

International Conference on Computational Science, ICCS 2012

Parallel Multi-Level Genetic Ensemble for Numerical Weather Prediction Enhancement[☆]

Hisham Ihshaish^{1,*}, Ana Cortés, Miquel A. Senar

Departament d'Arquitectura de Computadors i Sistemes Operatius, Escola d'Enginyeria, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

Abstract

The need for reliable predictions in environmental modelling is well-known. Particularly, the predicted weather and meteorological 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 meteorological prediction process. Lack and uncertainty of input data and parameters constitute the main source of errors for most of these models. In recent years, evolutionary optimization methods have become popular to solve the input parameter problem of environmental models. We propose a new parallel meteorological prediction scheme that uses evolutionary optimization methods based on Multi-Chromosome Genetic Algorithm to enhance the quality of weather forecasts by focusing on the calibration of input parameters. This new scheme is parallelized and executed on a HPC environment in order to reduce the time needed to obtain the final prediction. The new approach is called Multi-Level Genetic Ensemble (*M-Level G-Ensemble*) and it has been tested using historical data of a well-known weather catastrophe: Hurricane Katrina that occurred in 2005 in the Gulf of Mexico. Results obtained with our approach provide both significant improvements in weather prediction and a significant reduction in the execution time.

Keywords: Numerical weather prediction; Evolutionary computing; HPC; Ensemble prediction; Parameter estimation

1. Introduction

The pre-knowledge of the atmosphere future state has been continuously demanded for thousands of years. Agriculture, education, entertainment, industry, astronomy, etc. usually benefit from an accurate knowledge of the weather future state. Weather time evolution is represented by Numerical Weather Prediction models (NWP models) that are commonly solved by means of computing facilities. Precisely, it was in the early 1950s, when the USA National Weather Service (NWS)[1] began to utilize some of the early versions of computers to make large-scale weather forecasts, running NWP models.

[☆]This research has been supported by the MICINN-Spain under contract TIN 2007-64974.

*Corresponding author:

Email address: hisham.ihshaish@caos.uab.es (Hisham Ihshaish)

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 [2].

Computationally, NWP models are considered as soft-real time applications. The importance of having a certain degree of accuracy in the prediction in a certain time is a real challenge. Thus, ongoing research concentrate on methods to enhance the process of prediction and to get results of this process faster. As most simulation software works with well-founded and widely accepted models, the need for input parameter optimization to improve model output is a well-known and often-tackled problem. Particularly, in 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 behavior. 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.

This work presents a new parallel meteorological prediction scheme that uses evolutionary optimization methods to enhance the quality of weather forecast by focusing on the calibration of model input parameters. This new scheme is suitable to be run in HPC environments and thus, it aims to improve prediction quality while also reducing total execution time.

The rest of the paper is organized as follows: Section 2 gives an overview of NWP models, a NWP general scheme, and a brief description of the Weather Research and Forecasting Model (WRF), which constitutes the most commonly used model for weather and meteorological predictions. Section 3 focuses on the importance of accuracy in NWP models and also describes the most widely used methods for NWP enhancement in practice. In section 4, the proposed prediction scheme (*M-Level G-Ensemble*) is presented and described. Section 5 discusses experimental results obtained with a test case, where we compare our proposal with other enhancement methods. Section 6 describes the parallelization of the proposed scheme along with performance results of execution time on a cluster of computers. Finally, conclusions and future work are described in section 7.

2. Numerical Weather Prediction

Weather stems from the constant evolution of the atmosphere governed by physical laws. Using high-speed computers to solve a complex set of mathematical equations that represents the governing laws, NWP is a technique for simulating the atmospheric evolution in order to delineate the resultant weather changes. The variables involved in the equations include wind, temperature, pressure and moisture content. In principle, given the initial and boundary conditions, the atmospheric variables can be numerically solved as functions of time and form the basis of weather forecast. That is, NWP is described generally as "an initial-boundary value problem": - given an estimate of the present state of the atmosphere (initial conditions), and appropriate surface and lateral boundary conditions, the model simulates (forecasts) the atmospheric evolution. The more accurate the estimate of the initial conditions, the better the quality of the forecasts.

