



Research paper

Dynamically Optimized Unstructured Grid (DOUG) for Analog Ensemble of numerical weather predictions using evolutionary algorithms

Weiming Hu^{a,*}, Guido Cervone^{a,b}^a Department of Geography and Institute for CyberScience, Geoinformatics and Earth Observation Laboratory, The Pennsylvania State University, University Park, PA, United States of America^b Research Application Laboratory, National Center for Atmospheric Research, Boulder, CO, United States of America

ARTICLE INFO

Keywords:

Computational algorithms
Ensemble modeling
Evolutionary algorithms
Unstructured grid

ABSTRACT

The Analog Ensemble is a statistical technique that generates probabilistic forecasts using a current deterministic prediction, a set of historical predictions, and the associated observations. It generates ensemble forecasts by first identifying the most similar past predictions to the current one, and then summarizing the corresponding observations. This is a computationally efficient solution for ensemble modeling because it does not require multiple numerical weather prediction simulations, but a single model realization.

Despite this intrinsic computational efficiency, the required computation can grow very large because atmospheric models are routinely run with increasing resolutions. For example, the North American Mesoscale forecast system contains over 262 792 grid points to generate a 12 km prediction. The North American Mesoscale model generally uses a structured grid to represent the domain, despite the fact that certain physical changes occur non-uniformly across space and time. For example, temperature changes tend to occur more rapidly in mountains than plains.

An evolutionary algorithm is proposed to dynamically and automatically learn the optimal unstructured grid pattern. This iterative evolutionary algorithm is guided by Darwinian evolutionary rule generation and instantiation to identify grid vertices. Analog computations are performed only at vertices. Therefore, minimizing the number of vertices and identifying their locations are paramount to optimizing the available computational resources, minimizing queue time, and ultimately achieving better results. The optimal unstructured grid is then reused to guide the predictions for a variety of applications like temperature and wind speed.

1. Introduction

Numerical Weather Prediction (NWP) has been a computationally intensive problem. The main drivers for the increasing computational requirement include the growing complexity in NWP and the increased model resolution (both spatial and temporal). Models with higher complexity integrate more Earth system components, for example, the atmosphere and ocean waves, to generate more accurate predictions. Another driver are the refining spatial and temporal resolutions. Lorenz (1969) pointed out that, 'the errors in estimating the current state of the atmosphere are due mainly to omission rather than inaccuracy'. As a result, models are run with finer regular grids. The Global Forecast System has the potential to be run on a 4 km grid (Charba and Samplatsky, 2011), and the operational NAM runs on a 12 km grid. However,

regular grids only employ uniform discretization while atmospheric phenomena tend to be localized in both time and space (Skamarock and Klemp, 1993). When a finer regular grid is used to capture the dynamics of a weather event, required computational resources soar. Studies have been done to seek alternative grid types. For example, nested structured grids have demonstrated strength in simulating flows over aircraft and tracking shock waves (Bacon et al., 2000).

A third driver for increasing computation lies in the case of ensemble modeling where the model output comes from an ensemble of different realization of perturbed model initialization. The required computation multiplies with the number of members in an ensemble. Despite the cost, the ensemble output can be helpful in identifying

Abbreviations: AnEn, Analog Ensemble; DOUG, Dynamically Optimized Unstructured Grid; EA, Evolutionary Algorithm; FLT, Forecast Lead Time; GA, Genetic Algorithm; NAM, North American Mesoscale Forecast System; NWP, Numerical Weather Prediction; OMEGA, Operational Multiscale Environment Model with Grid Adaptivity; RMSE, Root Mean Square Error

* Correspondence to: The Pennsylvania State University, Department of Geography, 205 Walker Building, University Park, PA, 16802, United States of America.

E-mail address: weiming@psu.edu (W. Hu).

URL: <http://geoinf.psu.edu/> (W. Hu).

<https://doi.org/10.1016/j.cageo.2019.07.003>

Received 20 July 2018; Received in revised form 19 May 2019; Accepted 2 July 2019

Available online 1 August 2019

0098-3004/© 2019 Elsevier Ltd. All rights reserved.

the prediction uncertainty. Therefore this method has been gaining momentum in the past several decades.

In this paper, a DOUG is proposed to address the computational characteristic of NWP generations. The DOUG is a fully unstructured 2-dimensional grid driven by an Evolutionary Algorithm (EA). Together with the AnEn method, DOUG is designed to reduce the number of operations performed while maintaining, or even improving, the prediction accuracy.

Section 1.1 introduces grid refinement and grid adaptation techniques, and their applications in NWP. Section 1.2 reviews literature and studies on the EA. Section 2 describes the research region and data. Section 3 talks about the methodology. Section 4 shows the prediction results of the DOUG. Section 5 compares results from different simulation and discusses the pros and cons of the DOUG.

1.1. Grid adaptation and refinement

Although grid adaptation and refinement techniques have been widely discussed and adopted in Computational Fluid Dynamics problems (Jablonowski, 2004), probably (Skamarock et al., 1989) were the first to apply adaptive grid technology to atmospheric flow modeling in meteorological science.

Most of the time, grid adaptation and refinement are related to each other, thus appearing together (Dabberdt et al., 2004; Bacon et al., 2000; Jablonowski, 2004). Generally, grid refinement can be classified into two categories, the local refinement and the global refinement. The local refinement technique is to add grid nodes into original computational grids according to properly predefined criteria. Grid nodes from other areas may be removed to compensate for the nodes added. The total number of grid nodes might fluctuate within an accepted range. Tomlin et al. (1997) developed an adaptive gridding method for modeling chemical transport to study plume features, which generates denser meshes when it is closer to the smoke plume. The Adaptive Mesh Refinement, developed by Berger and Olinger (1984), is a classic example of local refinement technique by adding points to improve solution accuracy. Although local refinement might generate 'suboptimal' grid (Luchini et al., 2017) in terms of global accuracy, it is able to capture the dynamics of the event of interest. Another category is global refinement which redistributes a fixed number of total grid points. Global refinement is different from local refinement that the number of total grid points is strictly fixed and does not change during the refinement. Another major difference is the scope of optimization. Global refinement tends to look at the overall domain while evaluating a certain mesh solution while local refinement focuses on a smaller partition and decides whether to add or remove grid nodes from this specific partition. The status of the other partitions can be agnostic for the current partition because local refinement is designed for approximating detailed features. Global refinement, on the other hand, is designed for problems with a computational limit and an overall goal to improve the accuracy of representation.

