Forecasting as a process that involves converting observational data into forecasts via a model that is done through three basic phases (figure 2.1): The first is the collection and analysis of meteorological data to define the initial conditions of the model as accurately as possible using data assimilation techniques. The second phase requires the use of deterministic numerical prediction models to project the initial conditions of the system into future states. The third phase the process of converting the output from numerical models into information of practical value for users.

Being that said, the idea of numerical weather prediction (NWP) models, was firstly recorded by **Vilhelm Bjerknes** (*Norwegian physicist and meteorologist, "the father of modern meteorology", 1862-1951, see [17]*), almost a century ago in his paper [18], when he discussed that it would be possible to forecast the weather by solving a system of nonlinear partial differential equations.

Later on, **Lewis Fry Richardson** (*English mathematician, physicist, and meteorologist, 1881-1953, see [19]*), proposed the first numerical weather prediction model in 1920s, and managed to carry out an 8 hour wind and pressure prediction by hand calculations. The results turned out to be very inaccurate; however, his effort marked the start of the era for modern numerical weather predictions [4].

Since that time, much progress has been realized in developing sophisticated and reliable NWP models. The technology of weather prediction has improved dramatically during the past decades as faster computers, better models, and much more data (mainly satellites) have become available.

Physically, the atmosphere is described as a fluid. Therefore, the idea of numerical weather prediction is to sample the state of the fluid at a given time and use the non-linear equations of fluid dynamics and thermodynamics to estimate the state of the fluid at some time in the future [20]. As there is no analytical solution for these complex and partial differential equations (PDE), the only possibility is to solve them numerically [4].

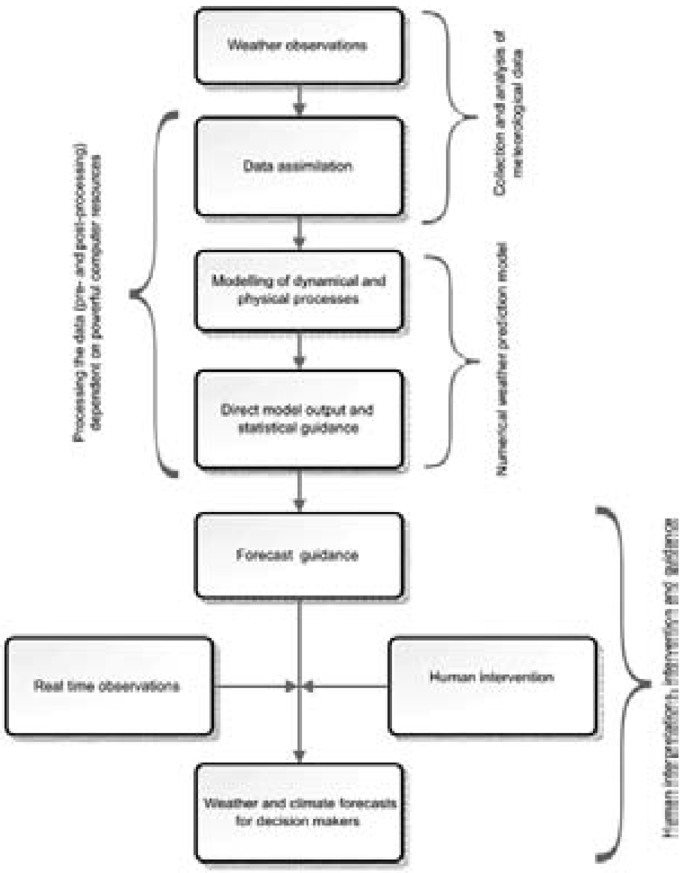


Figure 2.1: Weather Forecasting Procedure, (Source: [16]).

Basing on this idea, NWP is defined as the production of a forecast through the time-integration of a comprehensive set of mathematical equations that describe virtually all dynamical and physical processes in the atmosphere using numerical procedures [15]. In other words, it is the computational process of forecasting meteorological parameters such as temperature, wind and pressure by mathematical resolution of non-linear equations of these parameters diversity consequent with time and location. The behavior of atmosphere is convenient with physical rules which determined with mathematical equations. These physical rules are represented by a set of primitive equations which describe the atmospheric motions [21], by which, a NWP model comprises the changes of meteorological parameters such as temperature, wind speed and direction, humidity and pressure by the time.

For that, a 3D-dimensional grid has to be defined (figure 2.2) and the equations are discretized. With given initial conditions, the time integration of the governing equations allows to predict the new state after a predefined time step.

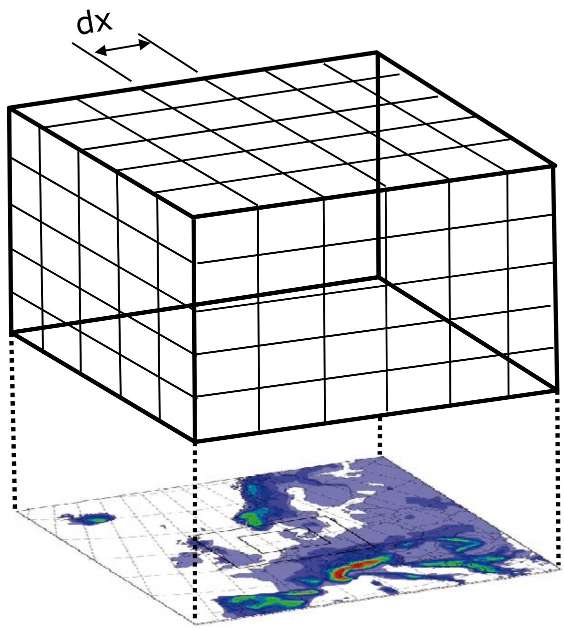


Figure 2.2: The principle of weather modelling: The atmosphere is fragmented into 3-D grid boxes. Topography and land type distribution are adapted to the model resolution (dx). Winds, heat transfer, solar radiation, relative humidity, and surface hydrology are calculated within each grid cell, and the interactions with neighboring cells are used to calculate atmospheric properties in the future. Sub-grid scale processes are parametrized, (Source: [16]).

Domain grid spacing in both directions, i.e, horizontal and vertical is of such importance in representing an atmospherical area in the used numerical model. The horizontal distance between grid points in a domain is generally known as the *spatial resolution*, or simply as domain resolution. As finer spacing between domain grid points, as more details are represented and resolved in the NWP, that is, some small-scale meteorological phenomena cannot be represented on coarser grids. However, the computational costs are higher, as more gird points are to be calculated within the same domain.

On the other hand, the vertical distance of atmospherical domains (vertical coordinate of the 3D-dimensional grid) is normally represented by a pressure coordinate system, in which the geopotential heights of constant-pressure surfaces become dependent variables. Actually, modern NWP models tend to use normalized pressure coordinates referred to as *sigma coordinates*, as such, pressure (P) levels are scaled by σ coordinates between the surface pressure of the domain (P_0), and the pressure in the top of the domain (P_T) [22]. That is, many pressure levels could be used between domain lower pressure to higher pressure layers.

In the next subsection, NWP process is discussed as well as prediction applications depending on domain resolutions.

2.1.1 NWP Models: Scales and Types

As it has been described previously, certain areas where atmospherical future state is to be predicted by NWP models, are represented by 3-D uniform-gridded-rectangles referred to as domains. The input data to these models, which describe an estimation of the actual state of the atmosphere, are called initial conditions. Those initial conditions (initial values of weather variables) are assigned to the points of the grid. Then, a NWP model is run on a high-speed computer, by which, new values of weather variables are produced each *time step* of model integration. Figure 2.1.1 illustrates the general steps for a numerical weather prediction process.

The spatial resolution of a NWP domain is used to describe the resolution of both the initial conditions and prediction results. Obviously using a finer resolution for the model grid will more accurately reflect the actual atmosphere, and thus, the prediction model will more accurately forecast the weather as more small-scale atmospherical processes can be represented and better representation of the topography is considered.

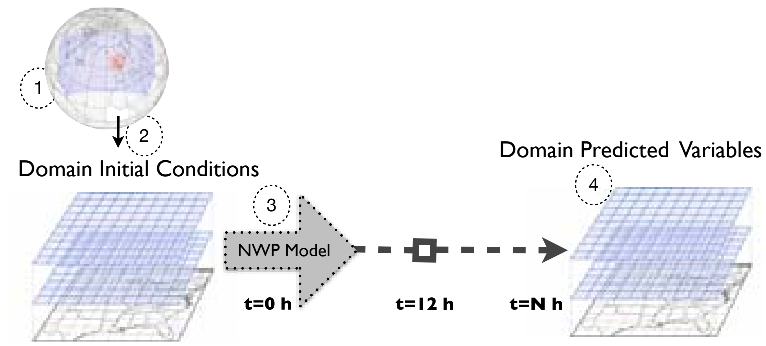
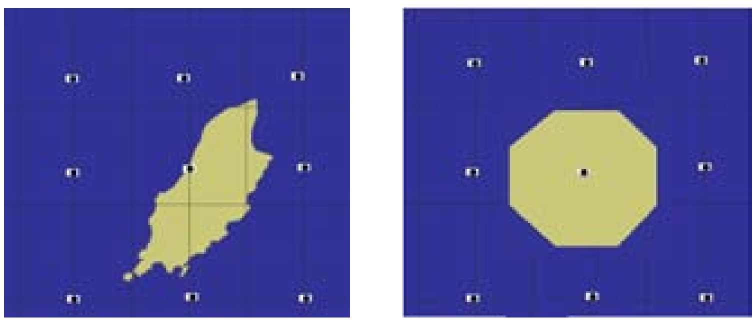


Figure 2.3: NWP: A simplified flow of the steps involved in a weather forecasting process.

For instance, suppose a weather forecast is to be conducted in The Isle of Man (an island located in the middle of the northern Irish Sea which covers an area of around 572 square kilometers. This island is 52 Kilometers long and, at its widest point, 22 Kilometers wide. If a NWP model is to be used to predict over this island, with a gird domain of 25 Kilometers of spatial resolution, one or two grid points would fall inside the island as depicted in Figure 2.4. Hence, for a more accurate forecast, a domain of higher resolution (less than 25 Kilometers) is needed to give a general outline of the island to the NWP model.



(a) (b)

Figure 2.4: The Isle of Man, forecast domain resolution of 25 Kilometers:
(a) with just on grid point over the island, the numerical model will consider the domain as in **(b)**. Source [23].

That is, domain spatial resolution is a crucial factor in determining the accuracy of a certain forecast. Actually, depending on the application and need, multiple types of NWP models are commonly used considering forecast domain resolution, as a result, different types of forecast products are delivered. In general, operational NWP models are commonly categorized depending on three related factors as follows (see [24]):

1. **Forecast Area**, or the coverage of the domain:
   - **global**: the whole, or approximately the whole planet.
   - **regional** (mesoscale): a certain region of the planet like North America or Europe, for example.
   - **local** (mesoscale): certain smaller domains, for example, Spain or eastern part of Spain.
2. **Resolution**, or the size of the grid: from few to over hundred of Kilometers ahead.
3. **Time Frame**:
   - **short-range**: from few hours up to 3 days.
   - **medium-range**: up to approximately 2 weeks.
   - **seasonal**: up to 12 months.

In contrast to global models, regional or mesoscale (also known as limited-area, or LAMs) models, produce shorter scale forecasts with much higher spatial resolution and more frequent output. Actually, this allows regional models (LAMs) to resolve explicitly smaller-scale meteorological phenomena that cannot be represented on the coarser grid of a global model. Generally, weather service centers in different countries and regions use in a daily basis mesoscale models to deliver short-range forecasts to their citizens or clients.

Worldwide, there are a couple of dozen NWP models being used, for instance, a relatively old, but illustrative comparison between various NWP models was presented in [25], where the performance six operational models where evaluated in short-range weather forecasts over the western United States of America. More recent works realized for the same reason can be found in [26, 27].

However, following the global, regional, and local categories described previously, listed bellow some of the mostly used NWP models:

Global Models :

- **GFS**: The Global Forecasting System, developed by the USA National Oceanic and Atmospheric Administration (NOAA). The GFS is a 35 to 70 km resolution, medium-range, global model [28].
- **IFS**: The Integrated Forecasting System, developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), normally referred to as the ECMWF model. ECMWF model is a 40 km resolution, medium-range, global model [29].

Regional and Local Models :

- **WRF**: The Weather Research and Forecasting model, developed cooperatively by the US National Center for Atmospheric Research (NCAR), the US National Centers for Environmental Prediction (NCEP), and the meteorological research community. WRF is one of the mostly used mesoscale NWP models for regional and local forecasts, which also serves for both operational forecasting and atmospheric research needs [7].
- **NAM**: the North American Mesoscale Model, which is used in the NCEP to serve for mesoscale short-range forecasts. Currently, the Weather Research and Forecasting Non-hydrostatic Mesoscale Model (WRF-NMM) model is run as the NAM [30].
- **MM5**: the Fifth-Generation Penn State/NCAR Mesoscale Model, a regional mesoscale model which was developed by Penn State university and the NCAR, however, although it is maintained and still operational in different weather centers, WRF model is considered as the successor of MM5, including the capabilities of MM5, WRF is generally known as a next-generation mesoscale model [8].
- **HIRLAM**: HIgh Resolution Limited Area Model, a highly-configurable, high-resolution system of short-range models, developed by the international HIRLAM programme (five Scandinavian and three Baltic countries, plus Ireland, the Netherlands, and Spain). It is operational in different weather centers in these countries [31].
- **ALADIN**: a high resolution short-range weather forecasting model, developed and operated by several European and North African countries [32].

A relatively old, but illustrative comparison between various NWP models was presented in [25], where the performance of six operational models where evaluated in short-range weather forecasts over the western United States of America. More recent works can be found in [26, 27]. Moreover, a complete taxonomy of operational NWP models up to 2002 could be found in [?].

During the discussion of the work presented in this thesis, the major focus is being oriented to enhance weather predictions for regional and local mesoscale NWP models (LAMs), and mainly, for short-range weather predictions, however, the presented work is also applicable for medium-range weather forecasts.

2.1.2 NWP Model Input/Output

The major steps of a numerical weather prediction process are generally classified into three phases:

- **Pre-processing**: which includes the process of data collection and assimilating observed data to the model.
- **Model Integration**: the execution of the numerical model itself, resolving the mathematical equations involved over computing resources.
- **Post-processing**: the generation of model outputs and graphics, which consecutively will be subject to human interpretation.

Subsequently, both Pre-processing and Post-processing phases are described in detail:

Pre-processing: Model Initialization

To make the forecast for a future time, the initial state of the atmosphere over the targeted domain must be provided, as an initial condition. These initial conditions (generally referred to as ICs) are assigned to all domain grid points to reflect the actual values of weather variables.

However, regional models use a global model outcomes besides the certain domain ICs, to specify conditions at the edge of their domain in order to allow the effects of the atmospherical motion coming from outside to be represented in the regional model calculations, these data is called normally the Boundary Conditions of the model (BCs) [33].

Commonly, initial conditions (ICs) are derived from both weather observations and previous model forecasts (pervious model forecasts are generally known as the *first guess*). These observations include many types of data and networks, which are generally classified as follows (see [33, 34]):

1- Surface Observations the most used set of observations used to initiate regional forecast models are surface observations, which are provided by weather stations at ground level over land and from weather buoys at sea (see figure 2.5.(a)). To unify the instrumentation used for getting observations in these stations, as well as the timing of these observations, the World Meteorological Organization (WMO) [35] works on standardizing this practice, as such, NWP models can be initialized by these surface observations worldwide.

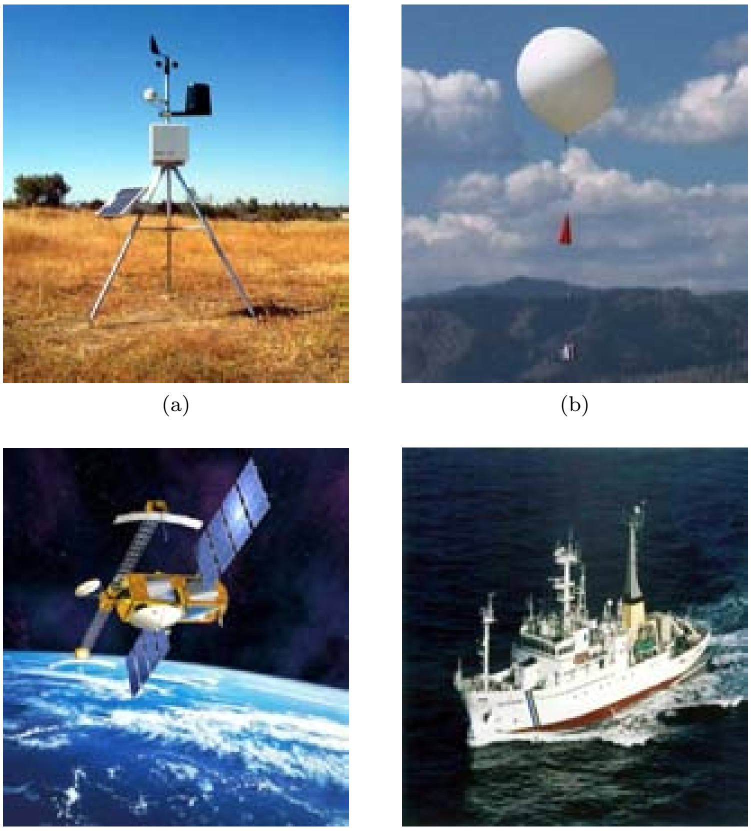
2- Upper-Air Observations another accurate and important data set used to initialize NWP models is upper-air observations. These observations are normally provided by a weather balloon called radiosonde (see figure 2.5.(b)), which measures the vertical profile of atmospheric variables and transmits them to a fixed receiver. Radiosondes are launched daily in large number of cities around the world, by which, various important weather variables are measured, like pressure, altitude, geographical position (latitude/longitude), temperature, relative humidity, wind speed and direction.

3- Satellite and Radar Imagery meteorological satellites and radars, especially meteorological satellites, are seen to be the dominant data source for providing initial conditions to NWP models (see figure 2.5.(c)).

4- Commercial Observations mainly coming from equipped and specialized aircrafts and ships in their routes (see figure 2.5.(d)). For example, some instrumented aircrafts fly in and around weather systems of interest gathering observations of weather variables, mainly wind related variables.

Besides all these gathered observations, NWP models are provided also by terrain maps, which are available at high resolutions globally. These maps help NWP models to better calculate atmospheric circulations depending on the topography of the area of interest. It should be mentioned furthermore, that certain weather fields (e.g., orography, surface roughness length over land, albedo and vegetation type), are kept constant in a forecast run (these fields are not changed by BCs).

By gathering observations which determine ICs of the targeted domain, and having boundary values (BCs) which aggregate the evolution of the atmospheric state at the lateral boundaries of the same domain to the NWP model, the data collection step is realized.



(a) (b) (c) (d)

Figure 2.5: Weather observation sources: **(a)** aweather station, **(b)** a radiosonde, **(c)** a meteorological satellite, and **(d)** meteorological ship.

However, another important step is to be done in the Pre-processing phase in order to start a correct forecast, that is, the assimilation of the gathered observations to fit in the scales of the targeted domain. Unfortunately, observations generally do include many types of errors such as instrumental and human errors, and moreover, they are irregularly spaced and distributed. Hence, it is needed to interpolate observations to grid points and also to insure that the various fields are consistent and physically plausible.

Because of this imperfectness in the observations, many techniques of

Objective Analysis and Data Assimilation (described in detail in Section 2.3.1) are applied in order to interpolate observations in domain grid points. These techniques basically combine the observed values of weather variables with a *first guess* prediction results (normally, a forecast for the same domain 6 hours earlier) to produce initial conditions for a NWP model.

Once the data assimilation process has taken place and the initial conditions are in their final form, the data is then sent to the NWP model to create a forecast. Model integration is the simulation process itself, which involves the execution of the numerical model on computers to calculate the evolution of the initial conditions of the area of interest.

Post-processing: Output and Visualization

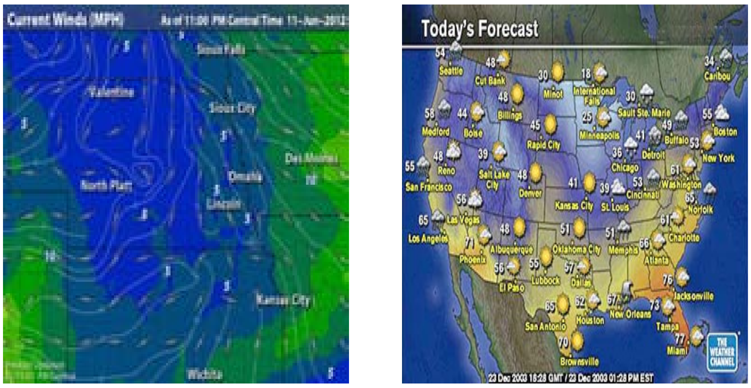
Once the NWP model is run, a forecast is produced as a result. Obviously, the results are the future values of the weather variables, calculated by the the model.

The Post-processing takes care of the output from forecast runs, including archiving in a suitable format. Field verification and verification against observations of selected meteorological fields are also parts of the Post-processing

On the other hand, managing the mass of forecast data created by the models is fast becoming a science of its own. Output from modern NWP models requires post-processing to make it intelligible and, most importantly, expert human interpretation in order to assess its meaning, qualities, and possible flaws. Since the first beginning of the use of modern NWP models in weather forecasting, visualization techniques and methods were being studied and developed [36]. Actually, without visualization techniques, human interpretation (which is still a fundamental step in weather forecasting) of the large amount of data produced by NWP models would not be possible.

Treinish, Lloyd A. discussed in [37] the importance and then, the need to develop effective visualization methods in order to better forecast the weather. Nowadays, many visualization products and methods are being used in practice, e.g., [38, 39, 40, 41, 42], as well, research continues massively towards improving this field.

Today, the level of detail in modern models allows for a wide variety of products and forecast fields to be delivered by different visualization techniques, for use not only in general meteorology, but also in specialized areas such as aviation and air quality. However, in general, weather and meteorological maps are the mostly common tools for visualizing weather forecasts, both for experts and also (simpler) for citizens and clients. An example of two weather maps is depicted in Figure 2.6.



(a) (b)

Figure 2.6: Weather forecast maps: **(a)** wind speed and direction map, **(b)** a simpler map delivered for citizens normally.

Summery: putting all together

As it has been described in the previous sections, the prediction process is composed of three major phases, Pre-processing, Model Integration, and Post-processing. Consequently, these three phases and their corresponding steps are summerized in Figure 2.7.

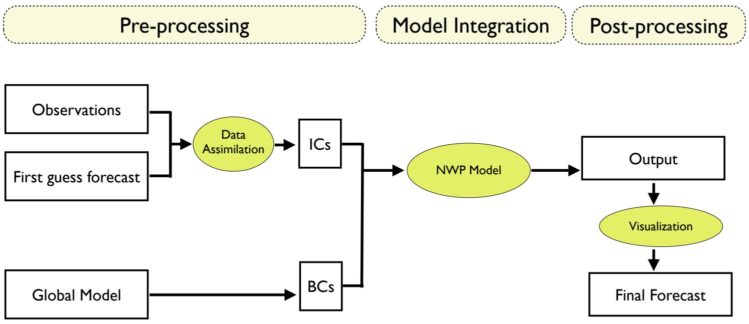


Figure 2.7: Complete numerical weather prediction process.

It should be mentioned as well, that both Pre-processing and Post-processing phases are as important as the NWP model processing itself, as such, a reliable weather prediction is usually relative to our capacity to reduce errors in all these three phases. Good model input, precise model, and good human interpretation, will lead to viable prediction results.

However, owing to the complexity of the weather system, and the nature of some meteorological processes which happen in small scales (usually smaller than the scales resolved by todays NWP models), more details must be considered to represent the effects of these small-scale processes on the scales resolved by the NWP model. For that, a process called *parameterization* is usually used to address this problem, in the next Section, this process and its applications are described.