Certain areas where atmosphere future conditions are to be predicted are represented by three-dimensional uniform-gridded-rectangles referred as domains. The input data, which describe an estimation of the actual state of the atmosphere, are called initial conditions. Those initial conditions are assigned to all points of the grid. The horizontal distance between grid points is referred as the spatial resolution of both the initial conditions and prediction results. Regional models (also known as limited-area models, or LAMs) allow for the use of finer grid spacing (higher resolution) than global models because the available computational resources are focused on a specific area instead of being spread over the globe. This allows regional models to resolve explicitly smaller-scale meteorological phenomena that can not be represented on the coarser grid of a global model. Hence, a NWP model will predict the new values of the initial conditions over future time scale.

The first step of a NWP process is to extract initial conditions that are usually obtained from a global forecasting. These initial conditions are assigned to the domain grid points and, by means of the NWP model applied over a time line, at each pre-defined time period, a new 3-dimensional domain is produced having new (predicted) values of meteorological variables at all grid points.

The Weather Research and Forecasting model (WRF) [3] is a widely-used numerical weather prediction system, which is considered as a next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. WRF is composed of a variety of programs to facilitate the prediction process. It includes modules for global terrain data extraction, modules for real observation injection while model integration, and modules for output post-processing. It should be mentioned that although we have applied our methodology to WRF, the proposed strategy is a model-independent design, which could also be used with other existing NWP models such as the PSU/NCAR Mesoscale Model [4] known as (MM5).

3. Related Work

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. 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., [5]). 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.

As it has been stated before, in the case of initial conditions, input data is extracted from global forecasts. Normally, global forecasts are conducted using domains of lower grid resolutions (the distance between grid points is large). This is due to the computational power needed if the whole globe is to be predicted using finer grid spacing. As a result, interpolations are needed to extract initial conditions from lower resolution domains to assign them to local domains of higher resolution. Unfortunately, this process is not perfect and the assigned values do not reflect the actual real state of the atmosphere. This problem is generally referred as the uncertainty of weather initial state.

On the other hand, physical parametrization is the representation of sub-grid scale physical processes, that is, some meteorological processes are too small-scale to be explicitly included in NWP models. Hence, parametrization enables the representation of these processes by relating them to variables on the scales (the points of the gridded domain) that the model resolves. For example, an important meteorological process is the surface flux of energy transmitted by the terrain which helps in enhancing the prediction of other important variables like near-surface temperature, sea surface temperature and even near-surface wind velocity variables. This process normally occurs in scales smaller than 1 kilometer, while NWP models predicts normally on domains of grid-scales higher than 1 kilometer. Parametrization is needed in such cases to represent this process on a certain domain scale.

By figuring out the main sources of error in predictability of NWP models, and over the past 20 years or so, stochastic or "ensemble" forecasting [6] became as a practical and successful way of addressing the predictability problem associated with the uncertainty in initial conditions. Early on moreover, several weather prediction centers have addressed this problem by developing operational ensemble prediction systems (EPS) (e.g., [7]). The main idea behind an EPS comes from the fact that the initial state of a certain variable should be seen as a probability distribution and not as a unique value, and thus, the ultimate goal of ensemble forecasting is to predict quantitatively the probability density of the state of the atmosphere at a future time. This is done by running multiple forecasts, each of which is initiated with small perturbations in the estimated initial conditions. Then, an ensemble forecast is usually evaluated in terms of an average of the individual forecasts (ensemble members) concerning one forecast variable, as well as the degree of agreement between various forecasts within the ensemble system, as represented by their overall spread [8].

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 [9, 10].

4. Multi Level Genetic Ensemble (M-Level G-Ensemble)

In this section, our Multi Level Genetic Ensemble (*M-Level G-Ensemble*) approach for prediction enhancement is described. Although *M-Level G-Ensemble* uses the same principles of the EPS, it clearly differs in the way of how ensemble members are obtained and computed. As mentioned before, the main idea of an EPS is to run multiple forecasts, each of which is initiated with small perturbations in the estimated initial conditions. Then, an ensemble forecast is usually evaluated in terms of an average of the individual forecasts. In the contrary, the presented methodology uses the same principles of an EPS, but the perturbations are used for physical parameters rather than for initial conditions. Moreover, the objective is to find the best possible values of these parameters and then, to use them in a single deterministic forecast, initiated with the same initial conditions, but with a better set of physical parameters. In this work, we focus on finding the best values of land surface physical parameterizations (*Landuse* and *Soil* parameters), which are provided to a NWP to calculate the evolution of some subgrid-scale variables in order to better predict near-surface meteorological variables. To achieve this objective, we propose a new scheme of prediction, shown in Fig.(1) where we introduce two pre-prediction phases, called Parameter Selection Phase and Calibration Phase. Hence, the whole prediction process will be formed of three stages: Parameter Selection, Calibration, and Prediction, which are described in the next subsections.