There are several techniques for grid adaptation and refinement. Nested grid involves sequentially putting finer grids into the computational grid where a higher spatial resolution is needed (Mass and Kuo, 1998; Juang et al., 1997). Several limitations are present. First, it requires a prior knowledge of the interest, for example, a rough trajectory of cyclones in cases of hurricane tracking. Second, propagating dispersive waves between discontinuously changing grids might be problematic, and the accuracy and the reliability on a smoothly varying grid are higher compared with those on an abruptly varying grid (Gravel and Staniforth, 1992). The elastic grid solves the problem of abruptly varying grids because grid vertices are gradually moved to their vicinity based on a calculated weight. Grids only become larger or smaller depending on the motion of vertices. However, this technique still encounters the issue that a prior knowledge is needed. An improved version of the elastic grid is the Continuous Dynamic Grid Adaptation (Dietachmayer and Droegemeier, 1992; THOMPSON,

1984). It is able to automatically adapt the spatial distribution of grid points to the target of interest. However, it might require extra CPU time because of the continuously changing grid-transformation terms in the governing equations. Additional computation to discover the optimized grid is unavoidable. Therefore the trade-off point between computation and accuracy is very important.

An unstructured grid has also been used in grid adaptation and refinement. The first study of our knowledge to use unstructured grid technique for atmospheric simulation is the Operational Multiscale Environment Model with Grid Adaptivity (OMEGA) (Bacon et al., 2000). It is designed for real-time hazard prediction and allows models to be simulated on different scales without sudden changes in grid shapes by incorporating triangular grids. Hanna and Yang (2001) have compared the OMEGA with some other state-of-the-art models and found that the OMEGA has similar accuracy with the models on a regular grid. Another example of an unstructured grid is 'Atlas', a library developed and maintained by European Centre for Medium-Range Weather Forecasts (Deconinck et al., 2017). 'Atlas' provides the flexibility of coupling different Earth system components, such as the ocean and the atmosphere, by using a hybrid meshed. It is specifically designed with the goal exploiting the emerging hardware architectures and adopting the many-core architecture.

In summary, although grid adaptation and refinement techniques are not yet a standard solution in atmospheric studies, there has been important research exploring applications and importance of these techniques. However, questions on how hard it is to generate and use an unstructured grid and how to generalize criteria for optimizing unstructured grids still remain less discussed.

1.2. Evolutionary algorithm

EA uses a computational model that resembles the biological evolutionary process (Fraser, 1962) in the design. EA has been widely adopted in numerical and non-numerical optimization problems where a relatively large amount of local optima exists.

EA has three major components (Spears et al., 1993): (1) Evolutionary Programming; (2) Evolutionary Strategy; (3) GA. At a higher level, three components are similar to each other. For the brevity of this paper, attention is focused on GA.

GA generally consists of 3 steps, population initialization, selection, and reproduction (Montero et al., 2005). Take the maximum-value problem of a continuous univariate function $y = f(x)$ as an example. First, a population of numbers, also referred to as chromosomes, is generated either randomly or by prior knowledge. Second, an indicator for each chromosome is calculated from a predefined fitness function. The fitness function can be defined as the function f itself so that a number with higher fitness score will be considered as superior (Beasley et al., 1993) and would have a higher probability to survive. This process is termed 'selection' or 'ranking' (Beasley et al., 1993). Third, parents are selected from the population to 'mate' with each other to reproduce a new population of offspring. Reproduction is usually composed of crossover and mutation operators. Crossover takes two or more chromosomes as parents, and mixes the genes from each of them, producing one or more new chromosomes. Mutation is designed to simulate the concept of gene mutation in a DNA sequence. It is applied to a single chromosome by alternating its one or more genes with a predefined probability. They are different by features: crossover generates different chromosomes to explore new potential global optimums; mutation generates similar chromosomes to exploit local optimization. After the selection and reproduction, the process iterates with the updated population until convergence of the algorithm where the population produces adequately similar fitness values.

GA is not guaranteed to find the global optimum solution to a problem, but it is generally good at finding 'acceptably good' solutions 'acceptably quickly' (Beasley et al., 1993). Analytic techniques, such as differential equations, are usually faster than a GA if they exist. Therefore the main ground for GA is in difficult areas where such analytical approaches do not exist.

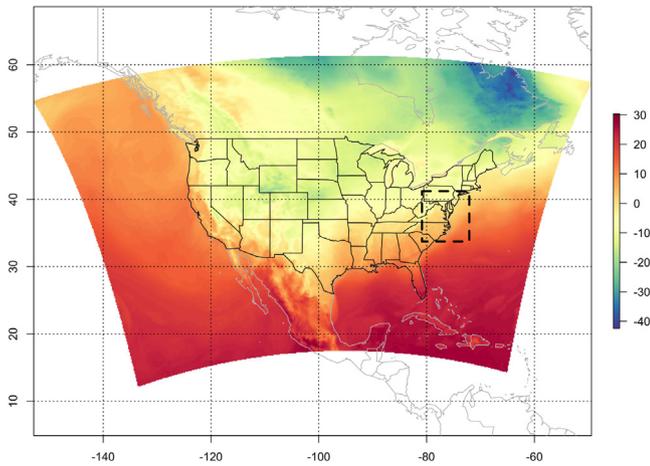


Fig. 1. A temperature map from NAM model forecast output for 3 AM on December 31, 2014, at 2 m above ground level. The black dashed rectangle shows the subset region with 4 424 grid points used in this work.

2. Data

NAM (Unidata, 2003) is one of the main weather models run by the National Centers for Environmental Prediction. It has four cycle runs per day starting at 00, 06, 12, 18 UTC, and for each cycle, it provides forecasts with a 3-hourly interval until 84 h into the future. It covers the North American continent at various vertical levels. In this work, the 12 km horizontal resolution is used. NAM forecasts from January 1, 2015 to March 5, 2016 were collected and then cleaned to remove days with missing values with 400 days left.

NAM analysis (Environmental Modeling Center, 2015) is used as an alternative 'ground truth'. Ground based observations are available but there are only 669 stations which does not meet the spatial resolution of NAM. The analysis of NAM would be helpful in this case. The analysis product assimilates more data than the original forecasts, for example observations from aircrafts and satellites. The analysis can be viewed as the theoretical 'best' that the model can perform. The analysis of the 12 km NAM have four cycles per day starting at 00, 06, 12, 18 UTC. NAM analysis from January 1, 2015 to March 5, 2016 were collected.

Fig. 1 shows the temperature forecasts for 3 AM on December 31, 2014, at 2 m above ground level from NAM. There are in total 262 792 regular grid points in the model output. The forecast data exceed 30 GB in size. This amount of data will be too large for developing and testing the work flow. Therefore, a subset region with 4 424 grid points is created from the east coast of the continental U.S.