2.1.3 Parametrization

In order to complete the treatment of the primitive equations of NWP models, it is necessary to consider the parameterization of the sub-grid scale processes that cannot be directly modeled yet. That is, some meteorological processes are too small-scale or too complex to be explicitly included in numerical weather prediction models.

For example, a typical cumulus cloud has a scale of less than 1 kilometer, and would require a grid even finer than this to be represented, while NWP models predicts normally on domains of grid-scales higher than 1 kilometer (see figure 2.8). Thus, parametrization is needed in such cases to represent this process on a certain domain scale.

Physical processes that are typically parameterized in modern NWP models are soil-vegetation processes, surface layer processes, turbulent exchange processes, micro-physics (cloud formation), convection and radiation. All these parameterizations are important because they have a strong influence on the skill of a weather forecast by interacting indirectly with each other by changing model variables.

Actually, the majority of these physical processes occur in scales less than 1 kilometer, hence, all these processes (between others) need to be parameterized [43, 44].

These physical processes play an important role in the atmosphere. And even in very high resolution models, physics on unresolved scales have important impacts on the evolution of the state of the atmosphere. Parameterizations are generally described as formulas (empirical or derived from physical hypothesis) which calculate the effect of sub-grid scale physics on the resolved scales by means of prognostic and diagnostic model variables [45].

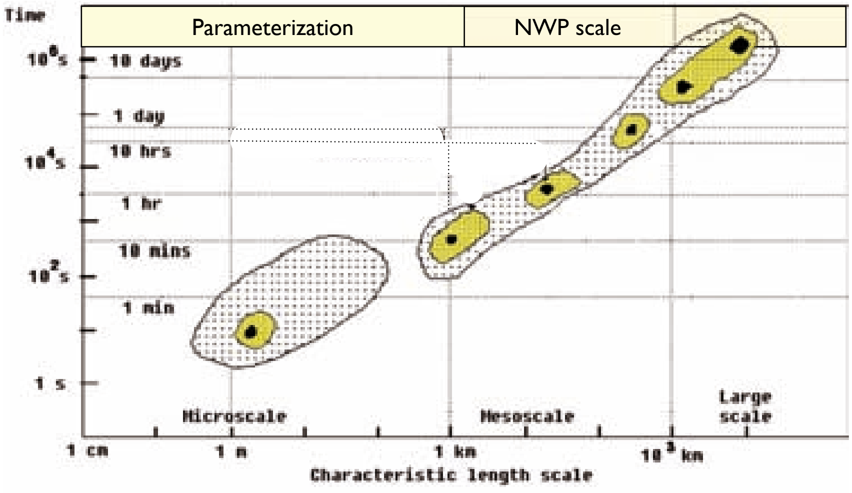


Figure 2.8: Atmospheric phenomena scales (spatial and temporal): modern NWP models of high spatial resolution are able to resolve weather phenomena which occur in scales down to few kilometers, however, there are lots of other important phenomena which are not resolved by NWP domain scales.

That is, either empirical formulas are applied to relate unresolved variability to explicitly simulated variables, or external information is needed to be inserted to the mesoscale model. Hence, parametrization enables the representation of these important meteorological processes by relating them to variables on the scales (the points of the gridded domain) that the model resolves, as such, the general concept of this process can be written as in Equation 2.1 (from [46]):

$$Output(\mathbf{x}) = \left[Input(\mathbf{x})\ \mathbf{y}\right] \tag{2.1}$$

Where the $Output(\mathbf{x})$ is a dependent variable that need to be computed, this variable is obtained from the input value $(\mathbf{x})$, and prescribed constant/constants $(\mathbf{y})$, through a transfer function $(\mathbf{T})$, which is the parameterization. The constant $(\mathbf{y})$ is obtained from limited observations, normally embedded within approximate models of physical processes that have been set to values that reproduce certain aspects of the certain observations.

The parametrization process itself, is an engineering module, i.e., it includes empirical equations with tunable coefficients that are derived form available (normally limited) observations. Hence, it is important to be aware of that parameterizations are simplified and idealized representations of complex processes and may therefore not always be appropriate. However, without parameterizations, NWP may not be useful because most subgrid-scale processes are key factors in weather forecasts that concern our daily life [47].

To enable modern NWP models to involve the effects of such subgrid-scale processes in the simulation process, many parameterization schemes (also known as models) have been developed and thus, coupled with operational NWP models.

David J. Stensrud in his book *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models* [46] put together all the categories of actual parameterization schemes and the entailed subgrid-scale processes resolved by them, as follows:

- **Land/water surface-atmosphere and soil/vegetation-atmosphere parameterizations**: also known as land surface modelling (LSM), these parameterizations intend to resolve the effects of the interaction between surface and the atmosphere, i.e., the exchange of surface water and energy fluxes at the soil-atmosphere interface, which depends highly on parameters like soil texture, vegetation type, soil moisture, land use, etc.

- **Planetary boundary layer and turbulence parameterizations(PBL)**: intends to resolve processes that happen in the lowest layer of the atmosphere where the wind normally is influenced by friction (a force that slows motion and dampens energy). Normally, surface friction from vegetation and topography causes turbulent eddies and chaotic wind motions.

- **Convection parameterizations (Cumulus)**: used in NWP models to predict the collective effects of (many) convective clouds that may exist within a single domain grid box.

- **Microphysics parameterizations**: to predict physical processes that lead to the formation, growth and precipitation of clouds. This process occurs in very small scales (the formation of water bubbles) and are impossible to be detected even by high resolution NWP models.

- **Radiation parameterizations**: these parameterizations intend to calculate the amount of solar radiation reaching ground level, also to calculate the surface flux of energy between the ocean and the atmosphere in order to determine realistic sea surface temperatures and type of sea ice found near the ocean s surface. That s done considering also soil type, vegetation type, and soil moisture, which all determine how much radiation goes into warming and how much moisture is drawn up into the adjacent atmosphere.
- **Cloud cover and cloud-sky radiation parameterizations**: which deal with calculations related to the amount of solar radiations reflected or absorbed by clouds (depending on location, type, size, and many other characteristics of clouds), as it is known that clouds reflect the solar short-wave radiation and they absorb the terrestrial long-wave radiation.
- **Orographic drag parameterizations**: which are used to represent atmospheric processes associated with orography.

Actually, all parameterizations are related to each other, that is, parameterizing land surface processes is affected by radiation parameterizations, which consequently, are affected by cloud parameterizations, and so on. Actually, this is due to the continuity nature of the atmospheric motion. A simplified scheme of direct interactions between parameterization processes is illustrated in Figure 2.9.

Nowadays, there are lots of different parameterization schemes to represent different small scale atmospheric processes in modern NWP models. Moreover, there are different schemes to represent the same class of subgrid-scale processes, as such, various parameterization schemes are developed basing on different assumptions and approximations (as well as variations in scheme constants and closure parameters) to represent the same physical processes.

For instance, in [49, 50], different operational Convection schemes (different physical schemes to represent convection process in atmosphere) are compared and evaluated in mesoscale weather predictions realized by NWP models. Many works are found which do comparisons also for other parameterization schemes, such as in [51] for different cloud-sky radiation parameterization schemes, and in [52, 53] for planetary boundary layer and turbulence parameterization schemes, and so on for the other parameterization schemes.

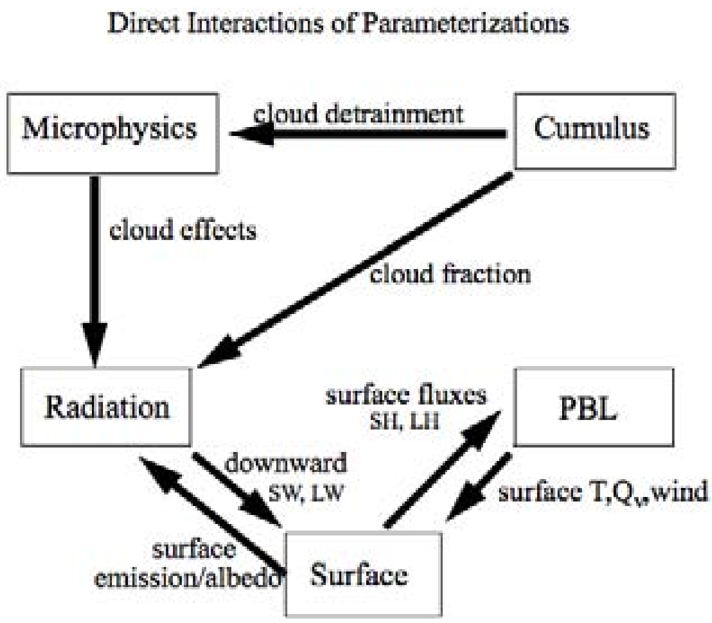


Figure 2.9: Interaction between parameterizations: SH is the surface sensible heat flux, LH is the surface latent heat flux, SW is the short wave radiation, LW is the long wave radiation, T is temperature, and Q_v is the water vapor mixed ratio. Source ([48]).

In fact, modern NWP models like WRF, provide the ability to select between different schemes for the same parameterization process. However, the overall conclusion on deciding wether to select one scheme or the other depends on different factors, as such, some perform better in certain regions with certain topographical characteristics (e.g., [54]), some perform better in a certain season, while others may be more stable in different weather conditions. Normally, weather prediction centers do large simulations over the same region for different seasonal periods, then they evaluate the quality of predictions resulted by applying different schemes, in order to select one or other scheme for their operational forecasts.

The presented work focuses mainly on Land/water surface-atmosphere and soil/vegetation-atmosphere parameterization, and their effects on some weather variables. Consecutively, more description about land surface parameterization process in numerical weather predictions is provided.

Land Surface Parametrization

Land surface parameterization in numerical weather prediction is of paramount importance. Actually, accurate forecasts of near surface weather variables are highly requested by the NWP users community. The quality of such variables as the near surface air temperature and humidity, winds, low level cloudiness and precipitation is to a large extent determined by the physical realism of the model representation of the surface-atmosphere interactions. For instance, a very wet soil on a sunny day gives rise to more evaporation, while a dry soil allows more solar radiation to warm the surface resulting in higher maximum temperatures.

Being that said, land surface is a critical component for the study of the weather, basically through its partitioning of solar radiation into sensible and latent heat fluxes, its redistribution of precipitation into evaporation, soil moisture, groundwater recharge, or runoff, and its regulation of biogeochemical cycles with processes such as photosynthesis and respiration. These water and energy exchanges between the atmosphere and the land surface are known to significantly impact atmospheric motion, which has motivated significant advancement in the understanding of the physical processes that govern these exchanges [55, 56, 57, 58].

On the other hand, as it has been shown in Figure 2.9, parameterization schemes exchange variables in all atmospheric layers, as a result of the continuos nature of the atmospheric motion. Thus, it is also crucial to well parameterize land surface processes in order to enhance the precision in the calculations realized by the other parameterization schemes (radiation, PBL, convection, cloud, etc.)

Zhang Ying in [59], stated that land surface processes are generally described in terms of physical fluxes and hydrological state of the land (depicted in figure 2.10):

The physical processes include:

- Radiative fluxes;
- Momentum flux; sensible and latent heat fluxes; partitioning of latent heat into canopy evaporation, soil evaporation and transpiration;
- Heat transfer in a multi-layer soil/lake/ocean.

The hydrological processes include:

- Snow accumulation and melt;
- Rainfall, interception, infiltration and runoff;
- Soil hydrology, including water transfer in a multi-layer soil.

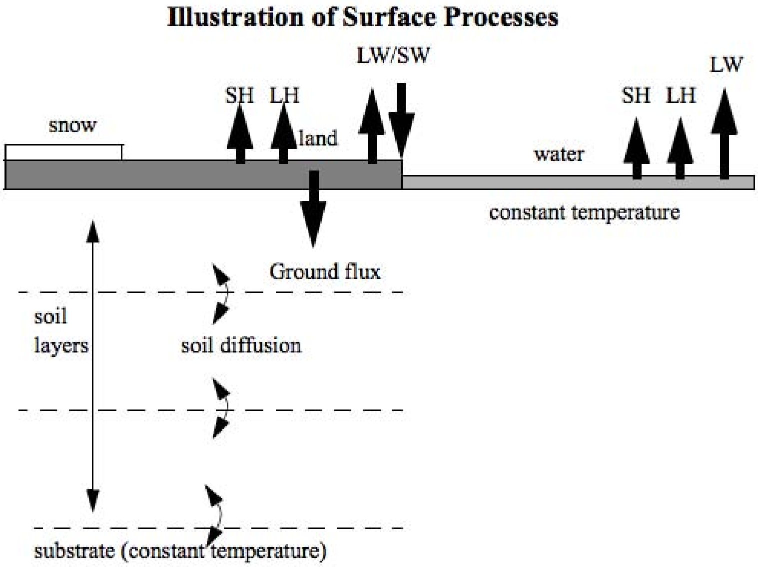


Figure 2.10: Summery of surface processes: SH is the surface/water sensible heat flux, LH is the surface/water latent heat flux, SW is the short wave radiation, and LW is the long wave radiation. Source ([48]).

Basing on physical hypothesis, many empirical formulas are derived and applied in land surface parameterization schemes (e.g., [60]). By applying these formulas, surface processes are simplified in order to represent their effect in NWP model scales.

These formulas depend fundamentally on surface characteristics (land use/vegetation types and soil texture) for each domain, as such, the heat/radiation transfer between the surface and the first layer of the atmosphere (the planetary boundary layer) depends basically on vegetation and soil related parameters. Actually, The vegetation type and the soil texture/cover are the primary variables to decide the land surface land surface characteristic fields [61, 55].

But the land surface parameterization process is even more complex, as NWP models normally run over domains composed of heterogeneous surface features (different vegetation types and soil textures). That is, it is crucial to consider this heterogeneity in parameterization processes. For that, modern NWP models use standard and global maps of vegetation and soil types in order to define the category of land use (vegetation) and soil texture for each surface grid point within the domain of interest.

For vegetation types and features, there are different categorization standards, known as land use categories or vegetation types (see [62]), which include the Global Ecosystems, the Simple Biosphere Model, and the U.S. Geological Survey Land Use/Land Cover System (USGS). An example of land cover categories map is depicted in Figure 2.11.

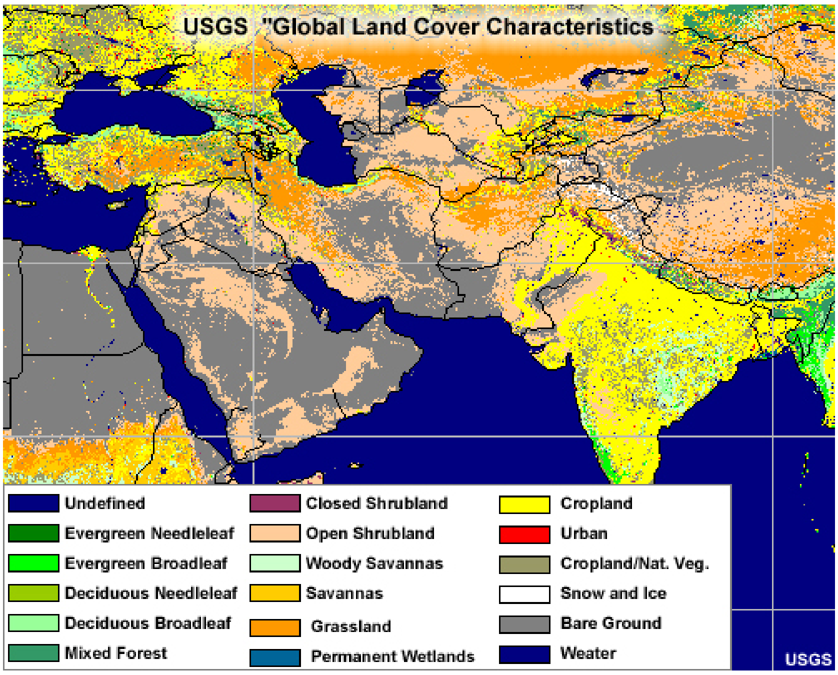


Figure 2.11: Land cover categories map, from USGS.

The surface features of each category are coefficients and parameters used by different land surface parameterization schemes, depending on their physical assumptions, by which, the heat/radiation exchanges with the atmosphere are calculated. These parameters (their values are relative to each land cover category) include: albedo and coefficients of long and short wave radiations, the vegetation height, the Noilhan parameter for the dependence of canopy stomatal resistance from solar radiation, the minimum stomatal resistance, the winter value of leaf area index, etc. (see [63, 59]).

The same happens for soil texture categorization, as such, there are different global standards and datasets which categorize soil textures according to their contents [64]. Figure 2.12 shows the major categories of soil texture types, where each category is defined according to its contents of the three major soil classes: silt, clay, and sand.

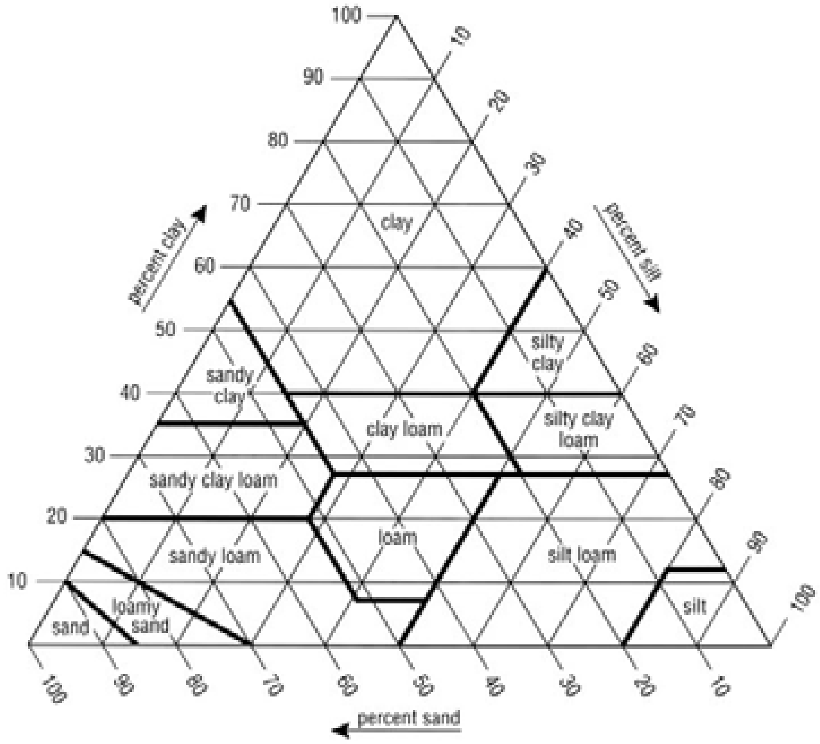


Figure 2.12: Soil texture categories, from the USDA-NCRS (1997) soil textural database.

As for vegetation categories, there are various coefficients and parameters (their values are relative to each soil texture category), which characterize each soil texture features, these include: the saturated moisture potential, a reference soil moisture, soil diffusivity, soil conductivity, etc. (see [59]).

The parameters which belong to each category of land use and soil texture, are assigned to different values, depending on the category to which they belong. Actually, NWP models are provided by standard tables, which include the default values of these parameters, depending on their category.

Normally, prediction domains exhibit heterogeneity in their surface characteristics. That is, the terrain of a certain domain (the first mesh of the 3-D grid) could include different vegetation types, and different soil textures.

For that, a prediction process, which uses land surface parameterization scheme, proceeds as follows: NWP models start a process of prediction over a certain zone using the initial and boundary conditions defined by their location (longitude, latitude and vertical distance) for each grid point of the domain.

Additionally, NWP models are also provided by terrain maps, which are available at high resolutions globally. These maps define the surface and topographical characteristics of the targeted domain, as such, the first mesh grid points of the domain are assigned with a number indicating its *landuse* category (LU-index) and with another number indicating its *soil texture* (SLTYP).

During the prediction process, the NWP model needs surface parameter values for each surface grid point in order to calculate the evolution of the other weather variables. These parameter values depend on their categories, and for each category, the NWP model is provided by its default parameter values which are provided in stand-alone tables like those shown in Figures 2.13 and for *landuse* parameters, and in Figures 2.15 and for *soil texture* parameters.

Then, for each surface grid point, the NWP model reads its assigned *landuse* category (LU-index) and, goes to LAND USE table to obtain the values of the surface physical parameters corresponding to that category. The process is done for all surface grid points and the same is done with soil texture parameters.

Albedo - SFC albedo (in percentage)	RGL - Parameter used in radiation stress function
Z0 – Roughness Length (m)	HS - Parameter used in vapor pressure deficit
SHDFAC - Green vegetation fraction	SNUP - Threshold depth for 100% snow cover
NROOT - Number of root layers	LAI - Leaf area index (dimensionless)
RS - stomatal resistance (s m-1)	MAXALB - Upper bound on max albedo snow

Figure 2.13: *Landuse* parameter description, (from [65]).

Category Type	Class	BB	DRYSMC	F11	MAXSMC	REFSMC	SATPSI	SATDK	SATDW	WLTSMC	QTZ
Sand	1	2.79	0.010	-0.472	0.339	0.236	0.069	1.07E-6	0.608E-6	0.010	0.92
Loamy Sand	2	4.26	0.028	-1.044	0.421	0.383	0.036	1.41E-5	0.514E-5	0.028	0.82
Sandy Loam	3	4.74	0.047	-0.569	0.434	0.383	0.141	5.23E-6	0.805E-5	0.047	0.60
Silt Loam	4	5.33	0.084	0.162	0.476	0.360	0.759	2.81E-6	0.239E-4	0.084	0.25
Silt	5	5.33	0.084	0.162	0.476	0.383	0.759	2.81E-6	0.239E-4	0.084	0.10
Loam	6	5.25	0.066	-0.327	0.439	0.329	0.355	3.38E-6	0.143E-4	0.066	0.40
Sandy Clay Loam	7	6.66	0.067	-1.491	0.404	0.314	0.135	4.45E-6	0.990E-5	0.067	0.60
Silty Clay Loam	8	8.72	0.120	-1.118	0.464	0.387	0.617	2.04E-6	0.237E-4	0.120	0.10
Clay Loam	9	8.17	0.103	-1.297	0.465	0.382	0.263	2.45E-6	0.113E-4	0.103	0.35
Sandy Clay	10	10.73	0.100	-3.209	0.406	0.338	0.098	7.22E-6	0.187E-4	0.100	0.52
Silty Clay	11	10.39	0.126	-1.916	0.468	0.404	0.324	1.34E-6	0.964E-5	0.126	0.10
Clay	12	11.55	0.138	-2.138	0.468	0.412	0.468	9.74E-7	0.112E-4	0.138	0.25
Organic Material	13	5.25	0.066	-0.327	0.439	0.329	0.355	3.38E-6	0.143E-4	0.066	0.05
Bedrock	15	2.79	0.006	-1.111	0.20	0.17	0.069	1.41E-4	0.136E-3	0.006	0.07
Land ice	16	4.26	0.028	-1.044	0.421	0.283	0.036	1.41E-5	0.514E-5	0.028	0.25

Figure 2.14: *Landuse* categories, used in the WRF forecasting model, (from [65]).

BB – Function of Soil type	SATPSI - SAT (saturation) soil potential
DRYSMC - dry soil moisture threshold (volumetric)	SATDK - SAT soil conductivity
F11 - Soil thermal diffusivity/conductivity coef.	SATDW - SAT soil diffusivity
MAXSMC - MAX soil moisture content (porosity), Volumetric	WLTSMC - Wilting point soil moisture(Volumetric)
REFSMC - Reference soil moisture (field capacity), Volumetric	QTZ - Soil quartz content

Figure 2.15: textitSoil parameter description, (from [65]).