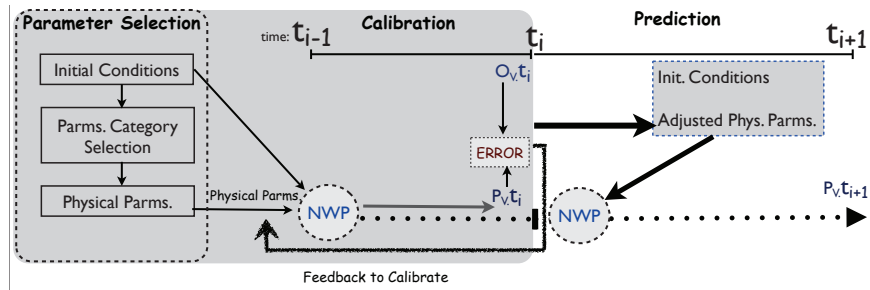


Figure 1: Three-phase prediction scheme; t_i is time 00:00 of prediction process, t_{i-1} is a time instant previous to Prediction Phase (initial time of Calibration Phase), t_{i+1} is the future time to be predicted. "O_v" is an observed meteorological variable at time t_i , "P_v" is the predicted variable at the same time using a NWP model.

4.1. Parameter Selection Phase

Before starting the process to enhance prediction, the parameters to be optimized should be selected. We focus our study on *Landuse* and *Soil* parameter values. They will serve as a prove of concept of our method, which could be applied to other parameters. Then, by now *Landuse* and *Soil* parameter values are the ones which will be optimized by our method. However, these parameters correspond to various globally pre-defined categories. Some of these categories of *Landuse* and *Soil* compose the terrain of a certain domain where prediction is to be conducted.

As it has been previously described, NWP model starts a process of prediction over a certain zone using the initial meteorological conditions defined by their location (longitude, latitude and vertical distance) for each grid point of the domain. Within these initial conditions and, for the first mesh of the grid (the surface or terrain mesh) each grid point is assigned with a number indicating its land use category (LU-index) and with another number indicating its soil type (SLTYP). During simulation, the NWP model needs surface parameter values for each surface grid point in order to calculate the evolution of the other meteorological variables. Actually, these parameter values depend on their categories, and for each category the NWP model is provided by its default parameter values provided in stand-alone tables. Then, for each surface grid point, the NWP model reads its assigned land use 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 parameters.

There are 33 land use categories, each of which has 7 surface physical parameters, and there are also 19 soil types, each of which has 10 physical parameters (see [11] for more description of land use and soil parameters). Prediction enhancement must look for optimal values of these input parameters. Therefore, the first step consists on selecting which category or categories correspond to the region terrain where meteorological prediction process is performed. In other words, it is necessary to determine exactly what class of parameters to optimize. For example, it does not make sense to optimize Grassland *Landuse* parameter values in a region that has 100% of it's terrain as water.

At this phase, a small program is developed to read domain initial conditions of 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 land use category and soil type. As a result, a table is constructed including each land use and soil type category, and each of which will have the number of domain grid points having that category. Table registers are ordered in descending order: the first land use category mostly repeated within the domain grid points is referred as the first dominant land use category, the second is the second dominant, and so on. The same is done for the soil type categories. Table (1) 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 weights (how often they are repeated between domain grid points). The first dominant category parameters of both *Landuse* and *Soil* parameters is referred to as first level parameters. It's 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.

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	17%
4	Rest	42%	Rest	19%

Table 1: Parameter Category Selection: Level (1) register contains the land use category (15) which covers 37% of the domain and the soil type category (6) which covers 41% of the same domain.