Fig. 2 shows the temperature variability of the full and the subset regions for the test and training period. The dates are the 9th of each month from January, 2015, to February, 2016, to show the seasonality both in the full and the subset regions. Data on the 9th of each month are the most complete, so the date is chosen.

Although the subset region only has roughly 2% of the grid points in the full region, because a relatively complex topography, including water bodies, coastal areas, and mountains, is preserved in the subset region, the spatial variability of the subset region is comparable to the full data set. Therefore, this can be a sufficient representative case to develop and test the methodology.

The AnEn is a multi-variate method. However, it is found that using more variables does not necessarily result to a better prediction (Clemente-Harding et al., 2016). 6 variables are used in total and they are presented in Table 1. Horizontal wind speed and wind direction are not directly used in NAM, but computed using the u and v components of wind provided by NAM model.

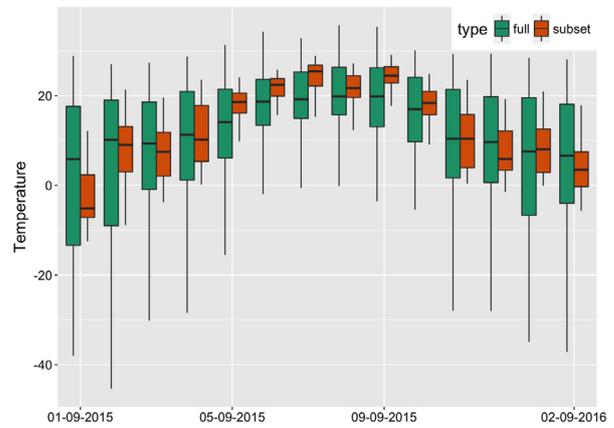


Fig. 2. A comparison between the temperature variability of the full and the subset regions from the data collected from January 2015 to February 2016. The dates are the 9th of each month.

Table 1

Weather variables used in the AnEn method.

Levels	10 m above ground	0 m above ground	100 000 Pa level
Variables	Temperature	Relative humidity	Total cloud cover
	Horizontal wind speed	Vertical velocity (pressure)	Wind direction

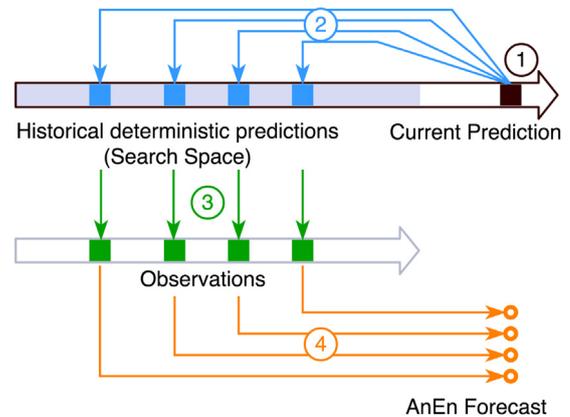


Fig. 3. The AnEn schema for generating a four-member ensemble forecast.

3. Methodology

The DOUG adaptation process is driven by a modified GA. The total number of grid points is fixed. Grid points will be automatically distributed to areas in need of more points. Then forecasts are generated at the selected grid points using the AnEn.

3.1. Analog ensemble

The AnEn method generates ensemble forecasts using a set of historical forecasts and the corresponding observations (Delle Monache et al., 2013, 2011). The data-driven technique takes the advantage of the massive amount of past available observations to avoid solving convoluted differential equations. As shown in Fig. 3, there are four steps to generate four-member AnEn forecasts.

1. Starts with a current multivariate prediction;
2. Searches through the historical multivariate predictions, and computes the similarity between current and past predictions;

3. Identifies predictions with the highest similarity;
4. Uses observations corresponding to the most similar historical predictions to generate the ensemble forecasts;

The similarity is computed using the following metric (Delle Monache et al., 2013):

$$\|F_t, A_{t'}\| = \sum_{i=1}^{N_v} \frac{\omega_i}{\sigma_{f_i}} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}, \quad (1)$$

where F_t is the current NWP deterministic forecast valid at the time stamp t and a specific location; $A_{t'}$ is the historical NWP deterministic forecast from the search space at the same location and the FLT, but for a different time t' ; N_v is the number of physical variables used in the comparison period; ω_i is the weight for each physical variable; σ_{f_i} is the standard deviation for the physical variable i calculated from historical forecasts for each FLT and location; \tilde{t} is equal to half of the window size of the FLTs to be compared; $F_{i,t+j}$ is the value of the current forecast for the physical variable i at the FLT $t + j$; $A_{i,t'+j}$ is the value of the historical forecast for the physical variable i at the FLT $t' + j$.

The method is able to generate ensemble forecasts from deterministic model without multiple simulation runs, which drastically reduces computation compared with the conventional ensemble models. Conventionally, multiple model simulations are carried out each with a slightly perturbed parameter initialization (Gneiting and Raftery, 2005) to test out the uncertainty. The AnEn method is flexible and scalable (Cervone et al., 2017) which takes advantage of a multi-core and multi-node infrastructure. It achieves 95% parallelization on a single node and has also been tested on the NCAR supercomputer Yellowstone.

3.2. Grid adaptation with an evolutionary algorithm

The DOUG adaptation process is driven by a modified GA. It is very important to properly design representation and reproduction. The spatial representation and reproduction procedures in the modified GA are introduced in the following sections.

3.2.1. Representation

The goal is to find the optimal unstructured grid given a fixed number of vertices. Once locations of vertices are identified, the corresponding unstructured grid will be the Voronoi graph determined by the vertices. Recall from Section 3.1 that predictions are generated only on vertices, and therefore there will be a resolution difference between the unstructured grid and the regular grid. To carry out grid-by-grid comparison, predictions on an unstructured grid need to be interpolated to the same extent and resolution of the regular grid. An unstructured grid becomes optimal when the final prediction from the unstructured grid has the lowest overall RMSE defined below.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2}, \quad (2)$$

where N is the total number of available grid points which ensures RMSE to be a grid-by-grid comparison between predictions and analysis; \hat{y}_i is the prediction for grid point i ; y_i is the analysis for grid point i .

The fitness function of the modified GA is defined as $f = -RMSE$ so that solutions with higher fitness measure are better than the others. The goal is to find a solution that maximizes the function f .