Category	Class	Albedo	Z0	SHDFAC	NROOT	RS	RGL	HS	SNUP	LAI	MAXALB
Urban and Built-Up Land	1	0.15	1.00	0.10	1	200.	999.	999.0	0.04	4	40
Dryland Cropland and Pasture	2	0.19	0.07	0.80	3	40.	100.	36.25	0.04	4	64
Irrigated Cropland and Pasture	3	0.15	0.07	0.80	3	40.	100.	36.25	0.04	4	64
Mixed Dryland/Irrigated Cropland and Pasture	4	0.17	0.07	0.80	3	40.	100.	36.25	0.04	4	64
Cropland/Grassland Mosaic	5	0.19	0.07	0.80	3	40.	100.	36.25	0.04	4	64
Cropland/Woodland Mosaic	6	0.19	0.15	0.80	3	70.	65.	44.14	0.04	4	60
Grassland	7	0.19	0.08	0.80	3	40.	100.	36.35	0.04	4	64
Shrubland	8	0.25	0.03	0.70	3	300.	100.	42.00	0.03	4	69
Mixed Shrubland/Grassland	9	0.23	0.05	0.70	3	170.	100.	39.18	0.035	4	67
Savanna	10	0.20	0.86	0.50	3	70.	65.	54.53	0.04	4	45
Deciduous Broadleaf Forest	11	0.12	0.80	0.80	4	100.	30.	54.53	0.08	4	58
Deciduous Needleleaf Forest	12	0.11	0.85	0.70	4	150.	30.	47.35	0.08	4	54
Evergreen Broadleaf Forest	13	0.11	2.65	0.95	4	150.	30.	41.69	0.08	4	32
Evergreen Needleleaf Forest	14	0.10	1.09	0.70	4	125.	30.	47.35	0.08	4	52
Mixed Forest	15	0.12	0.80	0.80	4	125.	30.	51.93	0.08	4	53
Water Bodies	16	0.19	0.001	0.00	0	100.	30.	51.75	0.01	4	70
Herbaceous Wetland	17	0.12	0.04	0.60	2	40.	100	60.00	0.01	4	35
Wooded Wetland	18	0.12	0.05	0.60	2	100.	30.	51.93	0.02	4	30
Barren and Sparsely Vegetated	19	0.12	0.01	0.01	1	999.	999.	999.0	0.02	4	69
Herbaceous Tundra	20	0.16	0.04	0.60	3	150.	100.	42.00	0.025	4	58
Wooded Tundra	21	0.16	0.06	0.60	3	150.	100.	42.00	0.025	4	55
Mixed Tundra	22	0.16	0.05	0.60	3	150.	100.	42.00	0.025	4	55
Bare Ground Tundra	23	0.17	0.03	0.30	2	200.	100.	42.00	0.02	4	65
Snow or Ice	24	0.70	0.001	0.00	1	999.	999.	999.0	0.02	4	75

Figure 2.16: *Soil* categories, used in the WRF forecasting model, (from [65]).

2.2 Weather Predictability and Sources of Error in NWP Models

The term predictability in NWP modelling, refers to the extent to which the future state of the atmosphere or a specific weather system may be predicted based on current ability of NWP models [**?**]. As such, to be viable, a NWP model must integrate an understanding of many different phenomena and their interactions: how the wind blows; how heat is received from the sun and transformed by the oceans, the ground, the air, and the clouds; how water vapor condensates into clouds and how droplets of water turn to rain, ice and snow; how friction near the ground mixes the lower layers of air. Thus, errors in handling one type of phenomenon can contaminate other parts of the model, or amplify errors in other model sub-systems.

By seeking further to analyze the nature of the atmospheric motion itself, atmospheric predictability research could be traced very early in a paper published over a century ago by **Bjerknes** in 1904 [18]. In that paper, he considered the problem of weather prediction from the standpoint of mechanics and physics, and proposed it as a deterministic initial value problem based on the physical laws such as the conservation of mass, momentum, and energy [66].

Later on, concretely in 1963, **Edward Lorenz** discussed his theory about the *chaotic nature of atmospheric flow* [5, 67], the so-called *chaos theory* in nonlinear dynamic systems. **Lorenz** stated that in such nonlinear dynamic systems (the atmospheric system), slightly differing initial states can evolve into considerably different states within a limited time. That is, the chaotic nature of the atmosphere determines that the predictability does not depend only on the realism of the model and the accuracy of initial conditions, but also on the system itself. Atmospheric motion, as a nonlinear dynamic (unstable) system, is supposed to have finite limit predictability.

In Figure 2.17, the problematic nature of chaos is illustrated by showing the motion trajectories of stable and unstable dynamic systems. As shown in the figure, even with two very close initial states (initial conditions) for a nonlinear dynamic system, results would have markedly different outcomes.

Unfortunately, besides the limited predictability of the atmospheric motion owing to its chaotic nature, atmospheric initial conditions are never known perfectly, what is generally known as the uncertainty problem of initial conditions.

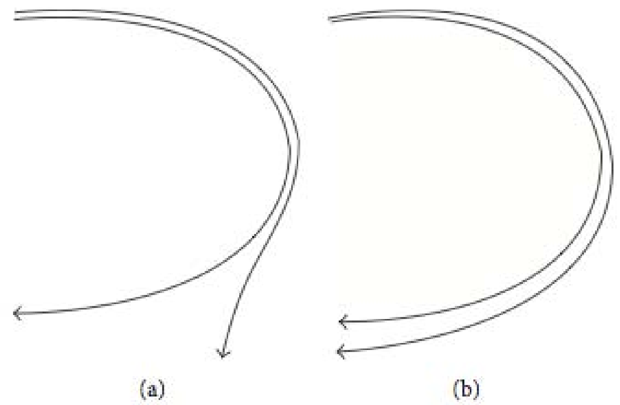


Figure 2.17: The evolutions of slightly different initial states (a) unstable trajectories; (b) stable trajectories (courtesy of Lorenz [67]).

Moreover, the numerical representations of the fluid dynamical equations involved in modern NWP models, depend on approximations and mathematical assumptions, which involve huge amount of model coefficients and tunable parameters, resulting in producing more uncertainties that limit the predictability of the evolution of the atmosphere. In particular, the uncertainty due to the parameterizations of sub-grid-scale physical processes is known to play a crucial role in prediction quality (e.g., [6]). Prediction errors caused by the uncertainty in physical parameterizations is commonly referred to as model errors.

As it has been stated previously (cp. Section 2.1.2), weather forecasting process has three basic steps: data collection (observations) and assimilation of observed data into initial conditions to be used by a numerical model, model integration to project the initial state into future, and the application of the forecasts to real world situations. However, the usefulness of any model depends mainly on the accuracy and reliability of its output.

In NWP models, Intrinsic uncertainties are introduced at each of those steps during a forecast process. In deed, numerical weather prediction is, by its very nature, a process that has to deal with uncertainties [68]. That is, the uncertainty is involved in all the phases of prediction process, for example, instrumental and human error introduced during the process of collecting data; errors introduced during data assimilation process due to mathematical assumptions and abstractions; imperfect model physics (approximations of real world such as parameterization of sub-grid processes) and numeric (e.g., discontinuity or truncation); and differences in human (both forecasters and end-users) s interpretation and decision to a same forecast depending on situations (who, what, when and where). All these kind of errors are intrinsic, unavoidable and sometimes even unknown to us in a real world operation.

Summarizing, sources of uncertainties in NWP models are classified as follows ([9]):

1. Initial conditions uncertainties
   (a) imprecision in specifying the boundary and initial conditions that impact the output variable values
   (b) imprecision in measuring observed output variable values.
2. Model uncertainties.
   (a) uncertain model structure and parameter values.
   (b) variability of observed input and output values over a region smaller than the spatial scale of the model.
   (c) variability of observed model input and output values within a time smaller than the temporal scale of the model. (e.g., rainfall and depths and flows within a day).
   (d) errors in linking models of different spatial and temporal scales.
   (e) uncertain model structure and parameter values.
3. Numerical errors.
   (a) errors in the model solution algorithm.

Over the last two decades, significant progress has been made in developing methods to reduce uncertainty in initial conditions, model development, and model diagnostics to enhance the predictability of weather systems. To have a deeper idea about the achieved progress in prediction quality over the last decades, figure 2.18 shows the evolution of mean forecast skill at the European Centre for Medium-Range Weather Forecasts (ECMWF) for the northern and southern hemispheres for the period of 19812004. Forecast skill is calculated with some measures of forecast errors defined by the difference between the forecast and the initial conditions estimated from observed data [66].

In spite of the significant progress realized in the field of weather prediction, many crucial demands are still evolving to be resolved by the scientific community, which include the necessity to have more reliable and precise predictions, over larger time scales, and higher resolution domains.

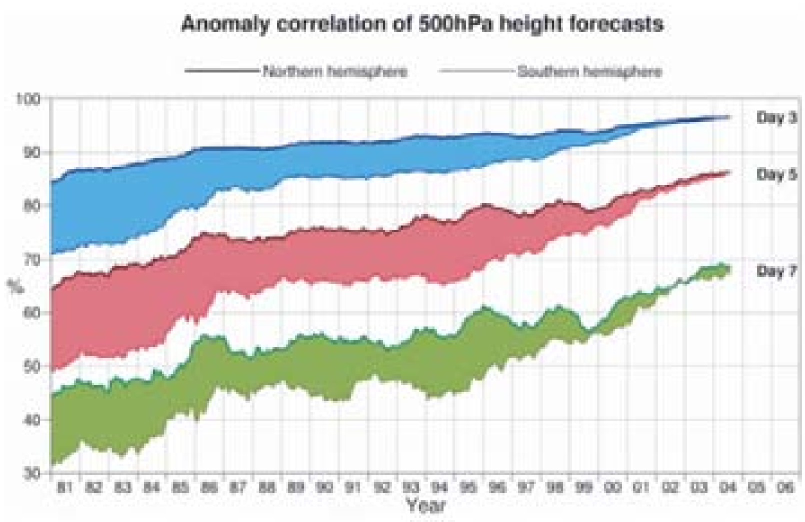


Figure 2.18: Evolution of mean forecast skill for the extratropical northern and southern hemispheres for the period of 19812004. Shading shows differences between hemispheres in anomaly correlation of 3, 5, and 7-day ECMWF 500-hPa height forecasts (from [66]).

While the main efforts are being oriented towards reducing the uncertainty in initial conditions and parameterizations, recent developments in ensemble forecasting and data assimilation have proved that there are promising ways to beat the forecast uncertainties [69]. These methods are described through the next sections consecutively.

2.3 NWP Enhancement Methods

Weather predictability has been a long-standing problem during the last decades. The advent of computers in the 1950s and their subsequent development in the 1960s,1970s along with the recent advance in parallel processing computers in the early 1990s, led to more and more numerically accurate representations of the fluid dynamical equations, which resulted in the development of sophisticated NWP models.

On the other hand, many studies were also devoted to improve initial conditions through advances in observing systems, the development of atmospheric data assimilation and ensemble forecasting techniques. On average, the accuracy of weather forecasts from such models improved steadily over these decades [70, 4].

That is, along with the realized advances in the development of modern and skillful NWP models in the last years, enormous efforts have been done also in order to reduce the effect of the other source of errors which limits weather predictability: the uncertainty in initial conditions. For this reason, two major strategies were developed; the first aims at reducing the uncertainty in initial conditions by data assimilation techniques, while the other aims at representing these uncertainties in the forecasting system, referred to as ensemble forecasting system. Consecutively, both strategies are described in more detail.

2.3.1 Data Assimilation

Data assimilation in atmospheric sciences started from the fact that NWP is an initial value problem. That is, If the initial state of the atmosphere was known accurately at a certain time point at the recent past or at the actual time (a complete and accurate specification of the three-dimensional (3D) structure of the initial values of a considered system), a forecast could be obtained by integrating the numerical model equations forward, from the past/present time into the future.

In practice, however, the state of the atmosphere is never known correctly, it is normally obtained through observations that are distributed non-uniformly in space and time, and they have different structures of random error. Having this limitation in the availability and quality of real observations, before producing a forecast, an optimization calculation must be performed to combine these irregular observations to generate the initial conditions that are distributed on regular model grids at the chosen time point, that are most consistent with the observations that have been made during the past few hours. This calculation is widely known as the data assimilation problem [71].

In other words, data assimilation proceeds by cycles. In each cycle, observations of the current (and possibly, past) state of the weather are combined with the results from a NWP model (denoted first guess, background, or prior information) to produce initial conditions (referred to as analysis), which are considered as the best estimate of the current state of the wether. Essentially, this step is done to balance the uncertainty in the observed data and in the forecast of the first guess. Then, these estimated initial conditions (analysis), as a result of data assimilation, are injected to the NWP to calculate the future state of the weather, the predicted future state of the weather is also considered as a first guess forecast or the analysis of a future prediction cycle.

Recent advances in many aspects of data assimilation and observing systems provide the opportunity for making substantial improvements in forecast skill. These advances include [72]:

- greatly increased availability of data, especially satellite and aircraft observation;
- the advent of adaptive observational techniques;
- and the improvements in assimilation algorithms, both in terms of their use of remotely-sensed observations and of their formulation

Being that said, the development of data assimilation methodology has mainly experienced three stages: simple analysis, statistical or optimum interpolation, and variational analysis dataAssimilation [10]. Nowadays, however, variational methods, in particular the three- and four-dimensional variational data assimilation (3D-Var)/4D-Var)) are the most used data assimilation techniques in the context of weather forecasting. Both approaches attempt to combine observations and background information in an optimal way to produce the best possible estimate of the model initial state. This technique not only has broad applications for the assimilation of atmosphere and ocean, but also can be used for many other applications in numerical weather prediction.

Actually, most weather prediction centers worldwide use data assimilation techniques to improve forecast skill in their predictions, especially 3D-Var and 4D-Var. Examples of operational 3D-Var and 4D-Var data assimilation techniques could be found in [73, 74, 75, 76, 77]. It should be mentioned here, that the major difference between the two methods is that in the case of 4D-Var, in addition to the assimilated three spatial dimensions of the NWP model domain variables, a fourth time dimension is added, as such, assimilation process is a continuous or dynamic process rather than discrete as in 3D-Var approach.

However, it should be noted that these approaches focus mainly on enhancing the estimate of the initial conditions of numerical weather models. That is, they do not deal with uncertainty problem related to the parameterization of subgrid-scale processes.

2.3.2 Ensemble Prediction Systems

Unfortunately, uncertainties always exist in both initial conditions and numerical models. Thus, reducing forecast errors caused by theses uncertainties remains a large area of research and operational implementation. In view of the uncertain properties of the atmospheric system, a theory of stochastic dynamic prediction was proposed by Epstein [78]. In his work, he described that the atmospherical motion is stochastic ; its behavior is non-deterministic. Hence, in a stochastic context, the initial and forecast states of the atmosphere must be represented as probability distributions. Basing on this idea, the developments in numerical weather prediction have led to current forecast systems that use an Ensemble Prediction System (EPS) approach to assess the probability of occurrence of possible forecast outcomes. Basically, with the following idea:- instead of using only one model with a single set of initial conditions, a group of forecasts with slightly different initial conditions (perturbed initial conditions) are made in an ensemble forecast. That is, an EPS is a collection of individual forecasts (forecast members) made from slightly different initial conditions and/or model parameters. The spread of the forecast member outcomes, defined as the standard deviation of the members from the ensemble mean, gives an estimate of EPS uncertainty.

So, estimates of the forecast uncertainty, for any forecast variable at any geographical location, are described by the probability density function (PDF) produced by a frequency distribution based on the various ensemble members [9].

Additionally, the mean of all outcomes for a certain forecast variable, is considered to be the best representative prediction.

The advances in parallel processing computers in the las decades, has led to operational EPS in some of the relative principal scientific centers in the field of weather prediction. In fact, EPS today, is considered as the most successful and merely operational enhancement approach, to deal with the uncertainty problem in weather predictions. This approach is being used as the principal prediction scheme in global weather centers like the European Centre for Medium-Range Weather Forecasts (ECMWF), U. S. National Centers for Environmental Prediction (NCEP), and the Meteorological Service of Canada (MSC).

Ensemble forecasting is being feasible in such centers due to the availability of huge computing power. However, in large number of weather services, EPS could be a non-adequate solution (if feasible at all) due to limitations in the available computing power, or if available, the corresponding limitations imposed to limit the number of ensemble members for a prediction, will minimize the needed precision in prediction results.

It should be noted that, EPS mainly is designed to include perturbations in the initial conditions, assuming that the error growth due to model deficiencies is small compared to that due to unstable growth of initial errors. However, in reality, uncertainties in model physical parameterizations cannot be ignored, since it has been realized that there is also a stochastic nature of physical parameterizations in weather prediction and the predictability is sensitive to variations in physical parameters.

Unfortunately, it has not been straightforward to develop theoretically sound, and also practical, formulations for how to insert parameterization uncertainty into ensemble development [12, 13]. Actually, some recent works were presented in order to deal with the uncertainty problem associated to model physics or physical parameterizations, e.g., [79, 80], however, EPS still need more efforts towards developing operational methods to deal with the uncertainties associated to NWP model closure parameters.

On the other hand, the gained forecast skill by EPS depends on one hand, on the way how perturbations are done in the initial conditions, and wether they reflect an acceptable distribution of probabilities, and on the other hand consequently, on the number of ensemble members by which a forecast is conducted. Indeed, the second issue depends on the available computing resources, that is, as more ensemble members are to be executed, more computing power is needed.

Chapter 3

Genetic Ensemble for Weather Prediction Enhancement

In the previous chapter, a study of numerical weather prediction modelling and methods were presented. Additionally, it has been discussed that NWP models suffer from the uncertainty of their initial conditions and input parameters. Actual operational enhancement methods for more reliable weather predictions were also outlined which are mostly dealing with reducing the uncertainty of model input data and parameters to assure better prediction results.

It should be highlighted however, that in the last 20 years or so, the major efforts to improve forecast skill in NWP models have been focusing on reducing the uncertainty in model initial conditions, mainly, by the advent of data assimilation techniques and ensemble forecasting methods.

On the other hand, as we discussed before (cp. Section 2.2), parameterization process is known to be a crucial factor in enhancing weather predictions. However, this process involves the processing of significant amount of tunable parameters and coefficients (denoted by *"model intrinsic or closure parameters"*), where NWP output is highly sensitive to their values. Currently, the values of these parameters are specified manually, as it has not been straightforward to develop theoretically sound, and also practical automatic approaches to find best suitable values of these parameters related to a certain prediction targeted domain [81].

In this work, we discuss a new methodology in the context of numerical weather prediction, which tackle this critical problem, in order to enhance quality of weather predictions as an ultimate goal.

Our hypothesis is that the NWP model forecast skill is sensitive to the specified parameter values. And thus, by finding `optimal values of these parameters, we aim to enhance prediction quality.

Basing on the concepts of EPS systems, in this chapter, the proposed methodology to enhance short-range and medium-range weather predictions is introduced. Basically, the proposed method which is called *"Genetic Ensemble for Weather Prediction Enhancement"* is based on developing an automatic and effective way to find optimal values of NWP model intrinsic/closure parameters in order to reduce the error produced in real predictions. The introduced methodology uses evolutionary computing methods, particularly, Genetic Algorithms in order to find the most timely `optimal values of these parameters, which appear in physical parametrization schemes that are coupled with NWP models.

On the other hand, the proposed methodology offers two different alternatives for the prediction process, a deterministic approach, where just one single forecast is to be conducted, and an ensemble prediction, where a set of forecasts (an ensemble) are used for a prediction. Furthermore, the proposed methodology is developed as a parallel application, which intends to overcome the cost and feasibility problems imposed by operational enhancement techniques, by reducing the overall execution time of the prediction process.

Firstly, a discussion concerning the subject of parameter calibration approaches used to find optimal input parameter values in environmental models in general is provided. Moreover, the classical weather prediction scheme is highlighted, then, the proposed weather prediction scheme is discussed, along with a detailed description of the implemented evolutionary computing technique in the process of searching for best possible values of model parameters.

3.1 Related Work

In the previous chapter (cp. Section 2.3.2), the concept of Ensemble Prediction System (EPS) was outlined. Actually, EPS predictions are being, nowadays, the most successful and merely operational enhancement methods, to deal with the uncertainty problem in weather predictions. However, EPS still need more efforts towards developing operational methods to deal with the uncertainties associated to NWP closure parameters.

Besides that, Ensemble forecasting normally requires huge computing power, actually, its a practice of running multiple number of forecasts to get their outcome means and spreads as a probabilistic prediction. However, in large number of weather services, EPS could be a non-adequate solution (if feasible at all) due to limitations in the available computing power, or if available, the corresponding limitations imposed to limit the number of ensemble members for a prediction, will minimize the precision in prediction results.

In this work, the proposed prediction scheme intends to enhance forecast skill in short and medium range weather predictions, by reducing the uncertainty of NWP closure parameters. As well, we intend also, to reduce the cost of this process compared to classical EPS solution for improving weather forecast skill. In the next subsection, parameter calibration approaches in environmental models are discussed.

3.1.1 Parameter Calibration Approaches

The problem of uncertainty in the modelling and simulation process is often overlooked. No model is a perfect representation of reality, so it is important to ask how imperfect a model is before it is applied for prediction. The scientific community relies heavily on modelling and simulation tools for forecasting, parameter studies, design, and decision making.

However, because almost all modelling systems and simulations relies widely on abstractions, idealization, and many assumptions. There exists a crucial need to what is generally know as parameter adjustment or calibration of models and simulation systems [82].

Model calibration [83] could be defined generally as the task of adjusting an already existing model to a reference system or, to a certain trusted reference model. This is usually done by adjusting the internal parameters of the model according to input-output sets of the system. Thus, in order to get trustworthy results from the model, input-output pairs of the model are fine-tuned to input-output samples of the reference system.

In various modelling systems, adjusting or calibrating model parameters is so crucial and necessary because of the difficulty (if possible in some models) to have timely accurate measures of the values of some model parameters, which are in complex modelling systems could be found in large amounts. As these parameters reflect some kind of abstraction and assumption in the basics of building simulation systems, they are normally tunable in order to enable the interested community to better-tune them according to the input-output pairs of a certain model.

Calibration is a far-reaching term and can mean quite different things to different people. This work however, deals only with a specific form of model calibration which is actually a special case of inverse problem analysis, in that the objective is to use observations of the simulator output to make inference about simulator inputs.

This type of calibration analysis poses several problems in practice:

1. Due to the complexity of weather modelling, there exist an elevated number of model input parameters which belong to standard ranges of possible values.

2. The simulation for higher-resolution models is often expensive, rendering an exhaustive exploration of all model parameter spaces is highly expensive if possible at all.

3. Various ranges and/or combinations of input parameters may yield comparable ts to the observed data.

4. Manual steering, tuning, or adjusting the large number of combination probabilities of NWP model parameters is mostly infeasible, not practical, and not reliable.

Considering these mentioned problems in the context of NWP model parameters, an automatic calibration method could be considered obvious and a strong contribution to the interested community of weather predictions. Actually, automatic calibration methods (independent of manual interference) have been recognized early and have become more and more reliable in the last decade [84].

Fortunately, there are in practice many approaches for automatic parameter calibration. Many use standard numerical or mathematical optimization techniques, e.g., Kalman filter [85] and principal differential analysis [86]. Other approaches like Bayesian methods, including Monte Carlo sampling [84], and heuristic and evolutionary practices like Simulated Annealing and Genetic Algorithms are widely used for parameter calibration problem.

Precisely, and due to the continuos advances in computing power and high performance computing tools and platforms, heuristic and evolutionary computing techniques and algorithms, especially genetic algorithms, have become practicable and more reliable to solve parameter problem in environmental models (e.g., [87, 88, 89, 90]).

The presented scheme which intends to enhance forecast skill in NWP models, implements genetic algorithms to calibrate its NWP closure parameters. In the next section, evolutionary algorithms for optimization problems and in particular, the basics of genetic algorithms are described in more detail.

3.2 Evolutionary Algorithms for Optimization Problems

In practice, there exist many approaches, that have been applied for calibrating input parameters in different types of modelling and simulation systems.