4.2. Calibration Phase

As shown in Fig.(1), t_i is the instant time from which the meteorological variables are going to be predicted, Calibration Phase starts at a time prior to prediction time and ends at time 00:00 of prediction period, i.e. calibration is done within the period (t_{i-1}, t_i) . That is, as in EPS, we initialize a set of simulations randomly, each of which has a different physical parameterization combination. This initial set, which we call *initial ensemble*, is run by the NWP model to predict meteorological variables at time t_i , then we use Genetic Algorithms (GA) to obtain an improved ensemble set (which has less errors compared to observations at time t_i) and the process is repeated again many times to a certain number of iterations.

At the last iteration of the GA, Calibration Phase exits with calibrated ensemble members that we refer to as *G-Ensemble*, each of which has a calibrated combination of physical parameters, which produced less error (their average error) than those of the *initial ensemble*. At that point, we select the ensemble member of the *G-Ensemble* with minimum error, to be the single ensemble member of the simulation that will conduct the Prediction Phase. We refer to this approach as *Best Genetic Ensemble Member (BeGEM)*.

As described in the Parameter Selection phase, there are more than one level of physical parameters which belong to a certain region where a prediction is going to take place. Applying a classical GA method, one level of these physical parameters could be calibrated by considering this single level of parameters as an individual in an initial population of individuals each of which with a combination of parameters that belong to one single category of physical parameters (one category of *Landuse* and *Soil* parameters). However, domains normally exhibit more than one category of *Landuse* and *Soil* parameters. If a calibration process with one level of parameters enhances prediction quality then results produced by a calibration process with more levels will be even better. For such cases, a multi-chromosome GA (e.g., [12, 13]) is introduced where more than one level of parameters could be calibrated.

For example, suppose that a meteorological prediction is to be conducted in a certain domain. The domain physical parameters of its terrain are classified by the Parameter Selection phase according to its dominant *Landuse* and *Soil* categories as shown in Table (1).

A calibration process considering the first level parameters, i.e. 7 parameters of the the land use 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 Fig.(2(a)).

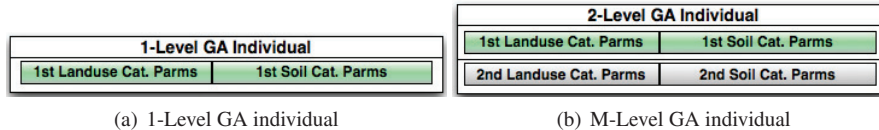


Figure 2: 1-Level GA individual in (a), and M-Level GA individual in (b).

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. However, to consider the parameters of the other levels of categories in the domain, a GA individual is constructed as shown in Fig(2(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. We call this approach as the *M-Level G-Ensemble*, and the output of the Calibration phase considering this approach is a single $BeGEM(x)$, where x refers to the number of parameter levels considered in the GA of the Calibration phase.

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, we propose two different error functions to be used, one referred as Single-Variable and the other referred 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.

4.2.1. Single-Variable G-Ensemble

The Calibration Phase is done with the goal of enhancing predictions for a single meteorological variable. The error function for the evaluation of ensemble members in our GA is the Root Mean Square Deviation *RMSD* or Error *RMSE*, shown below in equation (1). This error function is a frequently-used measure for the evaluation of meteorological predictions[14], which measures the differences between values predicted by a model or an estimator and the values actually observed from the variable being estimated.

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

$$RMS E = \sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{pre,i})^2}{n}} \quad (1)$$

Using *RMS E* error in the Calibration Phase limits our *G-Ensemble* to be oriented to enhance predictions for one meteorological 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*.

4.2.2. Multi-Variable G-Ensemble

The calibration is done with the goal of enhancing the prediction of multiple meteorological variables at the same time. To bypass the limitation imposed by *RMS E* error, we use the Normalized *RMS E*, see equation (2).

$$NRMS E = \frac{\sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{pre,i})^2}{n}}}{x_{obs(max)} - x_{obs(min)}} \quad (2)$$

The Normalized *RMS E* (referred as *NRMS E*) is the value of *RMS E* divided by the range of the observed values of a certain variable. *NRMS E* 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 *NRMS E* for all variables, and then, we consider the addition of all of them as the Multi-Variable error function. For example, the *NRMS E* of a model that predicts Temperature (*T*) and Precipitation (*P*) is the percentage obtained by the summation of two Percentages: *NRMS E(T)* and *NRMS E(P)*, as shown in equation (3).

$$Error = NRMS E(var1) + NRMS E(var2) = value\% \quad (3)$$

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.