There are currently two ways to encode spatial features as genes and chromosomes. One is to use a raster-like representation (a 2-dimensional array). However, this could potentially be time-consuming (Li and Parrott, 2016). Another way is the vector representation which is adopted in this work. Each gene is a pair of numeric numbers that stands for the x-y coordinates. A chromosome is therefore a vector of coordinates. For example, for a particular unstructured grid with 5

vertices $G = \{P_i\}, i \in \{1, 2, 3, 4, 5\}$ and coordinates $P_i = (x_i, y_i)$, the corresponding chromosome is defined as $C = \{x_1, y_1, x_2, y_2, \dots, x_5, y_5\}$. Note that the unstructured grid is always defined by a set of vertices. The optimal unstructured grid is found when the set of vertices are optimal.

3.2.2. Reproduction

Reproduction can be divided into crossover and mutation. Crossover is a genetic operator used to switch parts of the genetic information from two chromosome. It is a way to stochastically generate new chromosome from the existing population, and therefore contribute to the capability of exploration. This operator is found to be particularly helpful in combinational optimization problems, like the Traveling Salesman Problem, and in problems with a binary representation. Crossover operators are found to be helpful in speeding up the convergence of population and finding the best solution faster.

The mutation operator operates on one chromosome at a time. Each chromosome has a probability to be selected. If a chromosome is selected, a proportion of the genes, also referred to as pairs of coordinates, will be mutated with small values generate from the Gaussian distribution. The distribution mean is the mean of coordinates from the chromosome. The distribution deviation is determined by the range of coordinates and a dampening factor. As the iteration number grows, the dampening factor will decrease, therefore leading to a smaller distribution spread. The spatial points will be moved with smaller distances as the algorithm goes on.

3.3. Work flow design

Fig. 4 represents the work flow with the AnEn and the modified GA. A population of random sets of spatial points are first selected. Each set can be viewed as a chromosome. Predictions are generated only on spatial points. A bias correction of the AnEn is carried out to reconstruct the spatial consistency (Sperati et al., 2017). The unstructured grid is then interpolated to a regular grid for grid-by-grid comparison with model analysis. Then tournament selection randomly picks two solutions and selects the better one until enough solutions have been selected. The selection process does not greedily select the absolute best subset of the population, which is referred to as the greedy selection, but makes sure that some of the weaker solutions can be selected to retain the population variability. The reproduction process is carried out as discussed in Section 3.2.2. This process will loop until the process reaches the maximum number of iterations.

This proposed workflow will be compared with random evolutionary algorithms with and without elitism. The random algorithms are simply lack of the tournament selection step which makes the evolution direction unclear to the algorithms. A random algorithm with elitism includes some level of evolutionary pressure but it is shown later in the results that it takes much longer to get to the same performance of the algorithm with tournament selection.

4. Results

The modified GA has been run for 500 iterations to identify the location of spatial points to optimize the overall temperature prediction for 6 AM on January 1st, 2016. The population size is set to 100 to keep a relatively large sample of chromosomes. Each chromosome has 200 genes, or spatial points. As discussed in Section 3.2.2, crossover probability is set to 0.7 and mutation probability is set to 0.5. During mutation, half of the genes on the chromosome will be mutated. Elitism strategy is used so that the top 5% solution will always survive after each iteration.

In Fig. 5, (a) shows the best solution of the population after 500 iterations. Figure (b) shows the evolution of fitness measures for each generation. Recall that fitness is defined as $-RMSE$. Therefore, a higher fitness value indicates a lower RMSE. The fitness value can

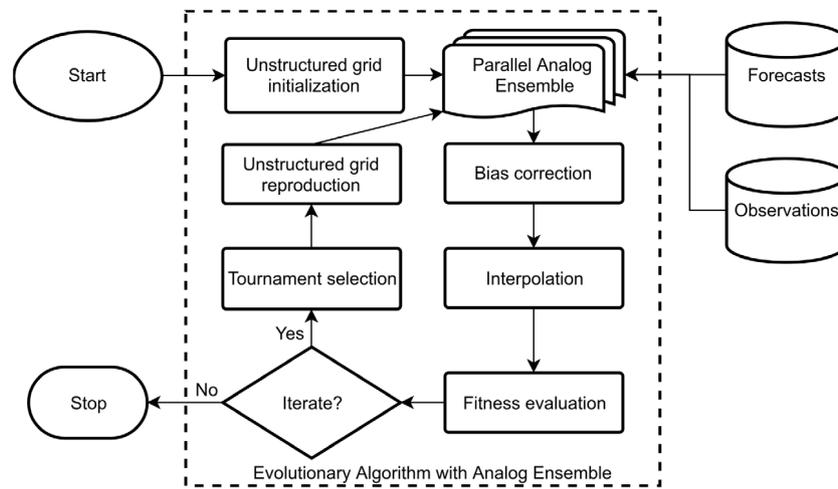


Fig. 4. The process for identifying the optimal subset of points to compute AnEn using a modified GA.

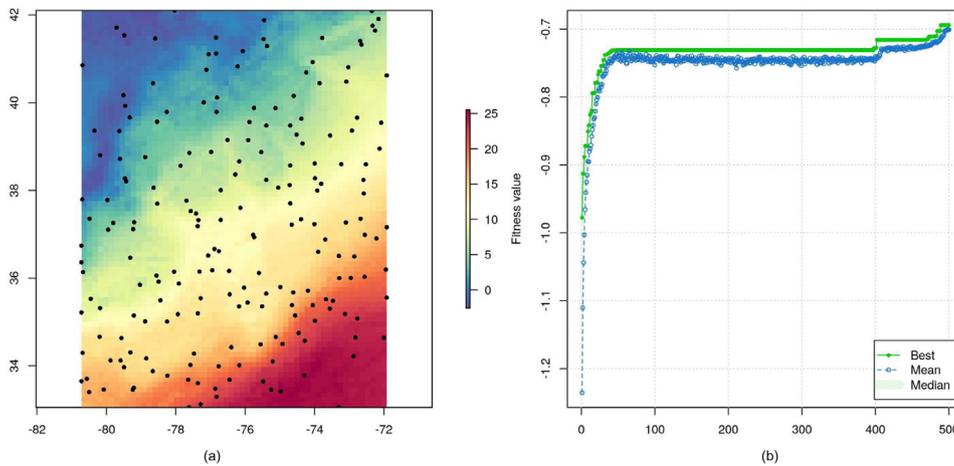


Fig. 5. (a) shows the best set of vertices from the population after 500 iterations. (b) shows the evolution of fitness measures for each generation.

never go across the 0 horizontal line because RMSE can never be negative. The accuracy is greatly improved during the first 50 iterations and then it reaches a plateau until about 400 iterations. The fast convergence before 400 iterations shows that the algorithm is able to find a good enough solution for unstructured mesh construction relatively fast. It is the result of both mutation and crossover operations. After 400 iterations, the algorithm shows a slight increase in forecast accuracy. But the algorithm is terminated at the 500th iteration in terms of the time consumption constraint. There is a 10-minute run time limit for the algorithm for finding the solution, and in this case, it can be translated to about 500 iterations. Another consideration for termination is that grid points from all solutions in the final population tend to have similar spatial patterns (not shown in the paper) which can be a sign of convergence.