In particular, heuristic and evolutionary computing practices have demonstrated over the years to posses desirable properties for parameter calibration problem [84]. These have been considered as a successful practice to overcome irregularities contained widely in environmental models.

Evolutionary Algorithms, as a representative of this field, have become therefore, a standard strategy to solve complex search problems in environmental modelling. Actually, this is due to the fact that they do not make any assumptions about resulting fitness landscape

That is, in searching problems, Evolutionary Algorithms, which are inspired by the biological evolution, consider all the individuals of a population as candidate solutions for that underlying search problem. These algorithms apply a fitness function (also refereed to as cost function), by which, the encountered solutions are evaluated. Then, evolution is regarded as a repetitive process, applying genetic operators in a consecutive manner.

On the other hand, Genetic Algorithms (GA), as one of Evolutionary Algorithms s practices, have been evolving as the best investigated and most popular algorithms to solve parameter calibration problem in environmental models. Other approaches of Evolutionary Algorithms have been also applied for such calibration problems, however, these approaches (e.g., Genetic Programming and Neuroevolution) are known to be less generic, and more suitable to be applied in different problem domains.

In environmental modelling, however, GA are applicable and widely applied applied to optimize model parameters [88, 89, 90, 91, 92].

In the following section, a description of the basic information of Genetic Algorithms is provided. Including the most important genetic operators and their functionalities.

3.2.1 Genetic Algorithm Basics

As outlined in Section 3.2, Genetic Algorithms are considered as the most popular approaches applied to solve parameter estimation problem in environmental models. As it has been presented in [93], GA are an effective tool for parameter optimization in environmental modelling due to their remarkable properties: they are able to rapidly locate optimal solutions, even for large searching spaces, and especially suitable for application in problem domains that have a complex fitness landscape.

GA are generally considered as a global population-based search heuristic, which mimic the process of natural evolution. That is, they repetitively apply operators which include elitism, selection, fitness calculation, crossover, mutation and reinsertion over the involved solutions, referred to as individuals. They usually evolve individuals in a targeted population of individuals for a large number of generations, until a stopping criterion is met. The fitter individuals of a population survive and transmit their properties to offspring, thus replacing poor solutions and increasing the average fitness.

The operational sequence of the application of GA for parameter calibration in environmental modes is depicted in Figure 3.1.

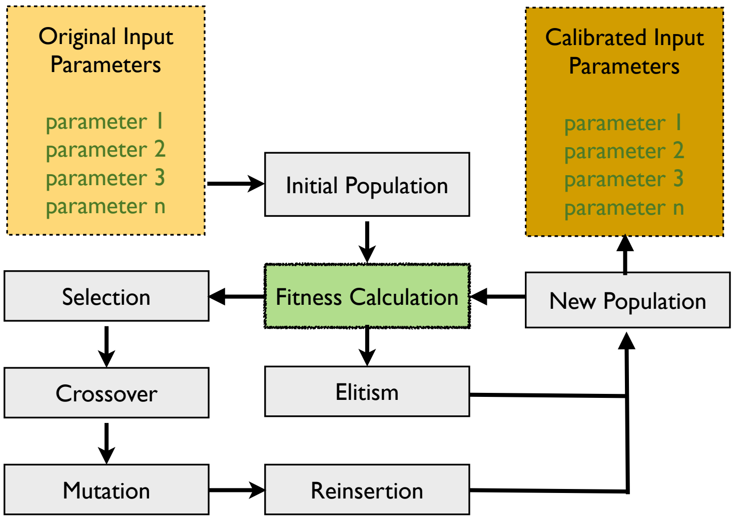


Figure 3.1: Operational sequence of the implementation of GA to calibrate input parameters.

That is, a GA normally operates over a population of individuals, with a certain size. Each individual (commonly denoted by chromosome) in the targeted population represents a partial parameter solution. Whereas these individuals are made up of a set of parameter values, as such, an individual is a combination of a set of attributes having certain values, referred to as gens (parameter 1, parameter 2, to parameter n as shown in figure 3.1).

The gens of each chromosome are encoded as real values within a previously predefined range. Usually, individuals are initialized randomly in the first population, hence, genes are assigned with random vales that fall within the predefined lower and upper pounds.

GA operators, as seen in the figure, are applied to produce offspring and create new individuals, maintaining genetic variety. Typically, the three main operators of GA are selection, crossover, and mutation, which operate over chromosomes in order to reproduce fitter chromosomes for subsequent generations. Following, the main operators of the GA are described in more detail.

Selection

The selection operator of the GA is responsible for determining which individuals will be selected for reproduction, as well as the number of offspring to be produced. Typically, individuals of a population are selected applying many techniques, which include tournament, roulette wheel, or local selections.

The mostly used is the roulette wheel selection, where the selection process depends on the fitness of each individual. It operates over these individuals defined by their fitness, when the wheel goes through these individuals, the probability of an individual to be selected is relative to its fitness, as such, the fitter an individual, the more probable to be selected. It should be highlighted that in the presented work, the roulette wheel technique is used for the selection operator in the applied GA.

Crossover

After selection, crossover is applied to generate offspring by exchanging segments (gens) from the selected chromosomes (generally referred to as parent chromosomes). It is supposed that the resulting individuals will inherit the favorable parts of their parents. Nevertheless, usually some selected chromosomes are reserved for the next generation without crossing it over, this process is generally known as the elitism, which intends to retain n ($n > 1$) most promising selected individuals to be reproduced for the next generation without any further change in their gens.

On the other hand, there are many crossover types; one, two, or multipoint crossover which could be applicable. A certain type could be chosen according to the characteristics of the individuals in each domain problem. These types define the number of segments (gens) are to be exchanged during the cross operation.

Crossover operators are normally configured to a certain probability, referred to as crossover probability . This probability determines how often crossover is performed over a selected set of individuals. The implementations however, recommend a crossover probability of 0.6 to 0.7, by which, GA normally do have better performance.

In the presented work however, one and two-point crossover are applied in the used GA, as well, the probability is chosen to be within the pre-mentioned ranges.

Mutation

In order to increase the population diversity, mutation is applied in GA. As in nature, mutation occurs very infrequently, and can often result in a weaker individual. Occasionally, however, a better individual could be obtained.

Actually, besides crossover, mutation process in GA is regarded the most important operation, because the performance of GA is highly influenced by these two operators.

As it has been stated before, it is intended to increase the diversity of the population by the implementation of mutation, that is done by perturbing some gens in some individuals after realizing the crossover process. Typically, this operator is configured to have very low probability. Actually, applying a mutation process can prevent that the algorithm becomes stuck in local optima.

Following, the fitness evaluation, which is a fundamental part of GA is described.

Fitness Evaluation

The quality of each individual (each possible solution), is evaluated by the so-called fitness function. Actually, this function is a problem dependent function, by which, the goodness of possible solutions are evaluated by their influence in reducing the irregularities and enhancing the performance in a certain model system. This work uses a statistical fitness function, which is widely used in evaluating the quality of weather predictions.

On the other hand, more than one fitness function are applicable in GA. That is, a GA application can use more than fitness function, where the resulted algorithm is widely known as multi-objective genetic algorithm.

In summery, It is widely agreed that GA are successful optimization methods for calibrating input parameters in environmental models. These algorithms operate on a population of possible solutions iteratively, that is, they start by selecting randomly a set of possible solutions/individuals, then, they evaluate these individuals according to a predefined fitness function, operators like crossover and mutation are consequently applied to reproduce an enhanced set of individuals to a next generation of solutions, this is repeated many times until a certain predefined criterion is met.

Basing on the ideas of the outlined topics; both parameter estimation approaches, especially Genetic Algorithms, together with the concept of the Ensemble Prediction System method for weather prediction enhancement, a new methodology was developed aiming at improving forecast skill in weather prediction models by optimizing their input parameters. Following, the proposed methodology is discussed in the next section.

3.3 Genetic Ensemble

In this section, the proposed prediction scheme is presented and described in detail. The principles of the proposed scheme depends on the basics of EPS (cp. Sections

As it has been stated before, the introduced methodology uses evolutionary computing methods, particularly, Genetic Algorithms in order to find the most timely optimal values of model closure parameters that appear in physical parametrization schemes which are coupled with NWP models. In order do that, a new phase is aggregated to the classical prediction scheme, which we call the calibration phase, in which, GA methods are used to find out optimal values of NWP model closure parameters.

Firstly, the traditional way of weather prediction process is summarized in Figure 3.2. As it can be seen in the figure, the traditional prediction scheme consists of a NWP model and a set of input data. The model is provided by input weather initial and boundary conditions, and also by certain values of model closure parameters (those which normally belong to coupled physical schemes).

Considering that the actual time (now) is time t_i (as depicted in the figure), in order to predict weather variables for the next hours or days, i.e., at time t_{i+n}, a numerical weather prediction model gets both initial and boundary conditions as input, as well as model closure parameters given for time t_i. These values are entered into the model, which then is integrated over computing facilities, to return the predicted values of weather variables at time t_{i+n}.

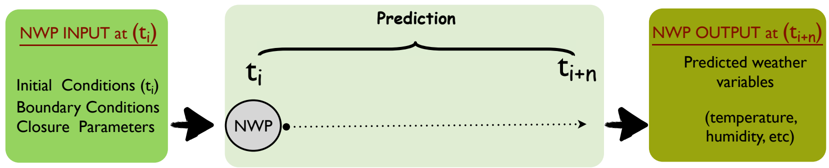


Figure 3.2: Classical weather prediction scheme by NWP models.

One execution of this classical prediction scheme is considered as a deterministic prediction, i.e., one single set of initial weather conditions and one combination of model closure parameters are processed by the NWP model to predict the future values of the weather. In the case of EPS, various sets of initial conditions (a distribution of perturbed initial conditions) are processed by the model, each of which in an independent forecast. The same happens for an EPS which counts for model closure parameters.

Due to the uncertainties in both initial conditions and model closure parameters ('physics' parameters) in NWP models, it is quietly often observed, that the forecasted values of weather variables tend to differ from the real observations of the same variables to a greater or lesser extent. As the prediction error accumulates gradually as the prediction time advances, deviations between real weather behaviour and forecasted weather variables become ever more significant. One reason for this incidence is that the processing of the classical prediction scheme is based upon one single set of input parameters or initial conditions, or in EPS, a limited set of input parameters and initial conditions, which for sure, do not represent a reasonable distribution of possible initial states of the weather.

In the next subsections, the Genetic Ensemble prediction scheme (G-Ensemble) for weather forecasting is introduced, which intends to calibrate NWP closure parameters in order to overcome the limitations imposed by EPS. Precisely, different versions of G-Ensemble scheme will be discussed, as well, a description of the implemented evaluation approach to compare errors of forecasts is provided. Finally, a parallel version of G-Ensemble scheme is presented.

3.3.1 G-Ensemble Prediction Scheme

To improve closure parameter quality and enable an automatic estimation and calibration of model input parameters for weather predictions, a new prediction scheme composed of two phases is developed. By doing so, an intermediate calibration phase in the style of feedback control systems is aggregated to the classical prediction scheme, as such, the quality of the last prediction outcome is evaluated comparing it to the really observed values of weather variables, before every new prediction step. Consequently, the set of input parameters is gradually refined. The approach is designed in a way that any set of NWP model closure parameters can be steered.

It should be mentioned however, that some similar two-phase prediction schemes were developed in order to calibrate input parameters in other areas of environmental modelling and different simulation systems. In a previous work in this university, Abdalhaq in [94, 95] proposed a similar scheme to calibrate input parameters for wildfire models. Other works dealing with similar parameter problems could be found, e.g., [96, 97]. The two-phase prediction scheme, as demonstrated in the previously mentioned works, has a strong potential to significantly enhance the quality of input parameters and hence to improve the prediction result.

The proposed scheme, which is called G-Ensemble consists of two phases: **calibration** phase and **prediction** phase as depicted in Figure 3.3.

Considering that t_i is the instant time from which the weather variables are going to be predicted, i.e., prediction is done within the period (t_i,t_{i+n}), calibration phase starts at a time prior to prediction time and ends at time 00:00 (t_i) of prediction period, i.e., calibration is done within the period (t_0, t_i).

During calibration phase, the genetic algorithm (parameter estimation method) is applied to search a set of input parameters values that would have reduced prediction error of a weather variable compared to real observations of the same variable at time t_i. The subsequent prediction phase then uses the encountered set of parameters to predict the weather variable the next time t_{+n}.

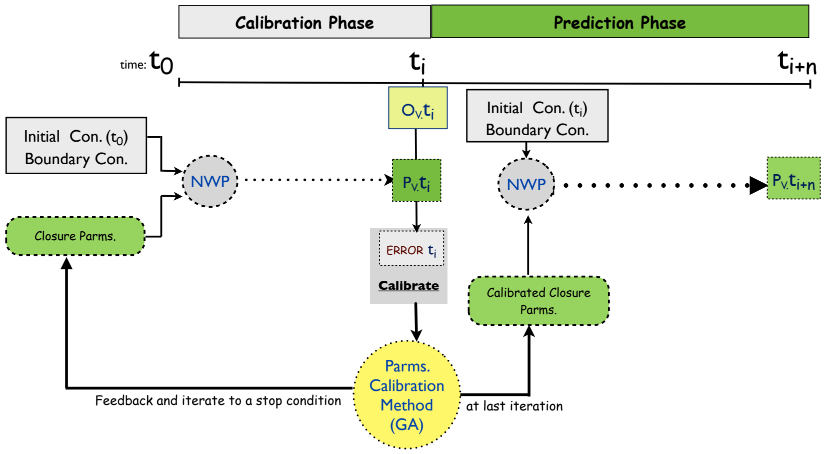


Figure 3.3: G-Ensemble: two-phase prediction scheme. NWP is a numerical weather prediction model, t_i is time 00:00 of prediction process, t_0 is a time instant prior to prediction phase, t_{i+n} is the future time to be predicted. O_V is an observed weather variable at time t_i, and P_V is the predicted value (the value to be predicted at time t_{i+n}) of the same weather variable.

Basing on EPS principles, the process of closure parameter estimation in calibration phase proceeds as follows:

1. at the beginning of calibration phase (time t_0 in fig. (3.3): a sample of the targeted parameter values from ensemble proposal distribution is generated (perturbations in closure parameter values);

2. the generated parameter values are inserted to the ensemble prediction model;

3. an ensemble of forecasts (the prediction model is different for each ensemble member regarding the targeted parameter values), is conducted to predict weather variables at time t_i, where real observations are available;

4. evaluation of a fitness function for each ensemble member is done at time t_i;

5. genetic algorithm functions (selection, crossover and mutation) are used to generate a new ensemble distribution from the set of combinations of closure parameters which score better predicting at time t_i; and

6. the process is repeated iteratively until an acceptable error value, or a predefined number of iterations is achieved.

At the last iteration in the calibration phase, the values of closure parameters, which produced the least error of prediction by the end of the phase, i.e., the ensemble member with the best forecast skill score at time t_i as shown in Figure 3.3, is selected to be used in prediction phase. More description about the prediction phase will be described later in this chapter (cp. Subsection 3.3.3).

That is, the objective of the additional calibration phase is to solve an inverse problem: Find a parameter configuration such that, given this configuration as input, the model output matches real variable observation. Having detected the model input that best describes the current environmental conditions, the same values, it is argued, could also be used to describe best the immediate future assuming model stability during the following prediction interval.

3.3.2 Error Evaluation- Fitness Function

A relevant point to be considered in the calibration phase is the error definition being one of the core elements of this phase. In this work, two different error functions are proposed, one referred to as Single-Variable and the other referred to as Multi-Variable. Depending on the error function used, we have designed two G-Ensemble strategies: Single-Variable G-Ensemble and Multi-Variable G-Ensemble, which are described below.

Single-Variable G-Ensemble

The calibration phase is conducted with the goal of enhancing predictions for a single weather variable. The error function for the evaluation of ensemble members in the used GA is the Root Mean Square Deviation $RMSD$ or Error $RMSE$, shown below in Equation 3.1. This error function is a frequently-used measure for the evaluation of weather predictions [98], which measures the differences between values predicted by a model or an estimator and the values actually observed from the variable being estimated.

In $RMSE$ equation, x_{obs} is an observed value of a variable x and x_{pre} is the predicted one for the same variable.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_{obs,i} - x_{pre,i})^2}{n}} \tag{3.1}$$

Using $RMSE$ error in the calibration phase limits G-Ensemble scheme to be oriented to enhance predictions for one weather variable at a time.

For example, we can use it to improve predictions of Temperature or Precipitation, but not for both at the same time. This occurs because the error used produces a value of the variable unit that cannot be compared with other variables. In order to overcome such a drawback, we proposed an alternative error function, which we refer as Multi-Variable G-Ensemble.

Multi-Variable G-Ensemble

The calibration is done with the goal of enhancing the prediction of multiple weather variables at the same time. To bypass the limitation imposed by $RMSE$ error, the Normalized $RMSE$ is used, see equation (3.2).

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_{obs,i} - x_{pre,i})^2}{n}}}{x_{obs(max)} - x_{obs(min)}} \tag{3.2}$$

The Normalized $RMSE$ (referred to as $NRMSE$) is the value of $RMSE$ divided by the range of the observed values of a certain variable. $NRMSE$ indicates the error percentage of the predicted value of a certain variable, compared to the range of its observed values. In order to consider more than one variable at a time, we evaluate $NRMSE$ for all variables, and then, we consider the addition of all of them as the Multi-Variable error function. For example, the $NRMSE$ of a model that predicts Temperature (T) and Precipitation (P) is the percentage obtained by the summation of two Percentages: $NRMSE(T)$ and $NRMSE(P)$, as shown in equation (3.3).

$$Error = NRMSE(var1) + NRMSE(var2) = value\% \tag{3.3}$$

Therefore, the calibration phase, and particularly the GA, considers this error function as the objective function used to sort the intermediate individuals of the ensembles.

3.3.3 Prediction Phase

As it has been described in the previous sections, the output of the calibration phase of G-Ensemble scheme is a set of combinations of NWP closure parameters, those which were being refined and thus, scored better forecast skill during the GA iterations.

That is, once the calibration phase is finished, it is the turn of the prediction phase. At this point, there are two alternatives to conduct a weather prediction process, described as follows:

G-Ensemble Set

Using this strategy, all the combinations of NWP closure parameters which are produced in the last iteration of the calibration phase, will form together an ensemble forecast in prediction phase. That is, each combination is considered to be an ensemble member and will be executed independently. As in EPS, the final prediction is represented by the average of all the results achieved by all ensemble members.

It is supposed that the prediction result of the forecasts of these calibrated combinations of NWP parameters, will be more accurate than the prediction result produced by the classical EPS, using non-calibrated combinations of NWP parameters.

Best Genetic Ensemble Member (BeGEM)

This strategy adapts a deterministic forecast, i.e., one single forecast is to be conducted in prediction phase. Precisely, at the last iteration in the calibration phase, the values of closure parameters, which produced the least value of $RMSE$ or $NRMSE$, i.e. the ensemble member with the best forecast skill score at the end of calibration phase, is selected to be used in prediction phase. This ensemble member is called: Best Genetic Ensemble Member (BeGEM). Our hypothesis is that, for short-range weather forecasts, if the forecast skill is improved in the calibrations phase by a set of a calibrated closure parameters, then, the same closure parameter values will also improve forecast skill during prediction phase.

However, the pre-described scheme (G-Ensemble), evaluates ensemble members in the calibration phase using a fitness function by considering one observation point for a certain weather variable. In other words, the evaluation of each ensemble member is done once at the end of calibration phase, i.e., at time t_i as shown in Figure 3.3. In the next Subsection, an enhanced version of G-Ensemble is presented, by which, a window of observations can be used to evaluate ensemble members during calibration phase rather than one single observation point.

3.3.4 G-Ensemble - Calibration Window

It is supposed that evaluating ensemble members during calibration phase according to one single observation for each weather variable is not that fair. Basically, due to the stochastic nature of weather, some ensemble members may change their performance over time. Hence, to help the used GA to take better decisions when selecting the set of ensemble members that will reproduce a consecutive generation of ensemble members in each iteration, G-Ensemble scheme is extended such that, it becomes capable to evaluate ensemble members according to a window of observations rather than one-point observation.

Back to Figure 3.3, ensemble members are evaluated according to real observations available at t_i. In contrast, in the extended version of G-Ensemble as shown in Figure 3.4, ensemble members are evaluated according to observations available in more than one point during Calibration Phase.

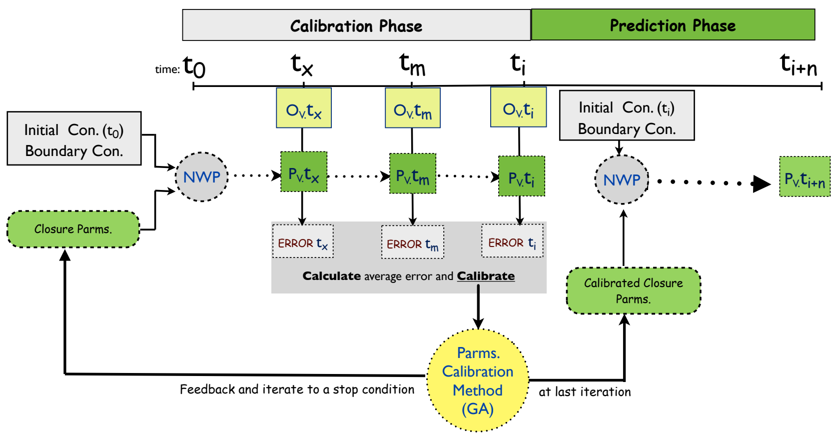


Figure 3.4: Extended G-Ensemble scheme: NWP is the a numerical weather prediction model, t_i is time 00:00 of prediction process, t_0 is a time instant prior to prediction phase, t_{i+n} is the future time to be predicted, t_x and t_m are time instants within the calibration phase where real observations are available as in t_i, O_V is an observed weather variable and P_V is the predicted value of the same variable.

If prediction is to take place from time t_i to t_i+n, calibration phase is to be conducted in the interval (t_0,t_i), however, observations could be available at times t_x, t_m (any model time steps that fall within calibration phase), as well as at time t_i. Being these observations available, the GA fitness function (cp. Subsection 3.3.2) considers the average error of the three error values calculated at times t_x, t_m and t_i, for each ensemble member according to the three observations available at the same time instants.

The main goal behind this extended version of the G-Ensemble, is to better guide the used GA in the selection process, as such, by considering more than one single observation point, ensemble members could be evaluated more fairly during calibration phase.

In the next Subsection, another extension of the G-Ensemble is presented, by which, an intelligent calibration process is introduced, which directs the calibration phase towards optimizing a certain set of model closure parameters relative to each weather domain characteristics.

3.3.5 Multi Level Genetic Ensemble (M-Level G-Ensemble)

The ultimate goal of the proposed G-Ensemble is to improve forecast skill in NWP models. This is done by calibrating the values of their input closure parameters, to which NWP models are known to be highly sensitive. These parameters belong to the different physical parameterization schemes (which are coupled with NWP models), thus, calibrating their values needs a further knowledge of their meaning and their relativity to the targeted domain where the prediction process is taking place.

Actually, this is because there are various parameterization schemes whose closure parameter values depend on the characteristics of the targeted domain. However, an effective method of calibration should be suitable for all domains on one hand, and on the other hand, it should be automatic, as such, capable to adapt itself to the particular characteristics of each targeted domain.

To discuss the proposed G-Ensemble which will be directed to solve this problem, we focus our study on one example of these parameterization schemes; the land surface models (LSM), which will serve as a prove of concept of our method. It should be mentioned however, that the proposed scheme can be used in all other parameterization schemes to solve the same problem, such as the case of planetary boundary layer parameterization schemes (cp. Subsection 2.1.3).