4.3. Prediction Phase

Once the Calibration Phase is finished, it is the turn of the Prediction Phase. At this point, the *BeGEM(x)* produced by the previous phase will be run by the NWP model. It is expected that this ensemble member will generate better predictions as it shown less error in Calibration Phase. In contrast to the classical EPS, only one simulation is executed here, while in EPS the whole ensemble set is executed.

5. Experimental Evaluation

To test our approach, we used historical data of hurricane Katrina[15], which occurred on August 28, 2005 in the Gulf of Mexico and unfortunately caused the death of more than 1,800 persons along with a total property damage that was estimated at \$81 billion (2005 USD). The objective of the experiments is to predict meteorological variables evolution 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 meteorological variables is produced every 3 hours and the spatial resolution of the domain was 12km.

To get the evolution of meteorological variables at 12:00 h. of 28/08/2005, we used initial conditions of the atmospheric state in the zone three hours before, i.e. model started prediction from time 09:00 of 28/08/2005. For our approach (*M-Level G-Ensemble*), the Calibration Phase started from time 00:00 of 28/08/2005 to time 09:00 of the same day. The variables predicted in our experiments were: *Latent Heat Flux LHF (W/m2)*, *Surface Skin Temperature TSK (K)*, *2-meter Temperature (K)*, *10-meter Wind Velocity components U10 and V10 (m/s)*, and the *Accumulated Precipitation RAINC (mm)*.

In a previous work [16], we showed a significant improvement in meteorological prediction quality by reducing prediction error using *1-Level G-Ensemble* scheme, where only one level of physical parameters was optimized using the classical GA of one single chromosome. Here we show our results when several levels of parameters were optimized according to the method described in the previous section. Additionally, all these results of this new scheme are compared to the classical EPS.

Firstly, Fig.(3) shows experimental results for a classical EPS prediction of 40 ensemble members (each of which has a different *Landuse* and *Soil* parameters, a classical EPS here is referred to an ensemble prediction initiated by different, randomly selected *Landuse* and *Soil* parameters' values) to predict (every 3 hours) the evolution of: a) *Accumulated Precipitation RAINC*, b) *Latent Heat Flux LHF*, and c) *Surface Skin Temperature TSK*. The evolution of the values of these meteorological 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 Fig. (3(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. Thus, it could be easily concluded, that there is a significant margin of enhancement in prediction which could be achieved.

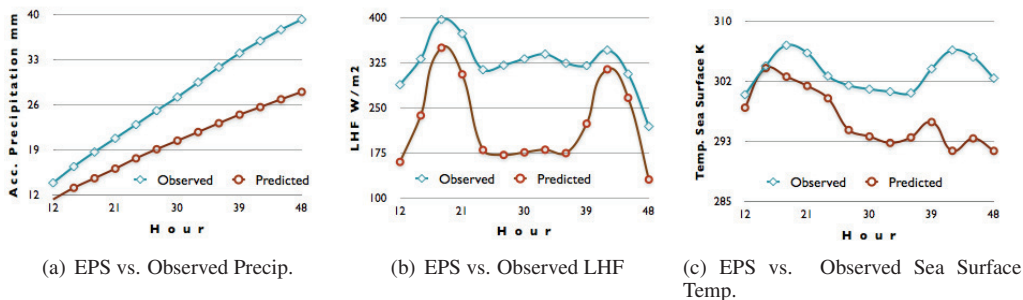


Figure 3: Classical EPS prediction results compared to observed values.

To use the *M-Level G-Ensemble* for prediction, Parameter Selection phase is run over the initial conditions of the domain of Katrina. Then, we applied our method (*M-Level G-Ensemble*) with *Single-Variable* error in two different cases: to predict Acc. Precipitation (results shown in Fig.4(a)) and to predict LHF (results shown in Fig.4(b)). In both cases, with the same initial ensemble members used in the EPS case, we obtained a significant improvement in prediction quality. Additionally, we also observed that better enhancements 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). Its three main operators were configured as follows: *Selection*: (best one of two) and (roulette), *Crossover*: (probability=0.7, type: two points crossover), and *Mutation*: (probability= 0.2).