Although the optimized unstructured grid is developed based on temperature prediction accuracy, it can be reused for other weather variables since weather events are inter-correlated. Fig. 6 shows comparisons between the forecasts on a regular grid and on the DOUG for six variables. Different variables might have different dynamics. For example, total cloud cover values are very different from temperature values, and therefore the optimal unstructured grid for temperature might not be optimal for the other variables. However, this limit can be partially offset in certain cases when the unstructured grid is reused for a period of time. Because these weather variables are related, one fast changing variable will also cause the other variables to change. More importantly, the benefit of reusing the grid lies in the amount of

computing power saved and the acceptable and sometimes improved prediction accuracy.

5. Discussion

The DOUG is compared with randomly generated unstructured grids. The first random method is termed the pure random adaptation. During each generation, random points are mutated and random solutions are kept to survive to the next iteration. Because random operations do not use the fitness selection, the pure random adaptation completely lacks a selection pressure to generate better solutions. The second random method is called random adaptation with elitism. This method is similar to the pure random adaptation except it uses elitism. The top one solution from each generation is kept, and the rest of the solutions are selected randomly. Elitism is the only source of selection pressure. Both methods have the same setting as the modified GA.

Table 2 compares the fitness values from the three methods respectively for the 500th generation. Fitness value is defined as $-RMSE$. Therefore a higher fitness value indicates better overall accuracy. The DOUG yields the highest fitness measure. The fitness mean and median of the DOUG both beat the two random methods. This indicates that after 500 iterations, the DOUG is better than the random grids. However, interestingly, the fitness range of the DOUG is the same with the pure random method, and random method with elitism yields a much smaller range. This is due to the random feature of these algorithms.

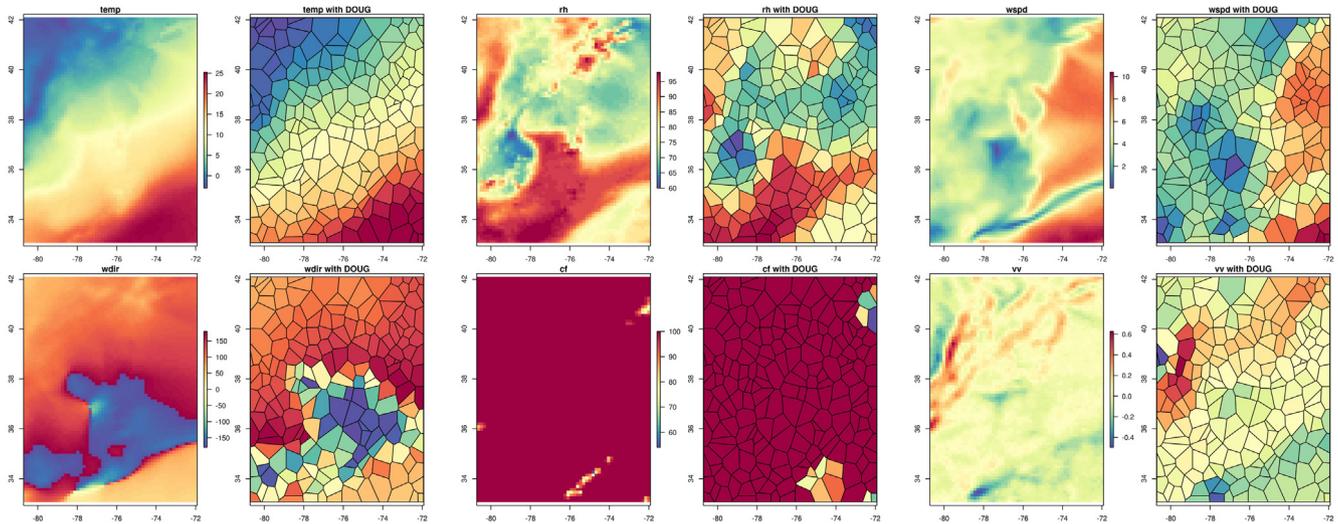


Fig. 6. The regular grid and the DOUG forecast comparison for different weather variables. Each variables are shown in regular grid on the left and in the DOUG on the right. The three variables on the first row are temperature, relative humidity at 2 m above ground, and wind speed at 10 m above ground; the three variables on the second row are wind direction at 10 m above ground, cloud fraction, and vertical velocity.

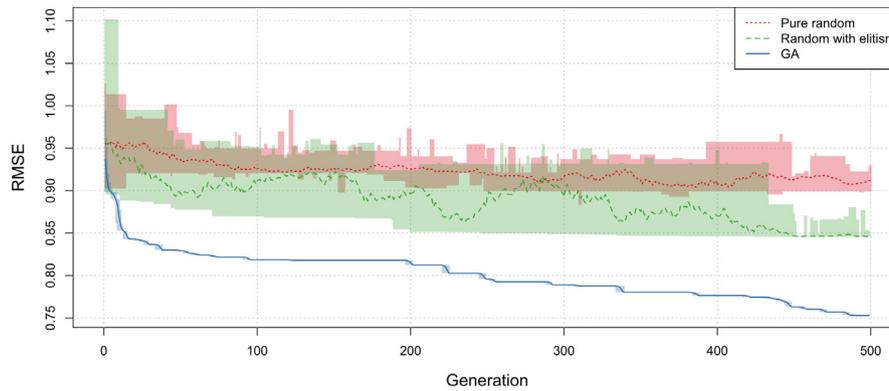


Fig. 7. The RMSE comparison between the random methods and the modified GA. Distribution shows the range of the RMSE of the population of a specific generation.

Table 2
Comparison of fitness values of the 500th generation from the DOUG population, the pure random population, and the random population with elitism.

	Number of vertices in each solution	Fitness value				
		Highest	Lowest	Range	Mean	Median
Pure random	200	-0.899	-0.93	0.031	-0.912	-0.917
Random with elitism	200	-0.846	-0.847	0.001	-0.846	-0.846
DOUG	200	-0.75	-0.78	0.031	-0.758	-0.753

Crossover is not used for the random method, so the range depends on mutation which can hardly be controlled.