Being that said, in the case of the closure parameters used by the coupled LSM, normally prediction domains exhibit heterogeneity in their surface characteristics (cp. Subsection 2.1.3). That is, the terrain of a certain domain (the first mesh of the 3-D grid) could include different vegetation types, and different soil textures. As it has been shown in Figures 2.14 and 2.16, for each *landuse* (vegetation) category, there are different values of vegetation parameters, as well, for each *soil type* category, there are different values of soil parameters. So, we are facing a problem summarized as follows:

1. each domain normally include various categories of *landuse* and *soil type*.

2. each category of *landuse*, has different values of closure parameters.

As well, each *soil type* category has different values of closure parameters.

The question to be answered in this case, which set of parameters should be targeted by the calibration process?

In order to solve this problem, before starting the process to enhance prediction, the parameters to be calibrated should be selected. In other words, it is necessary to determine exactly what class of parameters to calibrate. For example, it does not make sense to optimize parameters related to Grassland category in a region that has 100% of it s terrain as water.

As it has been previously described, NWP model starts a process of prediction over a certain zone using the initial and boundary conditions defined by their location (longitude, latitude and vertical distance) for each grid point of the domain. NWP models are also provided by terrain maps, which are available at high resolutions globally. These maps define the surface and topographical characteristics of the targeted domain, as such, the first mesh grid points of the domain are assigned with a number indicating its *landuse* category (LU-index) and with another number indicating its *soil type* (SLTYP).

During prediction process, the NWP model needs surface parameter values for each surface grid point in order to calculate the evolution of the other weather variables. These parameter values depend on their categories, and for each category, the NWP model is provided by its default parameter values which are provided in stand-alone tables like those shown in Table 3.1 (complete tables used by WRF models are provided in figures 2.14 and 2.16).

Then, for each surface grid point, the NWP model reads its assigned *landuse* category LU-index and, goes to LAND USE table to obtain the values of the surface physical parameters corresponding to that category. The process is done for all surface grid points and the same is done with soil texture parameters.

As shown in Table 3.1, there are 33 *landuse* categories, each of which has 7 surface closure parameters, and there are also 19 *soil type* categories, each of which has 10 closure parameters. Prediction enhancement must look for optimal values of these input parameters. Therefore, the first step consists of selecting which category or categories correspond to the region terrain where weather prediction process is going to take place.

Ⓐ

LU-index*(33)*	NAME	Land Use Parameters (7)			
		ALB	SLMO	SFEM	....
1	Urban land	15.	.10	.88	....
2	Agriculture	17.	.30	985	....
3	Range-grassland	18.	.50	985	....
4	Deciduous forest	18.	.25	.96	....
.	...	...	....	....	....
.	...	...	....	....	....
33	Industrial	10.	.10	.97	....

Ⓑ

SLTYP*(19)*	NAME	Soil Parameters (10)			
		BB	DRYSMC	F11	....
1	Sand	2.79	0.010	-4.472	....
2	Loamy Sand	4.26	0.028	-1.044	....
3	Sandy Loam	4.74	0.047	-0.569	....
4	Silt Loam	5.33	0.084	0.162	....
.	...	...	....	....	....
.	...	...	....	....	....
19	White Sand	2.79	0.010	-4.472	....

Table 3.1: Ⓐ: A snapshot of LANDUSE.TBL (*landuse* parameters: 7 parameters for 33 *landuse* categories). Ⓑ: A snapshot of SOILPARM.TBL (soil parameters: 10 parameters for 19 soil categories). Both tables are read by WRF model when NOAH LSM is coupled to calculate surface parameterizations. See [65] for more description of *landuse* and soil parameters.

In order to do that, we propose the Multi Level Genetic Ensemble (M-Level G-Ensemble) approach, which aims not only to select the targeted parameters to be calibrated, but also to enable the calibration phase to optimize more than one level of these parameters, i.e., to calibrate closure parameters which belong to more than one *landuse* and soil texture categories related to a certain domain.

For that, a new phase is aggregated to to the previously presented G-Ensemble in Subsection , this new phase is called parameter selection phase, which is described as following.

Parameter Selection Phase

The introduced scheme of the M-Level G-Ensemble approach is depicted in Figure 3.5, at parameter selection phase, a small program is developed to read the gridded data related to the domain surface grid points, then LU-index and SLTYP are extracted for each grid point and a counter is applied to each LU-index and SLTYP to find how many domain grid points are of each certain *landuse* category and soil texture type.

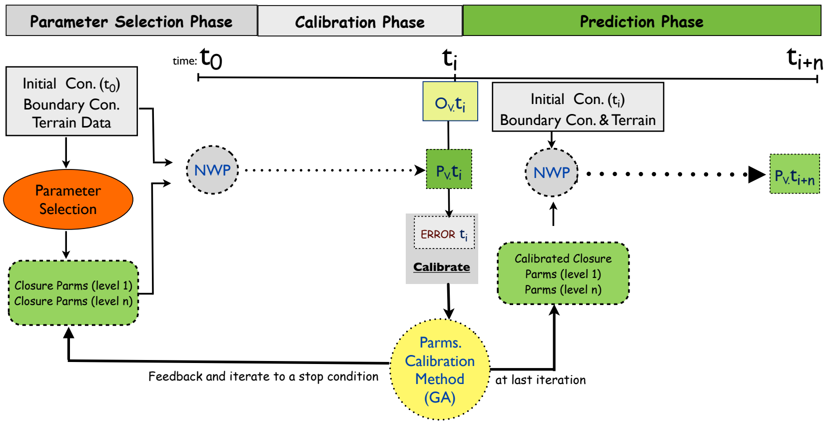


Figure 3.5: M-Level G-Ensemble: three-phase prediction scheme. NWP is a numerical weather prediction model, t_i is time 00:00 of prediction process, t_0 is a time instant prior to prediction phase, t_{i+n} is the future time to be predicted. O_V is an observed weather variable at time t_i, and P_V is the predicted value (the value to be predicted at time t_{i+n}) of the same weather variable.

As a result, a table is constructed including each *landuse* and *soil type* category, and each of which will have the number of domain grid points indexed to that category. Table registers are ordered in a descending order: the first *landuse* category mostly repeated within the domain grid points is referred to as the first dominant *landuse* category, the second is the second dominant, and so on. The same is done for the *soil type* categories. Table 3.2 shows an output of parameter selection phase.

Then, by the end of this phase, the categories of *landuse* and soil parameters for a certain domain are classified and ordered according to their weight (how often they are repeated in the targeted domain grid points). The first dominant category parameters of both *landuse* and soil parameters is referred to as first level parameters.

It is supposed that finding optimal values of the first level parameters will have more effect in reducing prediction error than the parameters of the second, third, to the end of the rest of categories.

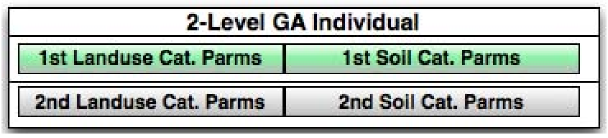
For example, suppose that a weather prediction process is to be conducted in a certain domain. The domain closure parameters of its terrain are classified by the parameter selection phase according to its dominant *landuse* and Soil categories as shown in Table 3.2.

Level	Land-Category	Coverage	Soil-Category	Coverage
1	15 -Mixed Forest	37%	6 -LOAM	41%
2	27 -White Sand	12%	15 -Bedrock	23%
3	16 -Water Bodies	9%	14 -Water	9%
4	7 -Grassland	5%	10 -Sandy Clay	8%
5	Rest	37%	Rest	19%

Table 3.2: Parameter Category Selection: Level (1) register contains the *landuse* category (15) which covers 37% of the domain and the *soil type* category (6) which covers 41% of the same domain. Register (5) represents the rest of *landuse* and *soil type* categories existing in the domain with their respective percentage of coverage.

A calibration process considering the first level parameters, i.e. 7 parameters of the the *landuse* category (Mixed Forest) and 10 parameters of *soil type* category (LOAM) will consider these parameters as a single individual in a population of individuals, each of which has different values of these 17 parameters. In this case, the GA deals with 1-Level parameters and its individual is shown in Figure 3.6.(a).

(a)



(b)

Figure 3.6: 1-Level GA individual in (a), and M-Level GA individual in (b).

The process of calibration at this point seeks to find the optimal values of these 17 parameters. In evolutionary computing terms, the GA used to deal with this case is usually known as single chromosome GA.

Actually, this is what two-phase G-Ensemble does (cp. Subsections 3.3.5), where the calibration phase is oriented to optimize closure parameters related to only one level, i.e., the closure parameters related to the first dominant *landuse* and *soil type* categories.

However, to consider the parameters of the other levels of categories in the domain, a GA individual is constructed as shown in Figure 3.6.(b), by which, a 2-Level parameters constitutes a GA individual and the process of calibration seeks to find the optimal values of the combination of 34 parameters (17+17) divided in two levels, and so on for more levels.

That is, by applying this applying this approach, model closure parameters are selected automatically, and furthermore, more than one level of these parameters could be calibrated, by which, the heterogeneity in land surfaces of NWP domains (see [55, 56]) is considered.

3.4 Parallel G-Ensemble

As it has been described earlier (cp. Section 2.2), the problem of the uncertainty in NWP initial conditions produces what is called imperfectness in prediction accuracy. The previously mentioned methods, among others are implemented to reduce the margin of the imperfectness in prediction accuracy. However, the trade-off between cost (execution time) and prediction accuracy is a crucial factor that should be considered to select the most suitable enhancement method.

Our proposed scheme intends to improve prediction quality in NWP models. However, a question still remains to be answered regarding the amount of time that must be spent to get better predictions. And how much time should be allowed under reasonable circumstances in practice?

Fortunately, most of NWP models are parallel programs, for example, in [99] a study of scalability of WRF NWP model over HPC platforms is provided. Other enhancement methods, such as EPS and 3D-Var, have parallel versions that may benefit from HPC platforms.

However, EPS may exhibit significant limitations when executed in environments with relatively small number of computational resources. A hypothetical situation where a prediction is needed for the evolution of meteorological variables for the next 20 hours might illustrate this limitation. If we assume that the time of the parallel execution of that prediction is 1 hour over a set of 10 available computers, then an EPS with 20 or more ensemble members will take more than 20 hours (as each ensemble member is a stand alone prediction), and the overall result will be useless in practice.

In order to provide an effective enhancement method, it should be first of all feasible, as such, an enhancement in prediction quality could be obtained within a reasonable waiting time.

Being that said, and by analyzing the proposed scheme, an overhead in execution time could be produced by the Calibration phase. That is because there is an added work to be done before prediction phase. However, it should be highlighted that, by using the proposed G-Ensemble scheme, a deterministic forecast could be used in the prediction phase, i.e., one single prediction process with a calibrated set of closure parameters. In contrast, an EPS system is conducted by a set of independent forecasts.

Additionally, G-Ensemble prediction scheme is paralleled by implementing a Mater/Worker parallel paradigm, as shown in Figure (3.7). That is, as it has been described earlier, the calibration phase of the G-Ensemble approach consists of GA iterative operations over a population of individuals, these individuals are executed independently for each iteration, which represent very short forecasts having different parameter value combinations.

As it can be seen in the figure, this process is paralleled, as such, the individuals of each iteration (short forecasts) are distributed over the available computing resources for execution and evaluation of their corresponding error. Then, results of all individuals are gathered back to the Master node, where GA operations are executed. This is repeated as many iterations as needed during the Calibration phase.

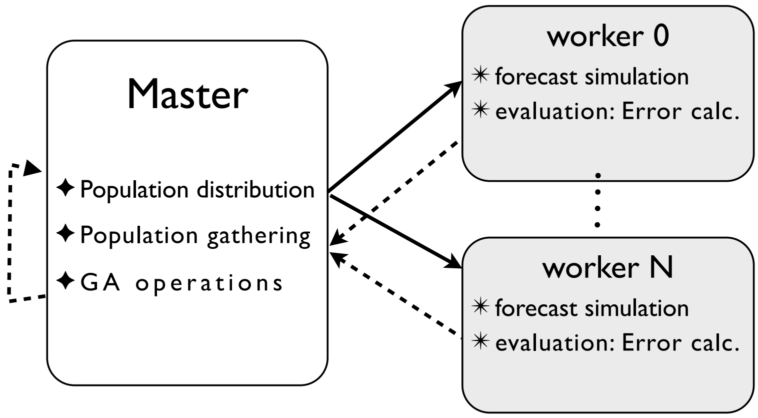


Figure 3.7: Master/Worker paradigm for Calibration phase.

It should be mentioned as well, that each individual in the calibration phase is a short prediction simulation, which is by itself a parallel application.

The ultimate goal of parallelizing the proposed scheme, is to get benefit of the parallel computing platforms, in order to reduce the overall waiting time needed for an improved prediction results.

3.5 Summery

Through this chapter, a new prediction scheme was presented, which aims at improving forecast skill in short and medium range weather forecasts, over the existing enhancement methods. The proposed scheme, which is called G-Ensemble, uses evolutionary computing techniques, particularly, genetic algorithms to calibrate NWP model closure parameters. These closure parameters are model input parameters which belong to physical parameterization schemes that are coupled with modern NWP models.

Actually, these parameters are tuned manually in most operational weather prediction centers, however, owing to the sensitivity of NWP models to the values of these parameters, it is of paramount importance to develop an effective and automatic approach capable to find optimal values of these parameters.

The presented scheme is developed in different versions, each is designed to deal with a certain particularity of the NWP process:

1. **G-Ensemble (1-point observation)**: which is a two-phase prediction scheme, where a combination of closure parameters are optimized within an aggregated calibration phase. The calibration process is realized using one single observation, available at the end of calibration interval. As such, all the possible combinations of closure parameters are used in short forecasts, each of which is evaluated using the available observation, then, GA operators are applied to reproduce a next generation of parameter combinations. This process is repeated iteratively until satisfying a predefined condition.

2. **G-Ensemble (window observations)**: an extended version of the G-Ensemble, where a set of observations are used in the evaluation process during calibration phase, instead of one single point observation. This approach is designed in order to fairly guide the used GA in deciding which are the best combinations to be selected for the next iteration.

3. **M-Level G-Ensemble**: this approach is developed in a three-phase scheme, in order to consider more than one combination of closure

parameters in the calibration process. This is accomplished by adding the parameter selection phase, which makes this approach capable of selecting automatically more than one level of model parameters to be calibrated, the selection process is done considering the particularity of the domain, that is, the selected parameters for one domain, would be totally different of the targeted parameters for another domain.

Besides these introduced approaches, G-Ensemble is designed to improve forecast skill for a set of weather variables together, which could be considered a significant contribution, that is because generally, enhancement methods are applied to improve forecast skill for a certain weather variable. However, by implementing a normalized evaluation fitness function in the calibration phase, more than one weather variable could be targeted by the proposed scheme, this approach is called the **Multi-Variable G-Ensemble**.

Additionally, the whole scheme is paralleled using a Master/Worker programming paradigm, this is done in order to enable the proposed scheme to get benefit of nowadays parallel computing environments, with the ultimate goal of reducing its execution time.

Finally, the output of the calibration phase of G-Ensemble scheme is a set of combinations of NWP closure parameters, those which were being refined and thus, scored better forecast skill during the GA iterations. Having this set of the calibrated combinations, two alternatives are feasible for conducting the prediction phase, the first is by running all these calibrated combinations in an EPS, which we call the **G-Ensemble Set**, and the second is to run a single deterministic forecast which is the ensemble member with the best forecast skill score at the end of calibration phase. This alternative is called the Best Genetic Ensemble Member or **BeGEM**. In the next chapter, the proposed prediction scheme with all of its approaches/versions are tested and evaluated over a real weather prediction case.

Chapter 4

Experimental Evaluation

During the presentation of this work, weather prediction concept has been clarified, as well as the problem of predictability in numerical weather prediction models. Furthermore, a study of ensemble prediction system (EPS), as one of the most successful method for enhancing weather prediction predictability in the last two decades, was presented. On the other hand, powerful parameter calibration approaches which deal with the problem of uncertainty in simulation input parameters were discussed, especially, evolutionary computing techniques and Genetic Algorithms (GA).

Additionally, G-Ensemble scheme, the proposed methodology to improve forecast skill in short and medium range weather predictions has been presented and discussed precisely. Thus, through the subsequent sections of this chapter, an evaluation of the overall performance of G-Ensemble in predictions of a real wether case will be presented and discussed in detail.

The objectives of the following experimentations are to show the enhancement in weather predictability gained by using the presented Genetic Ensemble scheme (G-Ensemble). As well, to demonstrate that the gained accuracy in the predictability is not time consuming and does not implicit added computational overhead.

Additionally, to evaluate the benefits achieved by applying G-Ensemble scheme of weather prediction enhancement in scenarios of limited computing resources, where other enhancement methods, such as EPS, could not be feasible in scenarios of limited computing resources.

In order to achieve the pre-mentioned objectives, G-Ensemble scheme by its various versions (cp. Section 3.3) is used to predict weather variables o the following real weather case: **Hurricane Katrina**, which occurred in 2005 in the Gulf of Mexico, and known as the strongest, deadliest and most destructive storm to impact the coast of the United States during the last 100 years. Hurricane Katrina historical data are well known to the scientific community, and widely used for experimentation and research [100].

In the provided experimentations, the results of G-Ensemble predictions are compared to the traditional ensemble prediction system (EPS), which is referred to as classical EPS , as such, these classical EPS forecasts are initiated with the same initial ensembles used by G-Ensemble, i.e., with a non-calibrated set of NWP closure parameters.

Additionally, the parallel version of G-Ensemble is evaluated over a parallel computing architecture, on the other hand, the scalability of the presented scheme over a HPC environment is presented and discussed. Finally, many configurations of the proposed G-Ensemble scheme are tested, these configurations are related to the settings of the used Genetic Algorithm in its calibration phase.

In the next section, a description of the configuration of G-Ensemble is provided, along with the computing platform specifications.

4.1 System Configuration

The proposed G-Ensemble scheme implements Genetic Algorithms in its calibration phase. During the conduction of these experiments, specific settings of the Genetic Algorithm were fixed, if not, it will be indicated otherwise in the individual case. These settings are listed following:

- An *elitism* rate of 5% for all GA executions, maintaining by this rate the 50% of the individuals (most promising ones) to be regenerated in the consecutive iteration of the applied GA.

- Roulette wheel *selection* and two-point *crossover* were applied. The *crossover* probability was set to 0.7 in all executions, however, both *crossover* probability and types are tested in different cases and will be indicated by occurrence. The probability rates of the applied *crossover* fall within the recommended rates of this operator where GA generally tends to have a maximized performance.

- Due to the relatively small size of the initial populations of individuals which are adapted in our experiments, *mutation* probability was set to 0.2. By applying this probability value, it is intended to overcome the limitation imposed by the relatively small size of GA initial populations.

- During GA executions, initial populations are initialized randomly, respecting the given upper and lower bounds of each parameter (gen). The same populations are used to initiate both the G-Ensemble and the classical EPS.

The targeted weather variables to be enhanced by the proposed scheme are selected for both importance and their relation with the calibrated model parameters. These six variables are:

- Accumulated Precipitation (mm),
- 2-meter Temperature (K),
- Sea Surface Temperature (K),
- 10-meter Wind Velocity components U10 and V10 (m/s), and
- Latent Heat Flux LHF ($w/m2$).

On the other hand, the used NWP model in all experiments to predict these variables is the Weather Research and Forecasting model (WRF).

Additionally, all the demonstrated results represent the average of a set of executions, more than three execution in the majority of the experiments. This is done to assure that the obtained results are reliable by avoiding the randomity which could be produced in GA operations in some cases.

Finally, all experiments were executed on a cluster of 32 computing nodes (128 CORES) (Intel(R) Xeon(R) CPU 5150 @2.66GHz 4MB L2, 8 GB Fully Buffered DIMM 667 MHz.

4.2 Hurricane Katrina Test Case

Hurricane Katrina - pictures shown in Figure 4.1 - occurred on August 28, 2005 in the Gulf of Mexico. It is considered as one of the strongest storms to impact the coast of the United States during the last 100 years. Unfortunately, Katrina caused the death of more than 1,800 persons along with a total property damage that was estimated at $81 billion (2005 USD) [100].

(a)

(b)

Figure 4.1: Hurricane Katrina: **(a)** Satellite picture of Katrina on Aug. 29, 2005 at 12:15 P.M, and **(b)** Flood caused by Katrina in New Orleans (USA).

The objective of the experiments is to predict the evolution of weather variables from time: 12:00 h. of the day 28/08/2005 to time 00:00 h. of 30/8/2005 (a period of 36 hours in which the major effects of the hurricane were produced). The evolution of these weather variables is produced every 3 hours and the spatial resolution of Katrina domain was 12km.

To get predict weather variables at 12:00 h. of 28/08/2005, WRF is provided by initial conditions of the atmospheric state in the zone three hours before, i.e., NWP model started prediction from time 09:00 of 28/08/2005 in advance. For the proposed G-Ensemble scheme, the calibration phase is conducted from time 00:00 of 28/08/2005 to time 09:00 of the same day.

Again, the targeted weather values for prediction, as described before in Section 4.1, are: Latent Heat Flux, Sea Surface Temperature, 2-meter Temperature, 10-meter Wind Velocity components U10 and V10, and the Accumulated Precipitation.

In the following subsections, the experimentations for Katrina test case are described and analyzed.

Classical EPS predictability

Firstly, Figure 4.2 shows some experimental results for a classical EPS prediction of 40 ensemble members, to predict (every 3 hours) the evolution of: a) *Accumulated Precipitation*, b) *Latent Heat Flux*, c) *Sea Surface Temperature*, and d) *10 m. Wind Velocity Component*.

The evolution of the values of these weather variables using EPS was under-estimated in this case. Concretely, in many cases, EPS gives a prediction error of more than 30% compared to observed values in a certain hour.

For example, a prediction of the accumulated precipitation variable, as shown in Figure 4.2.**(a)**, at hour 39. The observed value was (35 mm), however, the predicted value using EPS was (24 mm), so, EPS in this case produced about 32% of prediction error compared to the real observed value. Consequently, it can be easily concluded, that there is a significant margin of enhancement in prediction which could be achieved.

In order to enhance prediction for weather variables during the hurricane, G-Ensemble was applied and compared to classical EPS prediction results. Both G-Ensemble and classical EPS were ititialized by the same set of NWP closure parameters.

Referring to the two fitness functions in G-Ensemble (cp. Section 3.3.2), the proposed scheme could be oriented to enhance predictions for one weather variable applying Single-Variable G-Ensemble, or for a set of weather variables together, applying the Multi-Variable G-Ensemble. In the next subsection, Single-Variable G-Ensemble results are compared with predictions of classical EPS for the same variables.

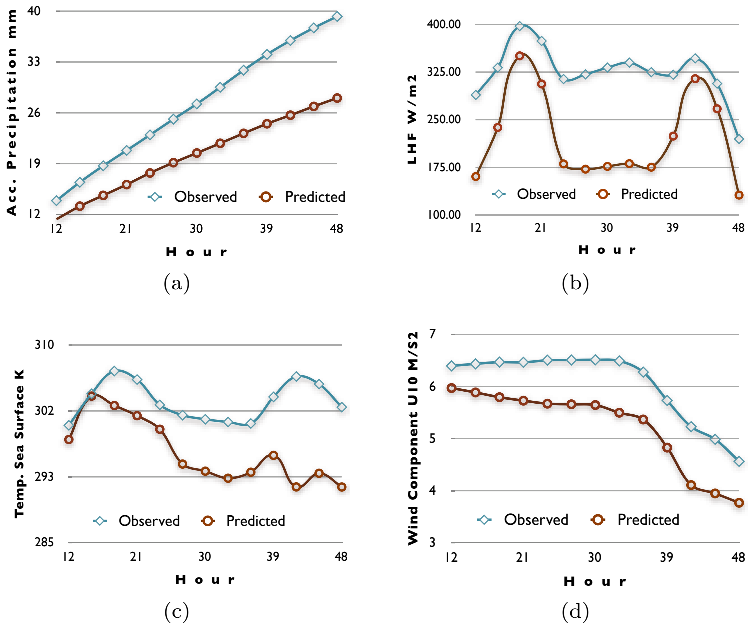


Figure 4.2: Classical EPS prediction results compared to observed values: **(a)** EPS vs. Observed Precip, **(b)** EPS vs. Observed Latent Heat Flux, **(c)** EPS vs. Observed Sea Surface Temp, and **(d)** EPS vs. 10 m. Wind Velocity Component (U10).