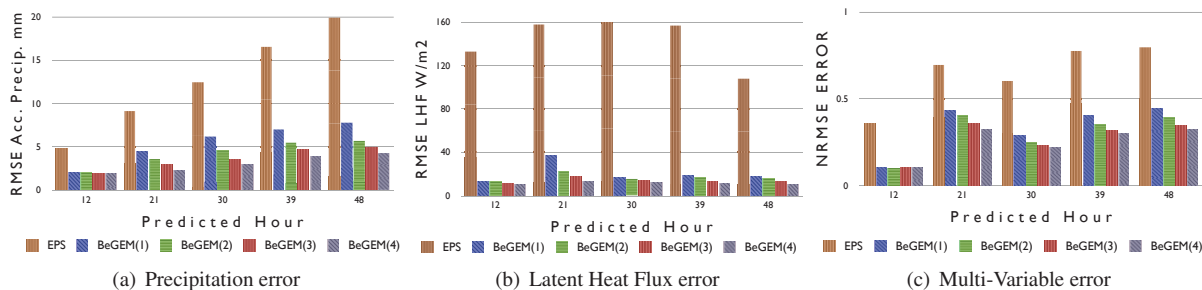


Figure 4: Single-Variable and Multi-Variable *M-Level G-Ensemble* ; (a): *RMSE* error in prediction of variable *Acc. Precipitation* and (b): variable *LHF*. Results are of classical *EPS* and the *BeGEM*(1, 2, 3, and 4) for both variables, (c):*Multi-Variable M-Level G-Ensemble*; *NRMSE* in prediction of variables: *LHF*, *Surface Skin Temperature*, *2-meter Temperature*, *10-meter Wind Velocity* components *U10* and *V10*, and the *Accumulated Precipitation*.

We also used our approach to enhance predictions of a set of meteorological variables at the same time, by applying the *Multi-Variable M-Level G-Ensemble* and using the error *NRMSE* (shown in equations 2 and 3) in Calibration Phase as the fitness function of the GA. In this case, we were also able to obtain significant improvements in the prediction of a set of meteorological variables at the same time. Fig.(4(c)) shows the results obtained in this case. Again, significant reduction of the *NRMS E* was obtained in the prediction of a set of meteorological variables together and, as more levels of parameters were calibrated, better results of prediction were obtained.

Additionally, we observed that a reduction in the *NRMS E* of a set of variables also provides an enhancement in the prediction of each meteorological variable alone. In other words, all variables were better predicted when *M-Level G-Ensemble* oriented to reduce the *NRMS E* of those variables together. To illustrate these results, we show in Fig.(5) how the corresponding prediction error of each variable was reduced when *M-Level G-Ensemble* was oriented to reduce the *NRMS E* of six variables together.