Fig. 7 compares the RMSE for the 500 generations. The blue solid line indicates the RMSE changes of the modified GA algorithm. The distribution comes from the RMSEs of the population. As a comparison, the red dotted line indicates the pure random grid adaptation. The green dashed line indicates the random adaptation with elitism.

The random methods show the importance of selection pressure. The performance of the pure random method serves as a benchmark. With elitism, the algorithm performs better than the pure random adaptation, but the decrease of RMSE is not as fast as the modified GA algorithm. Additionally, the modified GA has the smallest distribution shows that there is a higher confidence that the optimized solution exists in the population.

Fig. 8(a) compares the RMSE from different methods for 2 months. Because unstructured meshes have different resolution from the baseline observational grid, the generated meshes are first interpolated and resampled to the same extent and resolution of baseline observations using inverse distance weighted interpolation. And then, RMSE is calculated based on a grid-to-grid comparison. The horizontal axis represents the index for test days starting from January 1st, 2016, to March 5th, 2016. The random grid and the DOUG are both generated for the first day on January 1st, 2016, and then reused for the rest of the test days. Analogs are also computed for each grid point, and the overall accuracy represents the theoretical best accuracy that the DOUG can reach, because the DOUG consists of fewer grid points in the region so that errors would potentially be introduced by interpolation.

For most of the days, the AnEn on a regular grid has the lowest RMSE as expected. When fewer grids are used, the random unstructured grid generates the worst result while the RMSE of the DOUG is very close to NAM model forecasts and the AnEn on a regular grid. This indicates the importance of the optimization of the unstructured grid when the number of vertices is limited. The RMSE difference can be as large as 2 degrees with and without optimization.

In Fig. 8(b) shows the error differences between each of the two unstructured grids and the theoretical best. A horizontal reference line is plotted at $RMSE = 0$. The green background indicates when the accuracy of DOUG is closer to the theoretical 'best' than the random grid. Note that on day 55, which is February 25, 2016, there are no analysis data available. This can happen but is not common. Out of

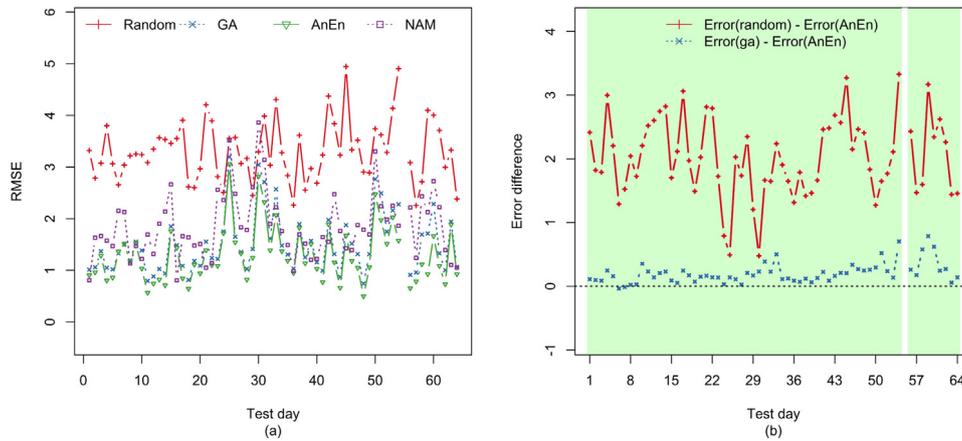


Fig. 8. (a) shows the RMSE for 2 months from different methods. 'Random' stands for reusing the pure random unstructured grid optimized for the first day; 'GA' stands for reusing the DOUG optimized for the first day; 'AnEn' stands for computing analogs for every regular grid point; and 'NAM' stands for the model forecasts on a regular grid. (b) shows the differences of RMSE over time for 2 months. A horizontal reference line is drawn and the green background indicates the error difference of the DOUG is closer to 0 than the that of the random grid.

400 days, there is only one day with missing data. All the other days show a green background indicating that using the DOUG is always better than the random grid. There are 2 test days, day 6 and 7, that the RMSE difference of the DOUG is negative. This suggests that the DOUG generates even better overall accuracy than the theoretical best RMSE. This demonstrates the benefit of the spatial abstraction. By optimizing the locations of spatial points, high error-prone areas can be avoided and therefore areas with higher predictability will have a larger probability to be selected. The benefit from avoiding high error-prone areas surpasses the errors brought by interpolation.

Another aspect to notice is that, in figure (b), the blue line shows an slightly upward trend after the 43rd test day while during the first 42 days the errors are closely restrained to the referential level. Reusing the unstructured grid optimized for the first day should cause the prediction error to accumulate because the grid is fixed while the weather dynamics is changing. There is a limit of how long a DOUG can be reused. In this case, the reuse cycle is 43 days. This limit can vary with a number of factors. First, predictands can affect the limit. Temperature is a gradually changing variable, unlike some other weather variables that can change very fast like total cloud cover and precipitation. Second, topology can affect the limit. In this case, the selected region is a rather complicated area with water, coasts, and mountains, which can contribute to a highly variable weather dynamics.

Fig. 9 shows the prediction errors from different methods. RMSEs are computed for each FLT. The distribution comes from the 64 test days from January 1st, 2016, to March 5th, 2016. Green box plots show the RMSEs of the AnEn method on each of the regular grids; orange box plots show the RMSEs of the original NAM model forecasts on a regular grid; purple box plots show the RMSEs of the AnEn method on the DOUG; and lastly red box plots show the RMSEs of the AnEn method on a randomly generated grid. Forecast lead times 06 h, 12 h, 18 h, 18 h, and 24 h UTC correspond to the local time in New York 02 h, 08 h, 14 h, and 20 h respectively. From forecast lead time 06 h to 18 h UTC, RMSEs of all methods, except the random method, increase because of the accumulation of model errors. However, at the third forecast lead time 18 h UTC, the DOUG generates significantly better results than the AnEn on a regular grid. This indicates that the DOUG can constrain the growth of model errors and achieve a lower RMSE when forecasting further into the future. At the fourth forecast lead time, RMSEs drops in all cases except for the random method because NAM model generally performs better during night time, and therefore there will not be a clear benefit in using grid abstraction technique and the DOUG.

To evaluate the relation between the computational performance and the prediction accuracy of both the random unstructured grid and

the DOUG, experiments have been carried out on a Dell desktop with 4 cores and 12 GB of memory. The processor model is Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz. With hyper-threading enabled, in total 8 threads can be used in parallel.