4.2.1 G-Ensemble: Single-Variable

Single-Variable G-Ensemble is applied on two different cases: to predict Accumulated Precipitation weather variable as demonstrated in Figure 4.3.**(a)**, and to predict V10 Wind Velocity Component as shown in Figure4.3.**(b)**.

BeGEM and G-Ensemble set (cp. Section 3.3.3) are compared with a classical EPS conducted by the same initial ensemble members of G-Ensemble scheme. The number of ensemble members in the experiment was 40, and the Genetic Algorithm of the calibration phase was configured to iterate 20 times.

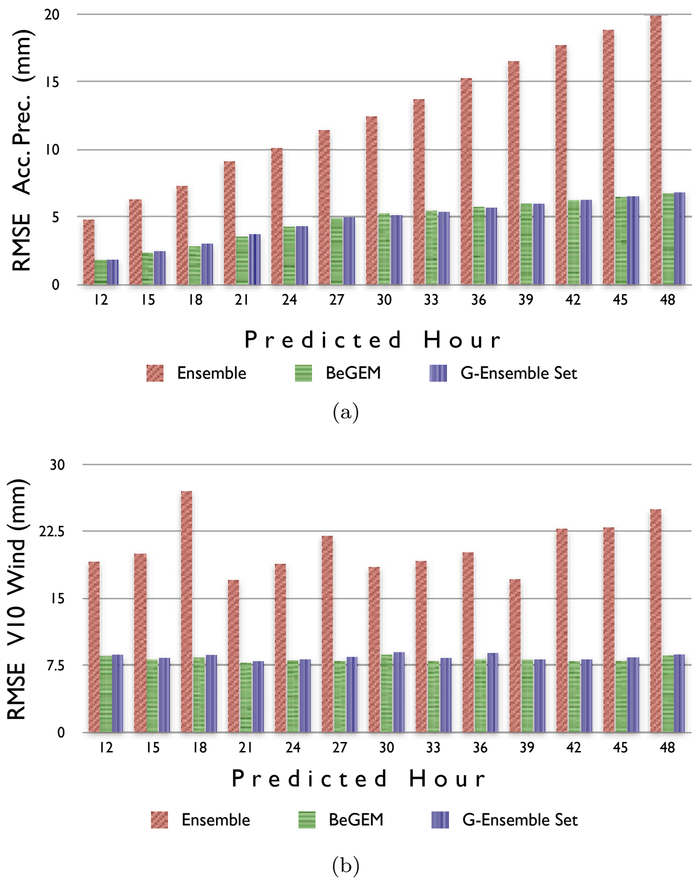


Figure 4.3: Classical EPS prediction results compared with Single-Variable G-Ensemble for the prediction of **(a)** Accumulated Precipitation, and **(b)** Wind Velocity Component (V10).

In both cases, with the same initial ensemble members, we obtained a significant improvement in prediction quality for the targeted weather variables. It can be observed, that both the deterministic forecast (BeGEM) and the ensemble forecast (G-Ensemble Set) produced by G-Ensemble scheme, have reduced significantly prediction error for both variables compared with the same results obtained by the classical EPS.

That is, a calibrated set of NWP closure parameters by the proposed scheme, increases forecast still of the final prediction over a classical EPS which is initialized by the same initial ensemble members before calibration. Moreover, the BeGEM, as a single deterministic forecast, produced even more enhanced prediction results.

In the following subsection however, Multi-Variable G-Ensemble and G-Ensemble (window) are also applied to enhance prediction results for a set of weather variables.

It should be highlighted that, through the rest of this chapter, the comparison will be made between the BeGEM (a calibrated single deterministic forecast) with the classical EPS approach.

4.2.2 G-Ensemble: Multi-Variable and Calibration Window

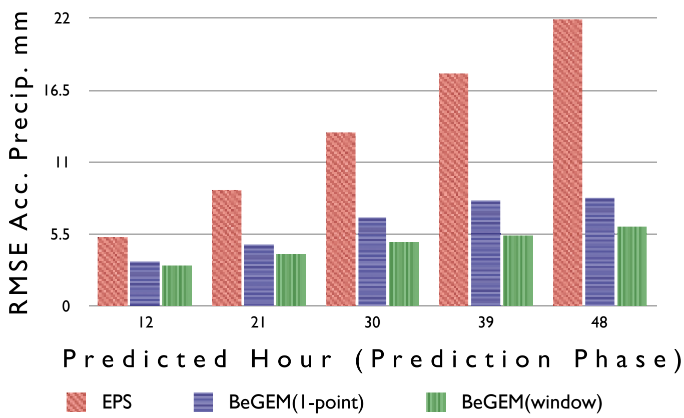
As it has been described in Section 3.3.4, G-Ensemble scheme has been extended such that, it becomes capable to evaluate ensemble members according to a window of observations rather than one-point observation. That is, to help the used GA to take better decisions when selecting the set of ensemble members that will reproduce a consecutive generation of ensemble members in each iteration of the calibration phase.

In the subsequent experiments, prediction errors ($RMSE$ and $NRMSE$) produced during prediction phase of three ways of prediction are compared;

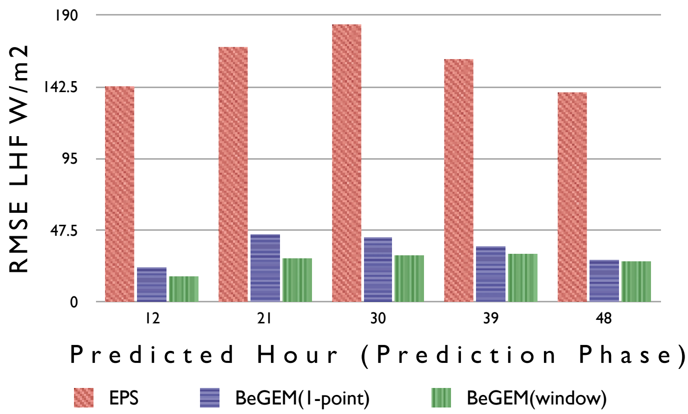
1. *G-Ensemble* approach, where calibration phase considers `one-point observation, at time 09:00 of 28/08/2005 ($BeGEM(1 - point)$)
2. *G-Ensemble* extended approach, where calibration phase considers a window of observations, at time 7:00, 8:00 and 09:00 of 28/08/2005 ($BeGEM(window)$).
3. EPS (referred to the classical EPS), which is used to refer to the average error of an ensemble forecast conducted by the initial ensemble members used in the first iteration of calibration phase (an ensemble forecast such that the prediction model is different for each ensemble member regarding the targeted parameter values, these parameters are not calibrated).

Firstly, we show experimental results for two different cases: to predict Accumulated Precipitation (results shown in Figure 4.14.**(a)**, and to predict Latent Heat Flux (results shown in Figure 4.14.**(b)**).

The Genetic Algorithm of the calibration phase was configured to iterate 15 times over an initial population size of 40 individuals (initial ensemble size).



(a)



(b)

Figure 4.4: Single-Variable G-Ensemble: **(a)** $RMSE$ error in prediction of variable Acc. Precipitation and, **(b)** variable LHF. Results are of classical EPS, BeGEM(1-point) and BeGEM(window) for both variables.

In both cases, with the same initial ensemble members used in the classical EPS case, a significant improvement in prediction quality is obtained by G-Ensemble approach over the classical EPS. Additionally, it could be also observed that better enhancements in predictions were obtained by the extended G-Ensemble which considers a window of observations in its calibration phase.

Furthermore, Figure 4.5 illustrates the results of the same experiment, but applying the Multi-Variable G-Ensemble strategy, in order to enhance prediction results for a set of three variables together: *Latent Heat Flux LHF*, *2-meter Temperature*, and the *Accumulated Precipitation.*

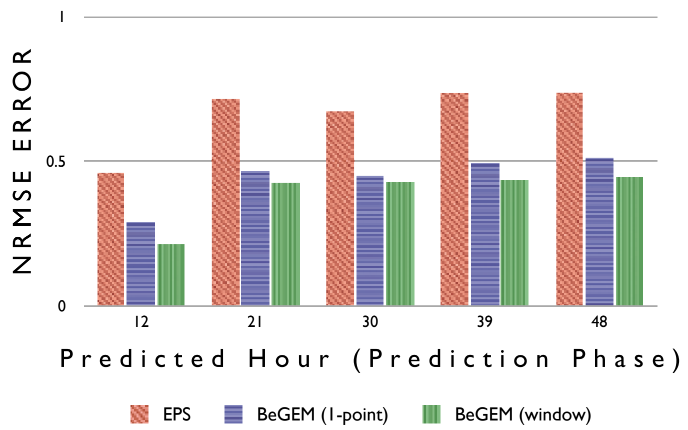


Figure 4.5: Multi-Variable G-Ensemble; $NRMSE$ in prediction of variables: *Latent Heat Flux LHF*, *2-meter Temperature*, and the *Accumulated Precipitation.*

Again, significant reduction of the $NRMSE$ was obtained in the prediction of a set of weather variables together and, the extended version of G-Ensemble, which considers a window of observations, produces a better forecast skill.

Additionally, it is observed that the reduction in the $NRMSE$ of the three variables together by using the Multi-Variable G-Ensemble approach, also provides an enhancement in the prediction of each weather variable alone.

In other words, all variables were better predicted when G-Ensemble oriented to reduce the $NRMSE$ of those variables together. To illustrate these results, Figure 4.12 shows how the corresponding prediction error of each variable was reduced when the extended G-Ensemble was oriented to reduce the $NRMSE$ of the three variables together.

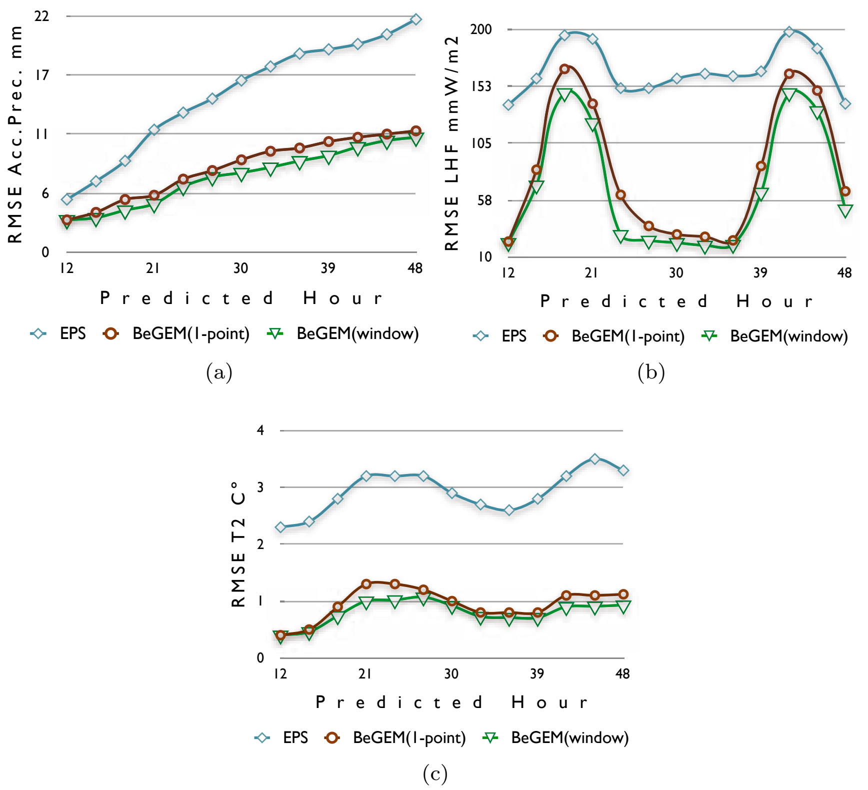


Figure 4.6: $RMSE$ prediction error of: **(a)** Accumulated Precipitation, **(b)** Latent Heat Flux LHF, and **(c)** 2-meter Temperature. Prediction using *BeGEM (1-point)* and *BeGEM (window)* produced after 15 iterations of the calibration phase of the Multi-Variable G-Ensemble.

The results obtained in these experiments approve our hypophysis that, G-Ensemble leads to better estimation of closure parameter values when it considers a window of observations rather than one single point observation for the evaluation of its ensemble members (GA individuals) during calibration phase. Actually, we believe that the reason behind this is that, the used GA in the calibration phase is better guided by more fairly error value when this error reflects an interval of time rather than one single point, to evaluate the performance of each ensemble member which determines its probability to be selected for subsequent iterations of the GA.

4.2.3 Multi-Level G-Ensemble

M-Level G-Ensemble, as it has been described before (cp. Subsection 3.3.5), is presented as a three-phase prediction scheme, which intends to enhance weather predictions by calibrating more than one level of parameters, using the Multi-Objective Genetic Algorithm. Concretely, the targeted parameters to be calibrated, belong to the coupled Land Surface physical schemes. These are relative to the category of each grid point of the surface mesh of NWP domain.

In the case of Hurricane Katrina, major part of its surface is water, however, some other parts of its surface are not. Figures 4.7 and 4.8, show the dominant *land use* and *soil type* categories of Katrina domain.

The objective of the subsequent experiments, is to demonstrate that by calibrating more than one level of these surface parameters, i.e., parameters which belong to the second, third, etc. dominant categories, prediction results are even better enhanced and forecast skill can be improved as more levels of parameters are calibrated.

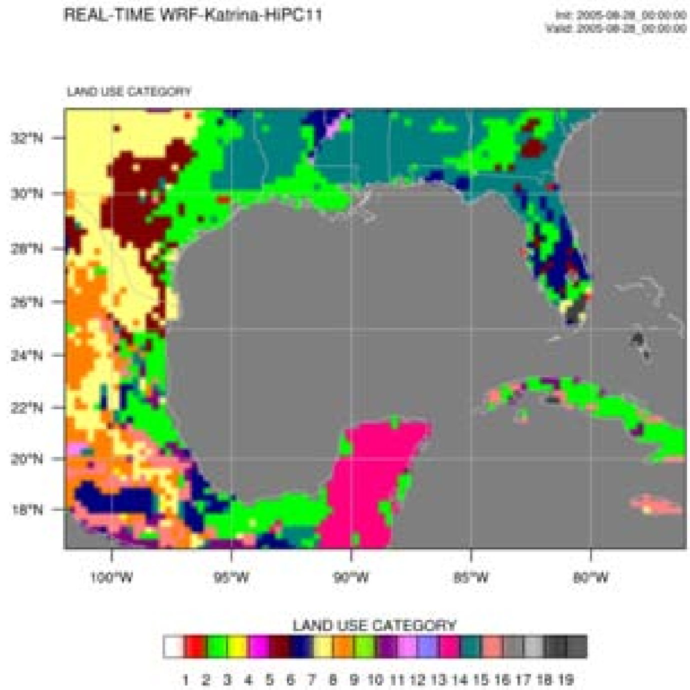


Figure 4.7: Hurricane Katrina *land use* dominant categories.

In order to calibrate more than one level of these parameters, parameter selection phase was executed over the domain, and its *land use* and *soil type* parameters dominant categories were classified upon there occurrence as shown in Figure 4.9.

Then, by having the categories classified by their occurrence in the domain, M-Level G-Ensemble was applied firstly by the Single-Variable strategy to predict: Acc. Precipitation (results shown in Figure 4.10.**(a)**), and to predict Latent Heat Flux (results shown in Figure 4.10.**(b)**).

In both cases, with the same initial ensemble members used in the classical EPS case, a significant improvement in prediction quality is achieved. Moreover, It can be observed as well, that more improvement in predictions were obtained as more levels of parameters were calibrated.

The Genetic Algorithm of the calibration phase was configured to iterate 20 times over an initial population size of 40 individuals (initial ensemble size).

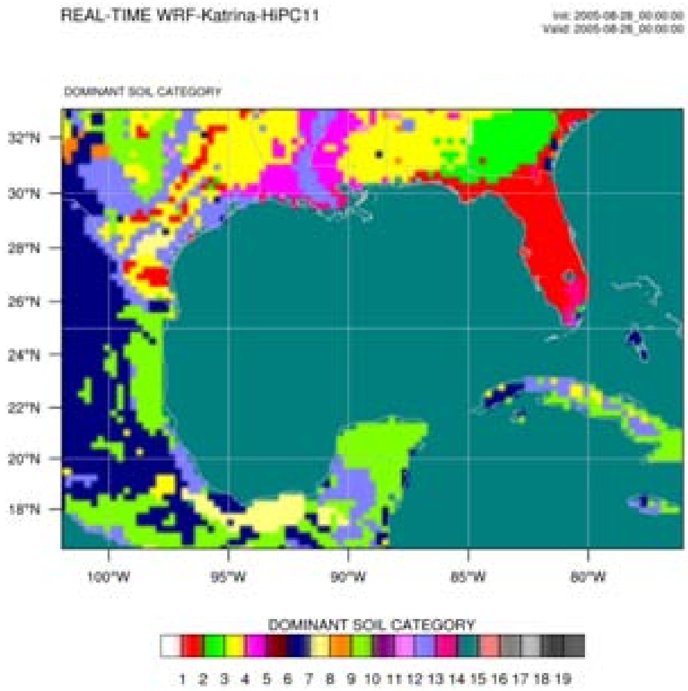


Figure 4.8: Hurricane Katrina *Soil* dominant categories.

The M-Level G-Ensemble approach is also tested to enhance predictions of a set of weather variables at the same time, by applying the Multi-Variable G-Ensemble strategy, using the the normalized error $NRMSE$ in calibration phase as the fitness function of the GA. In this case, as it can be seen in Figure 4.11, significant improvement in the prediction of a set of weather variables together, was obtained.

Additionally, it can be observed that a reduction in the $NRMSE$ of a set of variables also provides an enhancement in the prediction of each weather variable alone. In other words, all variables were better predicted when M-Level G-Ensemble oriented to reduce the $NRMSE$ of those variables together. To illustrate these results, we show in Figure 4.12 how the corresponding prediction error of each variable was reduced when M-Level G-Ensemble was oriented to reduce the $NRMSE$ of six variables together.

It can be concluded form the obtained results, that M-Level G-Ensemble achieves notable enhancement in prediction results as more levels of parameters are calibrated.

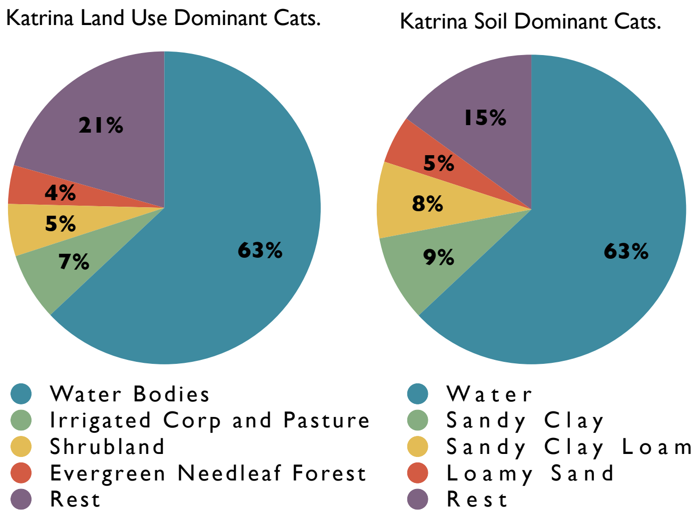
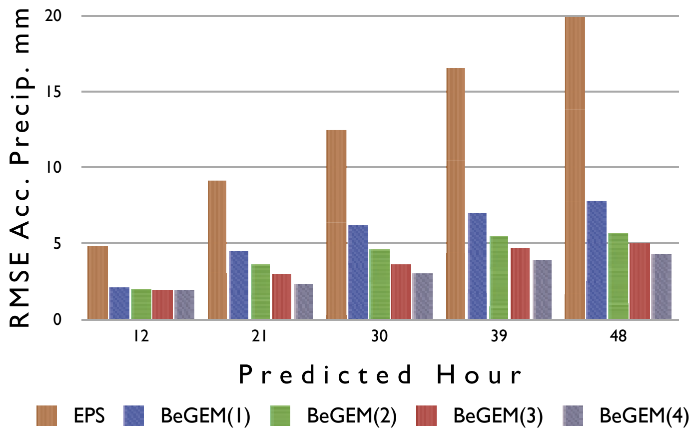
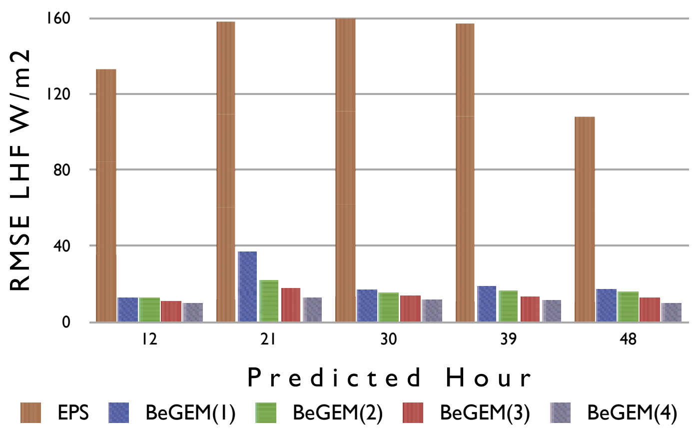


Figure 4.9: Katrina *Landuse* and *Soil*.

In the next Subsection, the parallel version of the proposed G-Ensemble scheme is evaluated.



(a)



(b)

Figure 4.10: Single-Variable M-Level G-Ensemble: **(a)** $RMSE$ error in prediction of Acc. Precipitation variable, and **(b)** Latent Heat Flux (LHF) variable. Results are of classical EPS and the BeGEM(x) for both variables, where x refers to the number of calibrated parameter levels.