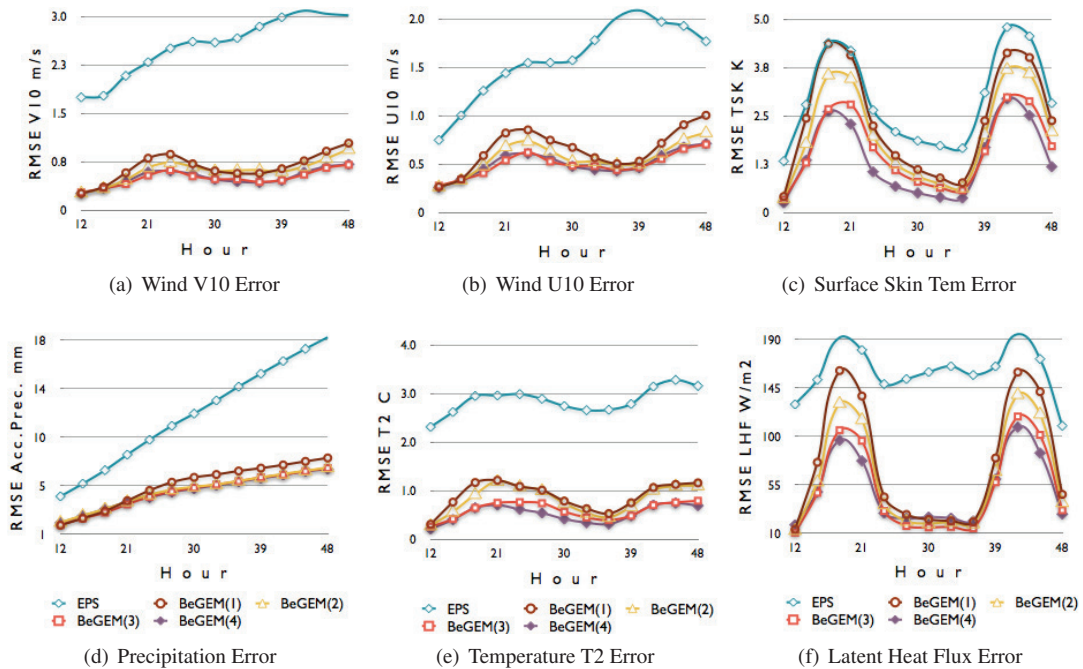


Figure 5: *RMS E* prediction error of: (a) 10-meter Wind Velocity component *V10*, (b) 10-meter Wind Velocity component *U10* (m/s), (c) Surface Skin Temperature, (d) Accumulated Precipitation, (e) 2-meter Temperature, and (f) Latent Heat Flux. Prediction using BeGEM (1, 2, 3, and 4) produced after 15 iterations of the Calibration Phase of the Multi-Variable *M-Level G-Ensemble*.

6. Parallel *M-Level G-Ensemble* Scheme

Our proposed scheme has shown a significant reduction in prediction error as described in the previous sections. 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, and this is also the case for WRF. See [17] for a study of scalability of WRF model execution over different HPC platforms and sizes.

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 contrast, *M-Level G-Ensemble* with its three phases, not only assures a significant reduction in prediction error, but also exhibits good execution times. It is important to note that, during the prediction phase, there is no need to execute all the initial ensemble members or the whole set of the calibrated ensemble members. We have shown in our experiments that by executing the BeGEM (a single calibrated forecast), prediction results are enhanced. However, this BeGEM is obtained after a Calibration phase.

As discussed previously, Calibration phase consists of GA iterative operations over a population of individuals. By implementing a Master/Worker parallel paradigm, as shown in Fig.(6), these individuals (very short forecasts having different parameter value combinations) are to be 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. At the last iteration, the parallel process exits with the selected individual, which we called the BeGEM that will be the selected set of parameters to conduct the Prediction phase.

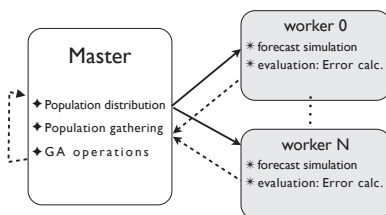


Figure 6: Master/Worker paradigm for Calibration phase.

To evaluate the parallel *M-Level G-Ensemble* scheme according to execution time and prediction enhancement, we show in Table (2) various scenarios of parallel *M-Level G-Ensemble* predictions with their respective execution time 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 using BeGEM(4) of 6 meteorological variables together in 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). Fig.(7) shows the respective prediction error of each scenario of those in Table (2).

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 2: Execution time Vs Scenario

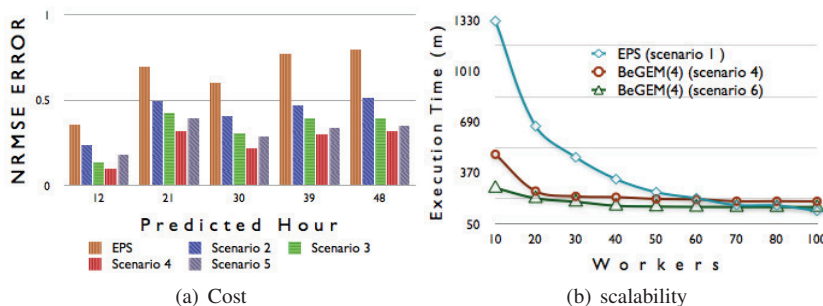


Figure 7: (a): Multi-Variable BeGEM(4); NRMSE of prediction of variables: Latent Heat Flux , Surface Skin Temperature , 2-meter Temperature, 10-meter Wind Velocity components U10 and V10, and the Accumulated Precipitation. (b): Scalability of EPS of 40 ensemble members over 100 computing noddos, BeGEM(4):(initial ens:40, iterations:15)

In all scenarios of *M-Level G-Ensemble*, we observed a significant reduction in execution time along with a corresponding reduction of prediction error. Additionally, as shown in Fig.(7(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 a 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.