Random grids and the DOUG have been developed with different numbers of vertices ranging from 10 to 1010 with an interval of 50. Similarly, the final prediction on a unstructured grid is interpolated to a regular grid and compared with the analysis data to calculate the grid-by-grid RMSE. The process is repeated 10 times for statistical accuracy to account for stochastic effect. Fig. 10(a) shows the distribution of RMSE. Given the same number of vertices, the DOUG generates predictions with lower RMSE than the random grid. When the number of vertices increases in both grids, their RMSE values tend to converge and become closer to the theoretical best accuracy specified by the reference line. The reference RMSE comes from the AnEn on a regular grid. This is expected because when more vertices are used in the region, the additional errors introduced by interpolation are reduced. Ultimately, when the number of vertices reaches the number of the regular grid, the random grid and the DOUG will become the same as the regular grid because they are using every grid point available. Another way to interpret Fig. 10(a) is to look at the box plots horizontally. For example, to achieve an RMSE of roughly 0.8, the random grid requires 360 vertices but the DOUG requires 160 vertices saving the computation to generate forecasts on 200 vertices.

Fig. 10(b) shows the RMSE difference by subtracting that of the DOUG from the random grid. This figure shows how much better the accuracy can be when the number of vertices is fixed to a certain level. When fewer spatial points are used, the benefit of using a DOUG is not very clear. For example, when only 10 spatial points are used, no matter how much the algorithm optimizes their locations, the overall accuracy will not be very difference from locations that are randomly selected. When more points are used, strategy becomes more and more important. But at last, these two methods will resemble to each other because ultimately both methods will at most use each of the regular grid.

6. Conclusion

The research explored the ability of an unstructured grid to optimize computation while retaining or even improving the forecast accuracy. An unstructured grid can optimize the computation because it uses fewer grids to represent the region of interest and therefore fewer computational resources are required to generate forecasts on an unstructured grid than using a regular grid.

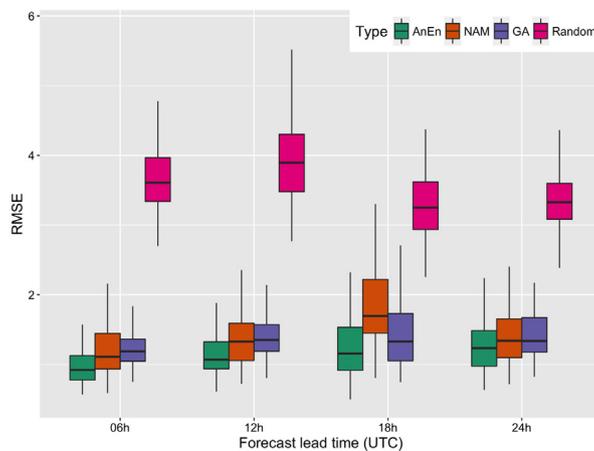


Fig. 9. The comparison of prediction errors across different FLT from different methods.

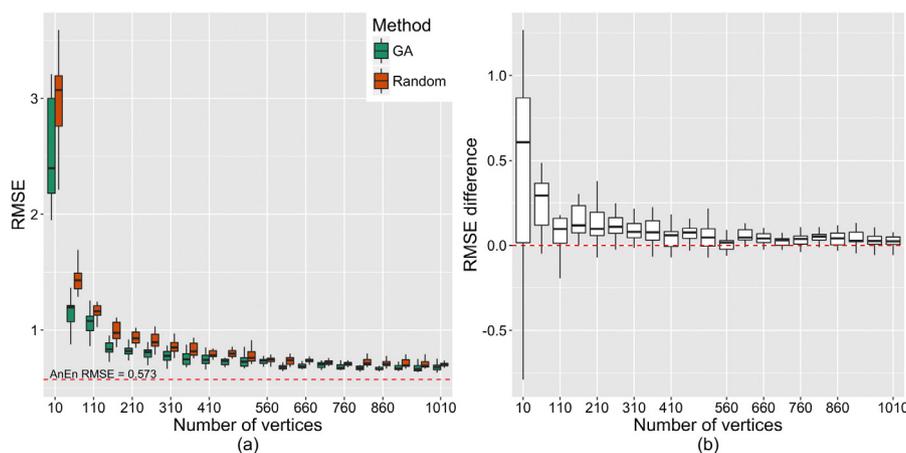


Fig. 10. (a) compares the RMSE for pure random grids and the DOUG given a different number of vertices in the grid. Experiments are repeated for 10 times to get the distribution. A horizontal reference line is shown with RMSE of the AnEn on a regular grid. (b) shows the RMSE difference between the DOUG and the random grid. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The proposed DOUG driven by a modified GA is found to beat random adaptation. In certain cases, the overall accuracy is very close to or even better than using a regular grid. This method can automatically identify areas in need of more vertices and allocate computational resources accordingly. When reusing the DOUG, errors are constrained in an acceptable range for the first 40 days in the case of temperature prediction. The limit of how long the optimal grid can be reused still needs to be thoroughly studied.

The DOUG shows promising performance in short-term temperature forecast carried out in this study. It provides an innovative solution to numerical weather forecast with a computational limit by trading off the overall accuracy and the computational performance. It is suggested to investigate the re-usability limit of an optimized grid for different weather variables and improve the performance of the GA in any future work.

Software availability

Name: Dynamically Optimized Unstructured Grid Algorithm
 Repository: <https://github.com/Weiming-Hu/DOUG>
 Developers: Weiming Hu, Guido Cervone
 Contact Address: Walker building, Room 205, University Park, 16802
 Telephone: 814-863-0179
 E-mail: weiming@psu.edu
 Year First Available: 2018

Program Language: R, C++
 Code Access: The complete codes are provided along with the submission of this manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Weiming Hu: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Resources, Software, Validation, Writing. **Guido Cervone:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing.

Acknowledgments

This work was supported by the National Science Foundation [grant numbers #1639707].
 We thank Laura Clemente-Harding and Martina Calovi for providing comments that greatly improved the manuscript. We also attribute special credits to R packages, ‘RColorBrewer’ (Neuwirth and Brewer, 2005), ‘raster’ (Hijmans and Van Etten, 2014) and ‘GA’ (Scrucca et al., 2013).