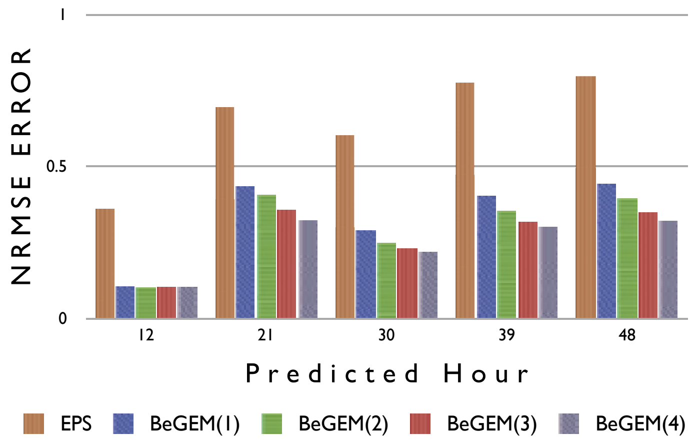
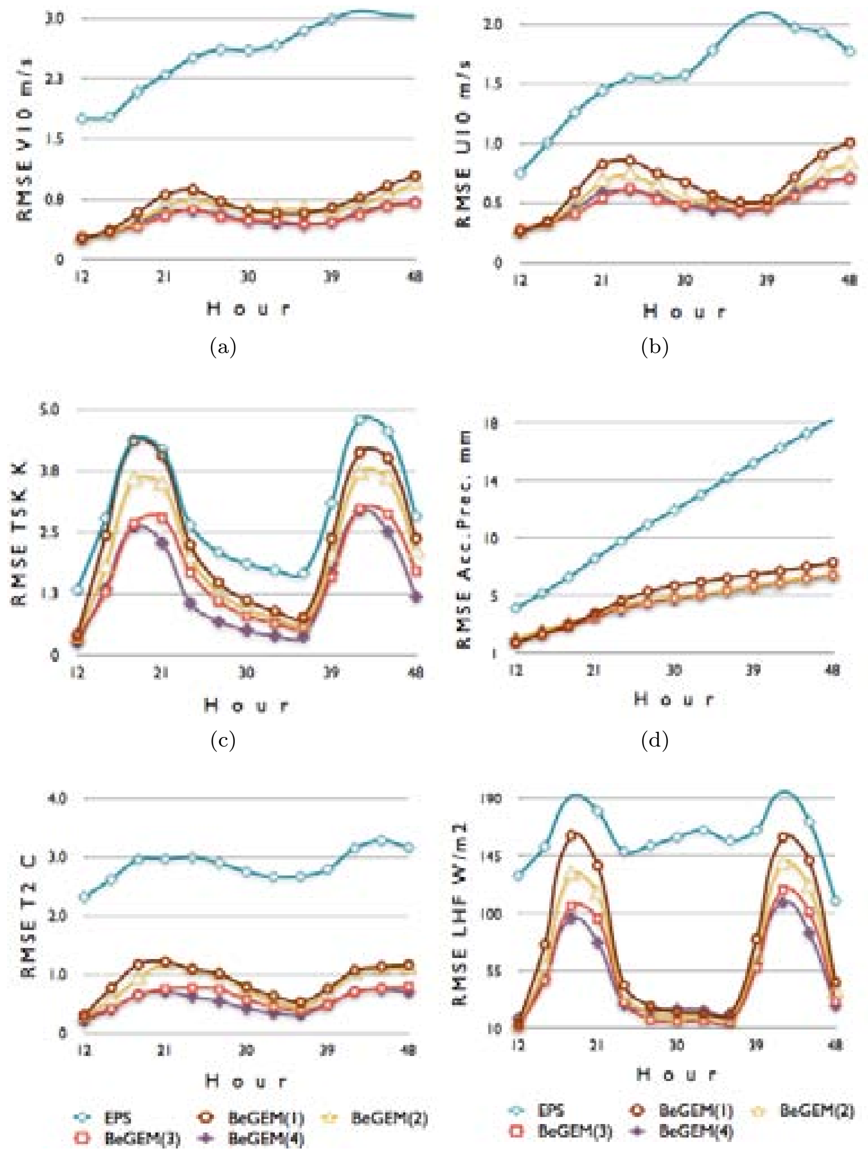


Figure 4.11: Multi-Variable M-Level G-Ensemble; $NRMSE$ in prediction of variables: Latent Heat Flux (LHF), Sea Surface Temperature, 2-meter Temperature, 10-meter Wind Velocity components U10 and V10, and the Accumulated Precipitation. Results are of classical EPS and the BeGEM(x) for both variables, where x refers to the number of calibrated parameter levels.



(a) (b) (c) (d) (e) (f)

Figure 4.12: *RMSE* prediction error of: (a) 10-meter Wind Velocity component V10, (b) 10-meter Wind Velocity component U10 (m/s), (c) Surface Skin Temperature TSK, (d) Accumulated Precipitation RAINC, (e) 2-meter Temperature, and (f) Latent Heat Flux LHF. Prediction using *BeGEM (1, 2, 3, and 4)* produced after 15 iterations of the Calibration Phase of the *Multi-Variable M-Level G-Ensemble*.

4.2.4 Parallel G-Ensemble and Cost

To evaluate the parallel M-Level G-Ensemble scheme according to execution time and prediction enhancement, we show in Table 4.1 various scenarios of parallel M-Level G-Ensemble predictions with their respective execution times compared to a classical EPS prediction conducted with 40 initial ensemble members.

Multi-Variable M-Level G-Ensemble was used in the calibration phase to enhance prediction of 6 weather variables together, applying BeGEM(x), where x refers to the number of the calibrated levels of closure parameters, this process is tested over 5 different scenarios, which correspond to different GA settings (number of iterations in calibration phase and the initial ensemble size).

Predictions were executed on a cluster of 30 computing nodes (Intel(R) Xeon(R) CPU 5150 @2.66GHz 4MB L2, 8 GB Fully Buffered DIMM 667 MHz). Figure 4.13 shows the respective prediction error of each scenario of those in Table 4.1.

Number	Scenario	Init. Size	# of Iterations	Ex.Time
1	*EPS*	40	-	468 m.
2	*BeGEM(4)*	40	5	109 m.
3	*BeGEM(4)*	40	10	168 m.
4	*BeGEM(4)*	40	15	223 m.
5	*BeGEM(4)*	40	20	279 m.
6	*BeGEM(4)*	20	20	189 m.

Table 4.1: Execution time Vs Scenario

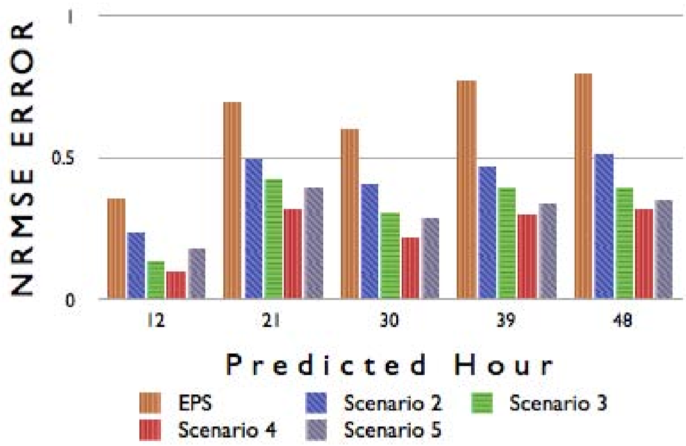
In all scenarios of M-Level G-Ensemble, a significant reduction in execution time along with a corresponding reduction of prediction error are observed. Additionally, as shown in Figure 4.13.(b), the parallel version of M-Level G-Ensemble achieved better execution times than EPS when both were executed on computing platforms with a relatively small number of processors (less than 50). As more computing nodes are available, EPS performance improves. And it achieved similar or slightly better execution times than M-Level G-Ensemble when more than 70 machines were used.

Actually, this is due to the combination between the number of resources available and the type of executions. In scenarios with limited number of computing resources, the parallel M-Level G-Ensemble is better because it constitutes a set of short executions, repeated each iteration in the calibration phase. In other words, Calibration executes each generation of individuals, which represent short forecasts.

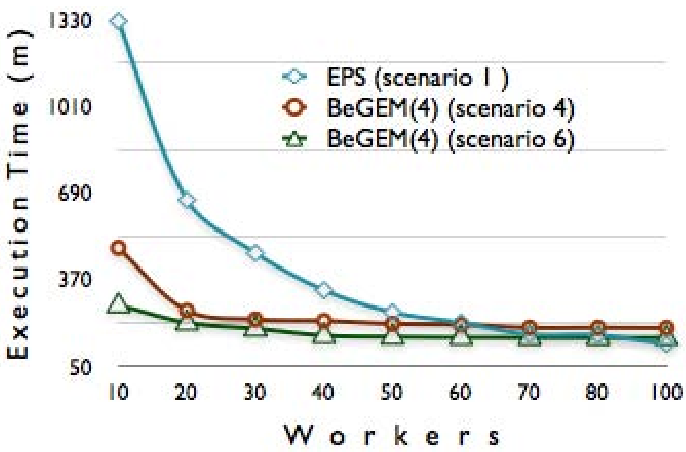
Therefore, the waiting time between each iteration is short. In contrast, in the case of EPS running on a limited number of resources, each ensemble member is a long forecast that will use the resources for a relatively long time. While the number of resources increases, this problem of waiting time in EPS is alleviated. This is why EPS shows the same performance or even better when the system size is increased.

In summary, parallel M-Level G-Ensemble method provides the possibility to select between various scenarios considering a balance between prediction quality and prediction cost, maintaining always a significant margin of enhancement in prediction quality. Moreover, in scenarios with limited number of computing resources, in which EPS could not be used due to its time constraints, parallel M-Level G-Ensemble stands to be a good alternative choice.

In the following Subsection, G-Ensemble is tested over the same prediction case, but with different configurations in the calibration phase regarding the operators of the used GA and the size of the used initial ensemble members.



(a)



(b)

Figure 4.13: (a): Multi-Variable BeGEM(4); $NRMSE$ of prediction of variables: *Latent Heat Flux* , *Surface Skin Temperature* , *2-meter Temperature*, *10-meter Wind Velocity components U10 and V10*, and the *Accumulated Precipitation. (b): Scalability of EPS of 40 ensemble members over 100 computing nodes, BeGEM(4):(initial ens:40, iterations:15)*

4.2.5 Tuning G-Ensemble

G-Ensemble scheme is tested with different GA configurations regarding its Crossover type and ratio, and by variating its initial population size in order to get better predictions.

The goal behind these tests is to provide a more completed insight of the scenarios and possibilities of how to configure an operational G-Ensemble according to the time allowed for prediction process, and to the number of computing resources available. In the subsequent experiments, prediction errors $RMSE$ produced during prediction phase for *Latent Heat Flux LHF* variable, applying two ways of prediction, are compared:

1. Single Variable G-Ensemble approach, with different initial ensemble sizes, Crossover type and ratio, and different number of iterations in calibration phase.

2. The EPS approach, which is used to refer to the average error of an ensemble forecast conducted by the initial ensemble members used in the first iteration of Calibration Phase (an ensemble forecast such that the prediction model is different for each ensemble member regarding the targeted parameter values, these variables are not calibrated).

In Figure 4.14.(a), prediction error is shown by using the G-Ensemble approach with different initial ensemble sizes to predict LHF variable compared to the classical EPS of the same ensemble sizes. The prediction error of the G-Ensemble approach is also depicted alone for the sake of clarity in Figure4.14.(b).

The Genetic Algorithm was configured to iterate 20 times over different initial ensemble sizes. Its three main operators were configured as follows: *Selection*: (elitism: best one of two), *Crossover*: (probability=0.7, type: two points Crossover), and *Mutation*: (probability= 0.2). As shown in Figure 4.14, in all cases with different initial ensemble sizes, G-Ensemble provides less error values in prediction compared to EPS predictions with the same initial ensemble members. A significant improvement in prediction quality is always gained.

Additionally, it can be observed that increasing the size of an EPS does not produce better results. Actually this happens because EPS results represent an average of the predictions of all ensemble members and, knowing that these members are variated regarding their closure parameters in a random way, using more members does not assure less average error. In contrast, increasing initial ensemble size, which will be calibrated iteratively by the G-Ensemble, provides better prediction results as observed in the same figure. That is, by increasing the initial ensemble size in G-Ensemble,

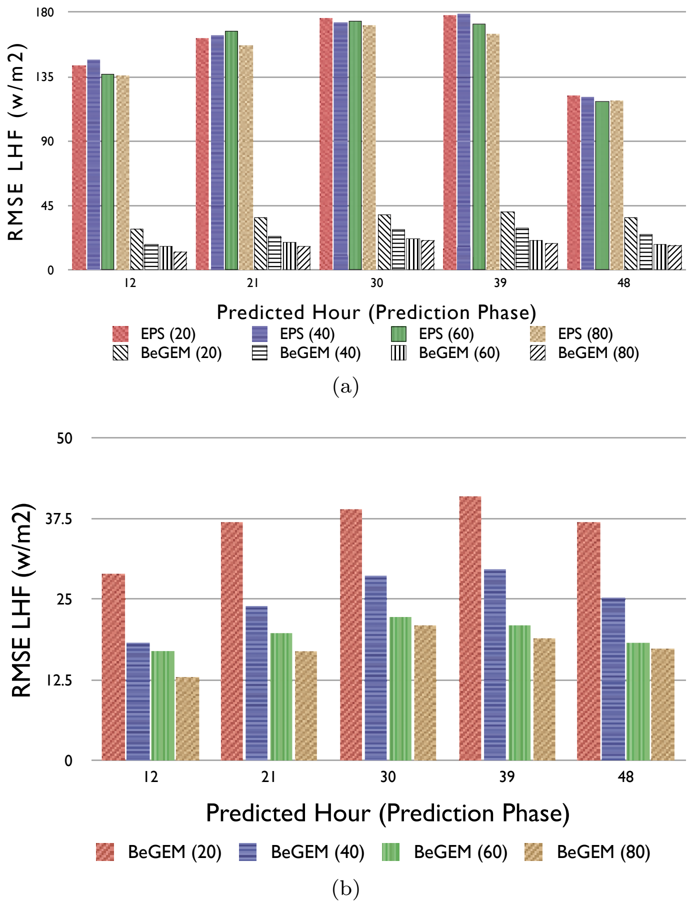


Figure 4.14: RMSE of LHF pediction. (a): Single-Variable G-Ensemble prediction error Vs. Classical EPS prediction error. Results are of classical EPS(x) and the BeGEM(x), where x refers to the initial ensemble size. (b): A snapshot of (a) to demonstrate RMSE of the different BeGEM(x).

the probability for finding better solutions through GA iterations, also increases.

On the other hand, Figure 4.15 shows the GA convergence in the calibration phase of G-Ensemble approach. As such, the error of the best ensemble member through GA iterations is depicted in the figure, using different initial ensemble sizes. As it could be observed, the BeGEM produced after 10 iterations when G-Ensemble was conducted using an initial ensemble size of 80 members, was equal or slightly better than the same BeGEM, produced after 20 iterations when G-Ensemble was conducted by 20 initial ensemble members.

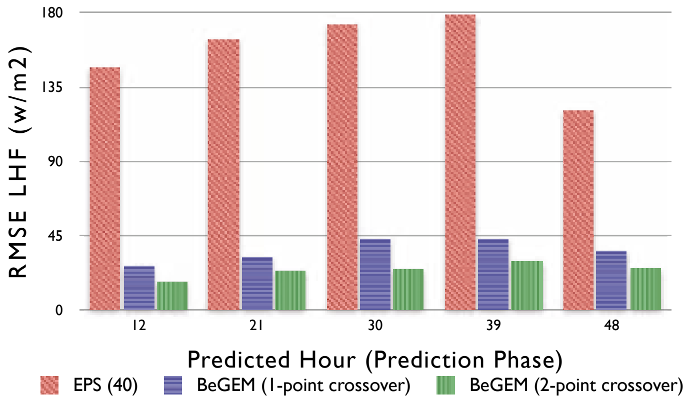
Figure 4.15: Calibration phase: BeGEM performance through the Calibration phase iterations for different initial ensemble sizes.

Then, according to the availability of computing resources, their number and the interval of availability, a certain scenario of the combinations between initial ensemble size and number of iterations, could be selected, referring to execution times provided previously (cp. Subsection 4.2.4).

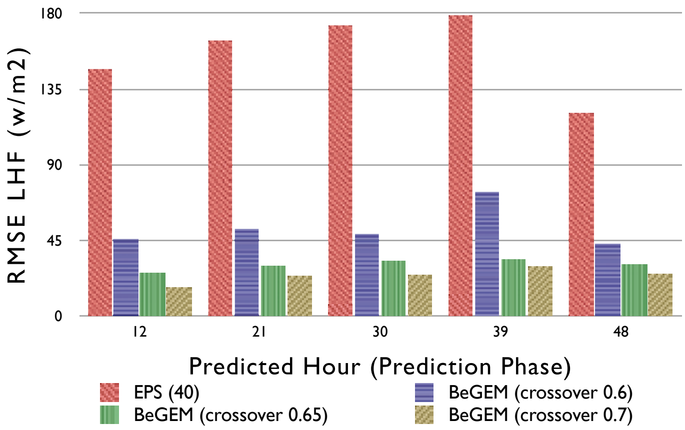
Moreover, G-Ensemble scheme is tested to predict the same weather variable (LHF) by changing the type of the GA Crossover during calibration phase, results are depicted in Figure 4.16.(a), and by changing the GA Crossover probability as shown in Figure 4.16.(b)

The obtained results show that when G-Ensemble used 2-points Crossover in its GA during Calibration phase, prediction results were slightly better, and the same happened when Crossover probability was higher.

That is, when configuring the GA implemented in the G-Ensemble scheme on a relatively small size of initial ensemble members, better prediction quality could be obtained by 2-points Crossover and higher Crossover probability. Actually, this is due to the size of the initial ensemble size (initial population size): by using 2-point Crossover and a higher probability of Crossover operations, more variations in ensemble members could be obtained during each iteration of the calibration phase. This enhances the ability of the GA to look for better solutions over small initial populations, which is normally the case of NWP executions, where ensemble sizes are normally up to 50 ensemble members.



(a)



(b)

Figure 4.16: BeGEM $RMSE$ in prediction of LHF produced in (a): using 1-point and 2-point GA Crossover in the Calibration phase, and (b): using 2-point GA Crossover but with different Crossover probability ratios.

4.2.6 Discussion of Results and Summery

In the previous experiments, all the approaches of the proposed G-Ensemble scheme were tested on the case of Hurricane Katrina, and consequently, their results we compared with predictions obtained by the classical EPS.

The results obtained in the provided experiments confirm our hypophysis that, better estimation of model closure parameter values enhances weather prediction quality. However, further detected remarks by observing the obtained results are summerized as follows:

The proposed G-Ensemble scheme, showed in all cases, a significant improvement in forecast skill for the targeted weather variables, over the classical EPS. As well, all the different approaches/versions of G-Ensemble were tested; G-Ensemble, G-Ensemble (window observations), Multi and Single-Variable G-Ensemble and the Multi-Level G-Ensemble. In all these approaches, weather variables were predicted better than the classical EPS used for the same reason.

By using the Multi-Level G-Ensemble, it has been observed that as more parameter levels are calibrated, as more improvements in the prediction outcomes are obtained. Obviously, this is due to the fact that, when calibrating more levels of parameters, more proportion of the targeted domain parameters are optimized, because each level of these parameters represents a set of domain grid points.

Moreover, results of G-Ensemble (window observations) were compared with the formal G-Ensemble (1-point observation). It was observed G-Ensemble leads to better estimation of closure parameter values when it considers a window of observations rather than one single point observation for the evaluation of its ensemble members (GA individuals) during calibration phase. Actually, it is believed that the reason behind this is that, the used GA in the calibration phase is better guided by more fairly error value when this error reflects an interval of time rather than one single point, to evaluate the performance of each ensemble member which determines its probability to be selected for subsequent iterations of the GA.

The EPS approach, which is used to refer to the average error of an ensemble forecast conducted by the initial ensemble members used in the first iteration of Calibration Phase (an ensemble forecast such that the prediction model is different for each ensemble member regarding the targeted parameter values, these variables are not calibrated).

Additionally, the whole scheme was paralleled using Master/Worker paradigm and was tested executing it over a HPC platforms. Again, the obtained results showed significant improvements in prediction quality and, less execution times over classical EPS, certainly, in scenarios of limited number of computing resources, where classical EPS shows to be infeasible due to the imposed time constraints. Additionally, results showed that G-Ensemble approach provides the possibility to select between various scenarios considering a balance between prediction quality and prediction cost, maintaining always a significant margin of enhancement in prediction quality. Moreover, in scenarios with limited number of computing resources, in which EPS could not be used due to its time constraints, parallel G-

Ensemble stands to be a good alternative choice.

The last experiments showed prediction results obtained by tuning the G-Ensemble, as such, by changing some configurations regarding the used GA in the calibration phase, and the size of the initial population of ensembles. The most remarkable observation was that by increasing the size of an EPS, better prediction results are not assured. Actually this happens because EPS results represent an average of the predictions of all ensemble members and, knowing that these members are variated regarding their closure parameters in a random way, using more members does not assure less average error. In contrast, increasing initial ensemble size, which will be calibrated iteratively by the G-Ensemble, provides better prediction results as observed in the same figure. That is, by increasing the initial ensemble size in G-Ensemble, the probability for finding better solutions through GA iterations, also increases.

In the next Chapter, the overall conclusions, as well as the open lines for future research are provided.

Chapter 5

Conclusions and Open Lines

Accurate numerical weather forecasting is of great importance. Indeed, the need for reliable predictions in environmental modelling is long known. Particularly, the predicted weather information about the future atmospheric state is crucial and necessary for almost all other areas of environmental modelling. Additionally, right decisions to prevent damages and save lives could be taken depending on a reliable weather prediction process.

Due to inadequate observations, our limited understanding of the physical processes of the atmosphere, and the chaotic nature of atmospheric motion, uncertainties always exist in modern numerical weather prediction (NWP). In recent years, much progress has been made in building more precise and sophisticated NWP models, However, weather prediction as by its nature, is mainly, is an initial value problem. That is, lack and uncertainty of input data and parameters constitute the main source of errors for NWP.

But beyond that fundamental problem of the uncertainty in input data and parameters, NWP models are considered as soft-real time applications. The importance of having a a certain degree of accuracy in the prediction in reasonable time period is a real challenge. Thus, ongoing research concentrate on methods to enhance the process of prediction and to get results of this process faster.

In recent years, evolutionary optimization methods have become popular to solve the input parameter problem of environmental models. Actually, its is well-known, that parameter calibration methods can improve prediction quality. It has therefore been the primary objective of this thesis to propose a valid framework for an automatic calibration of input parameters in NWP models, by which, forecast skill could be improved so far. As well, it is aimed that, the proposed solution is to be feasible considering the waiting time needed for better weather predictions.

5.1 Conclusions

During the course of this thesis, numerical weather prediction (NWP) models and their functionality were described. The accuracy problem in NWP models was also highlighted as well as, the importance of having a certain level of accuracy within a reasonable time to have prediction outcomes.

A new weather predictions scheme; Genetic Ensemble (G-Ensemble) was proposed as an automatic calibration framework for NWP models, to deal with the problem of uncertainty in model closure parameters, with the main objective of improving predictability of short and medium range weather predictions. This scheme uses an evolutionary optimization method, concretely, a Genetic Algorithm, by which, NWP model closure parameters are calibrated and optimized.

Additionally, G-Ensemble scheme was paralleled using a Master/Worker programming paradigm in order to get benefit from parallel and distributed computing architectures, and thus, reducing the waiting time for enhance weather predictions.

Then, the proposed scheme was evaluated by experimentation over a real weather cases, Hurricane Katrina, which occurred in the Gulf of Mexico in 2005. The results obtained in the experiments approve our hypothesis that, better estimation of NWP model closure parameter values enhances weather prediction quality.

Summarizing, the proposed G-Ensemble could be considered as a successful approach that estimates optimal NWP model closure parameter values in order to improve forecast skill in NWP models, due to the sensitivity of these models to variations in the value of these parameters.

Furthermore, G-Ensemble was extended, in order to consider more than observation point in the evaluation of forecasts during calibration phase. This addition enables the Genetic Algorithm, which is used during calibration phase, to make better decisions when selecting between forecasts through its iterations. By introducing this capability to the proposed scheme, it was shown by experiments, that forecast skill is improved while no computational cost is added.

Additionally, the proposed approach is an automatic calibration method, which not only calibrates a certain pre-defined set of parameters, but rather, it provides improvement in forecast skill independently of domain particular features, like the habitual heterogeneity of surface characteristics exhibited in weather domains. G-Ensemble approach can find the relative parameters to be calibrated, it classifies them according to their weights in the certain domain, and furthermore, it provides the ability to calibrate more than one level of these parameters.

The paralleled G-Ensemble scheme showed a significant improvement in prediction quality. Moreover, in scenarios of limited number of computing resources, it constitutes a good solution that guarantees a significant enhancement in meteorological prediction and, an overall reduction of execution time.

FInally, a list of publications have been realized corresponding to the progress during this research. At the beginning, a study had been conducted discussing the necessity of improving prediction quality in numerical weather predictions, as well as the sensitivity of weather prediction model to its intrinsic parameters. The results of these studies were published as follow:

- H. Ihshaish, A. Sairouni, A. Cortés, and M. A. Senar, La necesidad de Mejoras en Predicción para el Modelo MM5 , in Proceedings of the XX Jornadas de Paralelismo, La Coruna, Spain, September 2009.
- H. Ihshaish, A. Sairouni, A. Cortés, and M. A. Senar, MM5: Computational and Prediction Improvements , in Proceedings of 3d Palestinian International Conference on Computer and Information Technology, Hebron, Palestine, March 2010.