7. Conclusions and future work

In this work, we have briefly described Numerical Weather Prediction models, along with a description of WRF as one of the most widely used models in the field. We highlighted the importance of the accuracy in NWP models, discussing also the two major sources of error in prediction.

We have introduced the parallel *M-Level G-Ensemble*, as a new scheme that enhances weather predictions. It uses an evolutionary algorithm (multi chromosome genetic algorithm) to improve the estimation of possible physical parameters that will provide more reliable predictions.

The parallel *M-Level G-Ensemble* prediction scheme showed a significant improvement in prediction quality. Moreover, in scenarios with a limited number of computing resources, the parallel *M-Level G-Ensemble* also constitutes a good solution that guarantees a significant enhancement in meteorological prediction and an overall reduction of execution time.

Thanks to the enhancement in prediction accuracy, more sophisticated schemes might be developed in the near future by injecting observed meteorological variables at run-time. These results encourage us to continue our research efforts by adding methods that handle real observations and deciding their injection intervals at run-time in order to get more reliable meteorological predictions.

References

- [1] National Weather Service (NWS) homepage.
URL <http://www.nws.noaa.gov/>
- [2] P. Lynch, The origins of computer weather prediction and climate modeling, *Journal of Computational Physics* 227 (7) (2008) 3431–3444.
- [3] Weather Research and Forecasting Model homepage.
URL <http://www.wrf-model.org/index.php>
- [4] PSU/NCAR MM5 community model homepage.
URL <http://www.mmm.ucar.edu/mm5/>
- [5] 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* 127 (572) (2001) 279–304.
- [6] M. Leutbecher, T. N. Palmer, Ensemble forecasting, *J. Comput. Phys.* 227 (2008) 3515–3539.
- [7] The ECMWF Ensemble Prediction System homepage.
URL <http://www.ecmwf.int/research/predictability/projects/index.html>
- [8] D. J. Stensrud, H. E. Brooks, J. Du, M. S. Tracton, E. Rogers, Using ensembles for short-range forecasting, *Monthly Weather Review* 127 (4) (1999) 433–446.
- [9] T. N. Palmer, P. D. Williams, Introduction. stochastic physics and climate modelling, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 366 (1875) (2008) 2419–2425.
- [10] J. Teixeira, C. A. Reynolds, Stochastic Nature of Physical Parameterizations in Ensemble Prediction: A Stochastic Convection Approach, *Monthly Weather Review* 136 (2) (2008) 483–496.
- [11] The Community Noah Land-Surface Model homepage.
URL <http://gcmd.nasa.gov/records/NOAA.NOAH.html>
- [12] J. Peng, Z. S. Chu, A hybrid multi-chromosome genetic algorithm for the cutting stock problem, *International Conference on Information Management, Innovation Management and Industrial Engineering* 1 (2010) 508–511.
- [13] R. Cavill, S. L. Smith, 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, GECCO '06*, ACM, New York, NY, USA, 2006, pp. 1405–1406.
- [14] C. J. Willmott, K. Matsuura, On the use of dimensioned measures of error to evaluate the performance of spatial interpolators, *International Journal of Geographical Information Science* 20 (1) (2006) 89–102.
- [15] Hurricane Katrina homepage.
URL <http://www.katrina.noaa.gov/>
- [16] H. Ihshaish, A. Cortes, M. A. Senar, Genetic ensemble (G-Ensemble) for meteorological prediction enhancement, in: H. R. Arabnia (Ed.), *Proceedings of The 2011 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2011)*, Vol. 1, 2011, pp. 404–4010.
- [17] Parallel WRF benchmarks homepage.
URL <http://www.mmm.ucar.edu/wrf/WG2/bench/>