Later, a sensitivity study discussing the effect of model closure parameters on prediction results of NWP models, along with the first proposal of G-Ensemble scheme were presented and evaluated in:

- H. Ihshaish, A. Cortés, and M. A. Senar, Genetic Ensemble (G- Ensemble) for Meteorological Prediction Enhancement , in Proceedings of The 2011 Internacional Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2011), Las Vegas (US)., Ed., H. R. Arabnia, vol. 1, pp. 404-4010, July 2011.

Then, the parallel version of the proposed scheme, as well as the implementation of the Multi-Level G-Ensemble approach were published in:

- H. Ihshaish, A. Cortés, and M. A. Senar, Parallel Multi-Level Genetic Ensemble for Numerical Weather Prediction Enhancement , Procedia Computer Science, ELSEVIER, vol. 9, pp. 276-285, 2012 (Proceedings of the International Conference on Computational Science (ICCS 2012), Omaha, Nebraska, US., June 2012).

Additionally, the following work was published discussing the G-Ensemble approach, in which the calibration phase was extended to consider a window of observations, rather than one-point observation in its evaluation process during calibration:

- H. Ihshaish, A. Cortés, and M. A. Senar, Towards Improving Numerical Weather Predictions by Evolutionary Computing Techniques , Procedia Computer Science, ELSEVIER, vol. 9, pp. 276-285, 2012 (Proceedings of the International Conference on Computational Science (ICCS 2012), Omaha, Nebraska, US., June 2012).

Finally, tuning G-Ensemble, discussing its prediction results according to different GA configurations was accepted for publication in:

- H. Ihshaish, A. Cortés, and M. A. Senar, Tuning G-Ensemble to Improve Forecast Skill in Numerical Weather Prediction Models , in Proceedings of The 2012 Internacional Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2012), Las Vegas (US), July, 2012.

5.2 Open Lines

One of the remarkable conclusions obtained by the realization of this work is that, it was strongly approved that an automatic calibration method, which intends to estimate optimal values of NWP physical parameters is crucial to achieve more improvements in todays NWP models. Especially, due to the fact that today s skillful NWP models contain large amounts of tunable parameters, which makes it almost impossible to optimize their values manually.

Being that said, many important issues are opened for further research, these may include:

An implementation of another evaluation method (fitness function) aiming at providing enhancement in prediction quality related to a temporal precision. As such, to be able to reduce prediction errors considering the precision of each forecast, not only quantitatively, but also temporally. Obviously, a forecast which detects exactly the time of a coming storm might be more valuable than another forecast which was precise in predicting wind velocity, but not at the exact time, i.e., 6 hours later for example.

Another valuable research would be to provide statistical measures, basing on large amount of experiments, to asses to what extent the forecast skill would be improved for a certain weather prediction within different conditions and basing on different configurations, like the number of the available computing resources, the number of the parameters to be calibrated, the number of GA iterations, and the size of the selected population (number of initial ensemble members).

It would be also of great value to enhance G-Ensemble, by coupling it with operational ensemble prediction systems, that are dealing with initial conditions rather than physical parameters. In doing so, the resulted work would be an important contribution.

Finally, an important work would be to develop scheduling strategies or policies in order to better distribute the parallel workload of the proposed parallel application over the available computing environment.

Bibliography

[1] UK Met Office - Tropical Cyclone Facts, Accessed:10/6/2012. [Online]. Available: http://www.metoffice.gov.uk/weather/tropicalcyclone/facts

[2] Weather Forecasting - Wikipedia, the free encyclopedia, Accessed: 10/6/2012. [Online]. Available: http://en.wikipedia.org/wiki/Weather_forecasting

[3] National Weather Service (NWS), Accessed:10/6/2012. [Online]. Available: http://www.nws.noaa.gov/

[4] P. Lynch, The origins of computer weather prediction and climate modeling, *Journal of Computational Physics*, vol. 227, no. 7, pp. 3431 3444, 2008.

[5] E. Lorenz, Deterministic non-periodic flow, *Journal of Atmospheric Sciences*, vol. 20, pp. 130 141, 1963.

[6] T. N. Palmer, A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models, *Quarterly Journal of the Royal Meteorological Society*, vol. 127, no. 572, pp. 279 304, 2001.

[7] Weather Research and Forecasting Model, Accessed:10/6/2012. [Online]. Available: http://www.wrf-model.org/index.php

[8] The Fifth-Generation Penn State/NCAR Mesoscale Model, Accessed:10/6/2012. [Online]. Available: http://www.mmm.ucar.edu/mm5/

[9] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, 1st ed. Cambridge University Press, 2003, ch. 6: Atmospheric predictability and ensemble forecasting, pp. 205 253.

[10] B. Wang, X. Zou, and J. Zhu, Data assimilation and its applications, *Proceedings of the National Academy of Sciences*, vol. 97, no. 21, pp. 11 143 11 144, 2000.

[11] The ECMWF Ensemble Prediction System, Accessed:10/6/2012. [Online]. Available: http://www.ecmwf.int/research/predictability/projects/index.html

[12] T. N. Palmer and P. D. Williams, Introduction. stochastic physics and climate modelling, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 366, no. 1875, pp. 2419 2425, 07 2008.

[13] J. Teixeira and C. A. Reynolds, Stochastic Nature of Physical Parameterizations in Ensemble Prediction: A Stochastic Convection Approach, *Monthly Weather Review*, vol. 136, no. 2, pp. 483 496, Feb. 2008.

[14] Shoder. Cserd: What is Computational Science? Accessed:10/6/2012. [Online]. Available: http://www.shodor.org/cserd/Help/whatiscs

[15] G. Schulze, Atmospheric observations and numerical weather prediction, *South African Journal of Science*, vol. 103, pp. 318 323, 08 2007.

[16] W. Ruddiman, *Earth's Climate: Past and Future.* W.H. Freeman, 2001.

[17] Vilhelm Bjerknes - Wikipedia, the free encyclopedia, Accessed:10/6/2012. [Online]. Available: http://en.wikipedia.org/wiki/Vilhelm_Bjerknes

[18] V. Bjerknes, Das problem der wettervorhersage, betrachtet vom standpunkte der mechanik und der physik, *Meteor. Zeit.*, vol. 21, pp. 1 7, 1904, Translation by Y. Mintz: The problem of weather forecasting as a problem in mechanics and physics. Los Angeles, 1954.

[19] L. F. Richardson - Wikipedia, the free encyclopedia, Accessed:10/6/2012. [Online]. Available: http://en.wikipedia.org/wiki/Lewis_Fry_Richardson

[20] Numerical weather prediction - Wikipedia, the free encyclopedia, Accessed:10/6/2012. [Online]. Available: http://en.wikipedia.org/wiki/Numerical_weather_prediction

[21] P. Lynch, *The Emergence of Numerical Weather Prediction: Richardson's Dream.* Cambridge University Press, 2006, ch. 2: The Fundamental Equations, pp. 45 46.

[22] R. A. Pielke, *Mesoscale Meteorological Modeling*, 2nd ed. Academic Press, 2002, ch. 6: Coordinate Transformations, pp. 130 132.

[23] Grid Length and Resolution in Weather Prediction, Accessed:10/6/2012. [Online]. Available: http://weather.mailasail.com/Franks-Weather/Grid-Length-Resolution

[24] R. A. Pielke, *Mesoscale Meteorological Modeling*, 2nd ed. Academic Press, 2002, ch. 11.1: Grid and Domain Structure, pp. 347 363.

[25] B. G. White, J. Paegle, W. J. Steenburgh, J. D. Horel, R. T. Swanson, L. K. Cook, D. J. Onton, and J. G. Miles, Short-term forecast validation of six models, *Weather and Forecasting*, vol. 14, no. 1, pp. 84 108, 1999.

[26] Y. Ghile and R. Schulze, Evaluation of three numerical weather prediction models for short and medium range agrohydrological applications, *Water Resources Management*, vol. 24, no. 5, pp. 1005 1028, 03 2010.

[27] W. Skamarock, Evaluating mesoscale nwp models using kinetic energy spectra, *Monthly Weather Review*, vol. 132, no. 12, pp. 3019 3032, 2004.

[28] NOAA Global Forecasting System (GFS), Accessed:10/6/2012. [Online]. Available: http://www.emc.ncep.noaa.gov/index.php?branch=GFS

[29] The European Centre for Medium-Range Weather Forecasts (ECMWF), http://www.ecmwf.int/products/forecasts, Accessed:10/6/2012.

[30] The North American Mesoscale Model, Accessed:10/6/2012. [Online]. Available: http://www.emc.ncep.noaa.gov/index.php?branch=NAM

[31] The High Resolution Limited Area Model, Accessed:10/6/2012. [Online]. Available: http://hirlam.org/

[32] The ALADIN Numerical Weather Prediction Model, Accessed:10/6/2012. [Online]. Available: http://www.cnrm.meteo.fr/aladin/

[33] R. A. Pielke, *Mesoscale Meteorological Modeling*, 2nd ed. Academic Press, 2002, ch. 11.2: Boundary and Inictial Conditions-Initialization, pp. 364 378.

[34] *Numerical Weather and Climate Prediction.* Cambridge University Press, 2011, ch. 6.2: Model Initialization-Observations used for model initialization, pp. 199 209.

[35] The World Meteorological Organization, Accessed:10/6/2012. [Online]. Available: http://www.wmo.int/

[36] D. Middleton, T. Scheitlin, and B. Wilhelmson, *Visualization in Weather and Climate Research.* Academic Press, 2005, ch. 44, pp. 845 871.

[37] L. A. Treinish, How Can We Build More Effective Weather Visualizations? in *Eighth Workshop on Meteorological Operational Systems, ECMWF, Reading, UK.* European Centre for Medium-Range Weather Forecasting, 2001, pp. 90 99.

[38] B. W. Shen, W. K. Tao, and B. Green, Coupling Advanced Modeling and Visualization to Improve High-Impact Tropical Weather Prediction, *Computing in Science & Engineering*, vol. 13, no. 5, pp. 56 67, 2011.

[39] M. Koutek, 3D visualization of Hirlam data and use of virtual reality interfaces, in *European Working Group on Operational meteorological Workstations (EGOWS) 2009 Meeting, De Bilt, Netherlands.*, 2009.

[40] V. Lakshmanan, R. Rabin, J. Otkin, J. S. Kain, and S. Dembek, Visualizing model data using a fast approximation of a radiative transfer model, *Journal of Atmospheric and Oceanic Technology*, vol. 29, no. 5, pp. 745 754, 2012/06/11 2012.

[41] W. L. Hibbard, B. E. Paul, D. A. Santek, C. R. Dyer, A. L. Battaiola, and M. F. V. Martinez, Interactive Visualization of Earth and Space Science Computations, *Computer*, vol. 27, no. 7, pp. 65 72, Jul. 1994.

[42] S. W. Berrick, G. Leptoukh, J. D. Farley, and H. Rui, Giovanni: A Web Service Workflow-Based Data Visualization and Analysis System, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 106 113, Jan. 2009.

[43] M. J. Miller, T. N. Palmer, and R. Swinbank, Parametrization and influence of subgridscale orography in general circulation and numerical weather prediction models, *Meteorology and Atmospheric Physics*, vol. 40, pp. 84 109, 1989.

[44] T.-Y. Lee and S.-Y. Hong, Physical parameterization in next-generation nwp models, *Bulletin of the American Meteorological Society*, vol. 86, no. 11, pp. 1615 1618, 2005.

[45] T. T. Warner, *Numerical Weather and Climate Prediction*, 1st ed. Cambridge University Press, 2011, ch. 4.1: Physical-process parameterizations: Background, pp. 119 121.

[46] D. J. Stensrud, *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models.* Cambridge University Press, 2007.

[47] , *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*, 1st ed. Cambridge University Press, 2007, ch. 1: Why study parameterization schemes?, pp. 1 7.

[48] The Fifth-Generation Penn State/NCAR Mesoscale Model (MM5) V.3 Tutorial, Accessed:10/6/2012, pp.15-17. [Online]. Available: http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes-pdf/mm5.pdf

[49] A. J. Litta, S. M. Idicula, and U. C. Mohanty, A comparative study of convective parameterization schemes in wrf-nmm model, *International Journal of Computer Applications*, vol. 33, no. 6, pp. 32 40, November 2011, published by Foundation of Computer Science, New York, USA.

[50] W. Wang and N. L. Seaman, A comparison study of convective parameterization schemes in a mesoscale model, *Monthly Weather Review*, vol. 125, no. 2, pp. 252 278, 2012/06/16 1997.

[51] M. Choi, J. M. Jacobs, and W. P. Kustas, Assessment of clear and cloudy sky parameterizations for daily downwelling longwave radiation over different land surfaces in florida, usa, *Geophysical Rsearch Letters*, vol. 35, no. 20, p. L20402, 10 2008.

[52] X.-M. Hu, J. W. Nielsen-Gammon, and F. Zhang, Evaluation of three planetary boundary layer schemes in the wrf model, *Journal of Applied Meteorology and Climatology*, vol. 49, no. 9, pp. 1831 1844, 2012/06/16 2010.

[53] T. N. Krishnamurti, S. Basu, J. Sanjay, and C. Gnanaseelan, Evaluation of several different planetary boundary layer schemes within a single model, a unified model and a multimodel superensemble, *Tellus A*, vol. 60, no. 1, pp. 42 61, 2008.

[54] H. Badr, H. Kandil, B. Elhadidi, and A. Sherif, Evaluating the physics options of regional weather models for areas with complex land-use characteristics, in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, july 2011, pp. 3257 3260.

[55] R. Avissar and R. A. Pielke, A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology, *Monthly Weather Review*, vol. 117, no. 10, 1989.

[56] F. Ament and C. Simmer, Improved representation of land-surface heterogeneity in a non-hydrostatic numerical weather prediction model, *BoundaryLayer Meteorology*, vol. 121, no. 1, pp. 153 174, 2006.

[57] K. W. Oleson, K. L. Driese, J. A. Maslanik, W. J. Emery, and W. A. Reiners, The sensitivity of a land surface parameterization scheme to the choice of remotely sensed land-cover datasets, *Monthly Weather Review*, vol. 125, no. 7, pp. 1537 1555, 1997.

[58] M. L. Kavvas, Z. Q. Chen, L. Tan, S. T. Soong, A. Terakawa, Y. J., and K. Fukami, A regional-scale land surface parameterization based on areally-averaged hydrological conservation equations, *Hydrological Sciences Journal*, vol. 43, no. 4, pp. 611 631, 1998.

[59] Z. Ying, Coupling land surface process model with the weather research forecast modeling system, Ph.D. dissertation, Scuola di Dottorato in Scienza e Alta Tecnologia, Università degli Studi di Torino, Turin, Italy, 2010.

[60] C. M. Godfrey and D. J. Stensrud, An empirical latent heat flux parameterization for the noah land surface model, *Journal of Applied Meteorology and Climatology*, vol. 49, no. 8, pp. 1696 1713, 2010.

[61] K. W. Oleson, K. L. Driese, J. A. Maslanik, W. J. Emery, and W. A. Reiners, The sensitivity of a land surface parameterization scheme to the choice of remotely sensed land-cover datasets, *Monthly Weather Review*, vol. 125, no. 7, pp. 1537 1555, 1997.

[62] U.S. Geological Survey, Global Land Cover Characterization Version 1.2, Accessed:10/6/2012. [Online]. Available: http://edc2.usgs.gov/glcc/glcc.php

[63] J. F. Brown, T. R. Loveland, J. W. Merchant, B. C. Reed, and D. O. Ohlen, Using multisource data in global land cover characterization: concepts, requirements and methods, *American Society for Photogrammetry and Remote Sensing Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 6, pp. 977 987, 1993.

[64] European Comission, European Soil Portal-Soil Data and Information Systems, Accessed:10/6/2012. [Online]. Available: http://eusoils.jrc.ec.europa.eu/ESDB_Archive/soil_data/global.htm

[65] The Community Noah Land-Surface Model, Accessed:10/6/2012. [Online]. Available: http://gcmd.nasa.gov/records/NOAA_NOAH.html

[66] S. Yoden, Atmospheric predictability, *Journal of the Meteorological Society of Japan. Ser. II*, vol. 85B, pp. 77 102, 2007.

[67] E. Lorenz, The predictability of hydrodynamics flow, *Transactions of the New York Academy of Sciences Series II*, vol. 25, no. 4, pp. 409 432, 1963.

[68] M. Leutbecher and T. N. Palmer, Ensemble forecasting, *J. Comput. Phys.*, vol. 227, no. 7, pp. 3515 3539, Mar. 2008.

[69] H. Zhang and Z. Pu, Beating the uncertainties: Ensemble forecasting and ensemble-based data assimilation in modern numerical weather prediction, *Advances in Meteorology*, vol. 2010, Article ID 432160, p. 10 pages, 2010.

[70] Z. Pu, S.-Y. Hong, Y. Li, and H.-M. H. Juang, Advanced data assimilation and predictability studies on high-impact weather and climate, *Advances in Meteorology*, vol. 2010, Article ID 121763, p. 2 pages, 2010.

[71] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, 1st ed. Cambridge University Press, 2003, ch. 1: Historical overview of numerical weather prediction, pp. 1 31.

[72] A. Weerts and Y. Liu, Advances in data assimilation for operational hydrologic forecasting, *Eos Trans. AGU*, vol. 92, no. 7, p. 55, 02 2011.

[73] D. M. Barker, W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. N. Xiao, A three-dimensional variational data assimilation system for MM5: Implementation and initial results, *Monthly Weather Review*, vol. 132, no. 4, pp. 897 914, 2004.

[74] D. Zupanski, M. Zupanski, E. Rogers, D. F. Parrish, and G. J. DiMego, Fine-resolution 4dvar data assimilation for the great plains tornado outbreak of 3 may 1999, *Weather and Forecasting*, vol. 17, no. 3, pp. 506 525, 2002.

[75] M. Coniglio, K. Elmore, J. Kain, S. Weiss, M. Xue, and M. Weisman, Evaluation of wrf model output for severe-weather forecasting from the 2008 noaa hazardous weather testbed spring experiment, *Weather and Forecasting*, vol. 25, no. 2, pp. 408 427, 2010.

[76] A. T. Weaver, J. Vialard, and D. L. T. Anderson, Three- and four-dimensional variational assimilation with a general circulation model of the tropical pacific ocean. part i: Formulation, internal diagnostics, and consistency checks, *Monthly Weather Review*, vol. 131, no. 7, pp. 1360 1378, 2003.

[77] P. Gauthier, M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, Extension of 3dvar to 4dvar: Implementation of 4dvar at the meteorological service of canada, *Monthly Weather Review*, vol. 135, no. 6, pp. 2339 2354, 2007.

[78] E. S. Epstein, Stochastic dynamic prediction, *Tellus*, vol. 6, pp. 739 759, 1969.

[79] V. V. Kharin and F. W. Zwiers, Climate predictions with multi-model ensembles, *Journal of Climate*, vol. 15, no. 7, pp. 793 799, 2012/05/25 2002.

[80] T. N. Krishnamurti, C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, Multimodel ensemble forecasts for weather and seasonal climate, *Journal of Climate*, vol. 13, no. 23, pp. 4196 4216, 2000.

[81] H. Jarvinen, M. Laine, A. Solonen, and H. Haario, Ensemble prediction and parameter estimation system: the concept, *Quarterly Journal of the Royal Meteorological Society*, vol. 138, no. 663, pp. 281 288, 2012.

[82] X. Zhang, R. Srinivasan, K. Zhao, and M. V. Liew, Evaluation of global optimization algorithms for parameter calibration of a

computationally intensive hydrologic model, *Hydrological Processes*, vol. 23, no. 3, pp. 430 441, 2009.

[83] T. Trucano, L. Swiler, T. Igusa, W. Oberkampf, and M. Pilch, Calibration, validation, and sensitivity analysis: What s what, *Reliability Engineering & System Safety*, vol. 91, no. 10 11, pp. 1331 1357, 2006.

[84] J. A. Vrugt, B. Ó. Nualláin, B. A. Robinson, W. Bouten, S. C. Dekker, and P. M. Sloot, Application of parallel computing to stochastic parameter estimation in environmental models, *Computers & Geosciences*, vol. 32, no. 8, pp. 1139 1155, 2006.

[85] R. E. Kalman, A new approach to linear filtering and prediction problems, *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35 45, 1960.

[86] A. Poyton, M. Varziri, K. McAuley, P. McLellan, and J. Ramsay, Parameter estimation in continuous-time dynamic models using principal differential analysis, *Computers & Chemical Engineering*, vol. 30, no. 4, pp. 698 708, 2006.

[87] K. He, S. Dong, and L. Zheng, Service-oriented grid computation for large-scale parameter estimation in complex environmental modeling, in *Proceedings of the 2006 ACM symposium on Applied computing*, ser. SAC 06. New York, NY, USA: ACM, 2006, pp. 741 745.

[88] E. G. Bekele and J. W. Nicklow, Evolutionary algorithms for multi-objective, automatic calibration of a semi-distributed hydrologic model, R. Graham, Ed., vol. 200, no. 40856. ASCE, 2006, pp. 112 112.

[89] M. K. Muleta, P. F. Boulos, C.-H. Orr, and J. J. Ro, Using genetic algorithms and particle swarm optimization for optimal design and calibration of large and complex urban stormwater management models, R. Graham, Ed., vol. 200, no. 40856. ASCE, 2006, pp. 113 113.

[90] A. Ostfeld, G. Gal, and E. Salomons, Calibration of caedym using a genetic algorithm the lake kinneret case study, R. Graham, Ed., vol. 200, no. 40856. ASCE, 2006, pp. 114 114.

[91] R. Cavill, S. L. Smith, and A. M. Tyrrell, Variable length genetic algorithms with multiple chromosomes on a variant of the onemax problem, in *Proceedings of the 8th annual conference on Genetic*

and evolutionary computation, ser. GECCO 06. New York, NY, USA: ACM, 2006, pp. 1405 1406.

[92] J. Peng and Z. S. Chu, A hybrid multi-chromosome genetic algorithm for the cutting stock problem, Los Alamitos, CA, USA, pp. 508 511, 2010.

[93] Q. J. Wang, Using genetic algorithms to optimise model parameters, *Environmental Modelling Software*, vol. 12, no. 1, pp. 27 34, 1997.

[94] B. Abdalhaq, A Methodology to Enhance the Prediction of Forest Fire Propagation, Ph.D. dissertation, Universitat Autònoma de Barcelona, Barcelona, Spain, 2004.

[95] B. Abdalhaq, A. Cortés, T. Margalef, and E. Luque, Enhancing wildland fire prediction on cluster systems applying evolutionary optimization techniques, *Future Generation Comp. Syst.*, vol. 21, no. 1, pp. 61 67, 2005.

[96] Z. Lin and D. E. Radcliffe, Automatic calibration and predictive uncertainty analysis of a semidistributed watershed model, *Vadose Zone J.*, vol. 5, no. 1, pp. 248 260, 2006.

[97] L. Zhu, J. M. Chen, Q. Qin, J. Li, and L. Wang, Optimization of ecosystem model parameters using spatio-temporal soil moisture information, *Ecological Modelling*, vol. 220, no. 18, pp. 2121 2136, 2009.

[98] C. J. Willmott and K. Matsuura, On the use of dimensioned measures of error to evaluate the performance of spatial interpolators, *International Journal of Geographical Information Science*, vol. 20, no. 1, pp. 89 102, January 2006.

[99] Parallel WRF V.3 Benchmark hoempage, http://www.mmm.ucar.edu/wrf/WG2/bench/, Accessed:10/6/2012.

[100] Hurricane Katrina, Accessed:10/6/2012. [Online]. Available: http://www.katrina.noaa.gov/