References

- Bacon, D.P., Ahmad, N.N., Boybeyi, Z., Dunn, T.J., Hall, M.S., Lee, P.C.S., Sarma, R.A., Turner, M.D., Waight, K.T., Young, S.H., Zack, J.W., 2000. A dynamically adapting weather and dispersion model: The operational multiscale environment model with grid adaptivity (OMEGA). *Mon. Weather Rev.* 128 (7), 2044–2076.
- Beasley, D., David, B., Ralph, M., 1993. An Overview of Genetic Algorithms: Part 1, Fundamentals, Technical Report 2.
- Berger, M.J., Olinger, J., 1984. Adaptive mesh refinement for hyperbolic partial differential equations. *J. Comput. Phys.* 53, 484–512.
- Cervone, G., Clemente-Harding, L., Alessandrini, S., Delle Monache, L., 2017. Short-term photovoltaic power forecasting using Artificial Neural Networks and an Analog Ensemble. *Renew. Energy* 108, 274–286.
- Charba, J.P., Samplatsky, F.G., 2011. High-resolution GFS-based MOS quantitative precipitation forecasts on a 4-km grid. *Mon. Weather Rev.* 139 (1), 39–68.
- Clemente-Harding, L., Cervone, G., Delle Monache, L., Haupt, S.E., Alessandrini, S., 2016. Analog ensemble (AnEn): Optimal predictor weighting and exploitation of spatial characteristics in AnEn generation. In: AGU Fall Meeting Abstracts. pp. A41G-0131.
- Dabberdt, W.F., Carroll, M.A., Baumgardner, D., Carmichael, G., Cohen, R., Dye, T., Ellis, J., Grell, G., Grimmond, S., Hanna, S., Irwin, J., Lamb, B., Madronich, S., McQueen, J., Meagher, J., Odman, T., Pleim, J., Schmid, H.P., Westphal, D.L., 2004. Meteorological research needs for improved air quality forecasting. *Bull. Am. Meteorol. Soc.* 85, 563–586.
- Deconinck, W., Bauer, P., Diamantakis, M., Hamrud, M., Kühnlein, C., Maciel, P., Mengaldo, G., Quintino, T., Raoult, B., Smolarkiewicz, P.K., Wedi, N.P., 2017. Atlas : A library for numerical weather prediction and climate modelling. *Comput. Phys. Comm.* 220, 188–204.
- Delle Monache, L., Eckel, F.A., Rife, D.L., Nagarajan, B., Searight, K., 2013. Probabilistic weather prediction with an analog ensemble. *Mon. Weather Rev.* 141, 3498–3516.
- Delle Monache, L., Nipen, T., Liu, Y., Roux, G., Stull, R., 2011. Kalman filter and analog schemes to postprocess numerical weather predictions. *Mon. Weather Rev.* 139, 3554–3570.
- Dietachmayer, G.S., Droegemeier, K.K., 1992. Application of Continuous Dynamic Grid Adaptation Techniques to Meteorological Modeling. Part I: Basic Formulation and Accuracy. [http://dx.doi.org/10.1175/1520-0493\(1992\)120<1675:AOCDDGA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1992)120<1675:AOCDDGA>2.0.CO;2).
2015. N. W. S. N. U. D. o. C. Environmental Modeling Center, National Centers for Environmental Prediction, Ncep north american mesoscale (nam) 12 km analysis.
- Fraser, A.S., 1962. Simulation of genetic systems. *J. Theoret. Biol.* 2 (3), 329–346.
- Gneiting, T., Raftery, A.E., 2005. Weather forecasting with ensemble methods. *Science* 310, 248–249.
- Gravel, S., Staniforth, A., 1992. Variable resolution and robustness. *Mon. Weather Rev.* 120.
- Hanna, S.R., Yang, R., 2001. Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths. *J. Appl. Meteorol.* 40, 1095–1104.
- Hijmans, R.J., Van Etten, J., 2014. raster: Geographic data analysis and modeling. r package version 2.2-31, Google Scholar.
- Jablonowski, C., 2004. Adaptive Grids in Weather and Climate Modeling (Ph.D. Thesis). p. 272.
- Juang, H.-M.H., Hong, S.-Y., Kanamitsu, M., 1997. The ncep regional spectral model: An update. *Bull. Am. Meteorol. Soc.* 78, 2125–2144.
- Li, X., Parrott, L., 2016. An improved genetic algorithm for spatial optimization of multi-objective and multi-site land use allocation. *Comput. Environ. Urban Syst.* 59, 184–194.
- Lorenz, E.N., 1969. Atmospheric Predictability as Revealed by Naturally Occurring Analogues. [http://dx.doi.org/10.1175/1520-0469\(1969\)26<636:APARBN>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2).
- Luchini, P., Giannetti, F., Citro, V., 2017. Erratum to: Error sensitivity to refinement: a criterion for optimal grid adaptation. *Theor. Comput. Fluid Dyn.* 31 (5-6), 595–605. <http://dx.doi.org/10.1007/s00162-016-0413-x>; 2017. *Theor. Comput. Fluid Dyn.* 31, 665.
- Mass, C.F., Kuo, Y.-H., 1998. Regional real-time numerical weather prediction: Current status and future potential. *Bull. Am. Meteorol. Soc.* 79 (2), 253–264.
- Montero, G., Rodríguez, E., Montenegro, R., Escobar, J.M., González-Yuste, J.M., 2005. Genetic algorithms for an improved parameter estimation with local refinement of tetrahedral meshes in a wind model. In: *Advances in Engineering Software*, Vol. 36, pp. 3–10. <http://dx.doi.org/10.1016/j.advengsoft.2004.03.011>.
- Neuwirth, E., Brewer, R., 2005. ColorBrewer palettes.
- Scrucca, L., et al., 2013. Ga: a package for genetic algorithms in r. *J. Stat. Softw.* 53, 1–37.
- Skamarock, W.C., Klemp, J.B., 1993. Adaptive Grid Refinement for Two-Dimensional and Three-Dimensional Nonhydrostatic Atmospheric Flow. [http://dx.doi.org/10.1175/1520-0493\(1993\)121<0788:AGRFTD>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1993)121<0788:AGRFTD>2.0.CO;2).
- Skamarock, W., Olinger, J., Street, R., 1989. Adaptive grid refinement for numerical weather prediction. *J. Comput. Phys.* 80, 27–60.
- Spears, W.M., {De Jong}, K.A., Bäck, T., Fogel, D.B., {de Garis}, H., 1993. An overview of evolutionary computation. In: *Proceedings of the European Conference on Machine Learning - ECML 1993*, Vol. 667. Vienna, Austria, pp. 442–459. http://dx.doi.org/10.1007/3-540-56602-3_163.
- Sperati, S., Alessandrini, S., Delle Monache, L., 2017. Gridded probabilistic weather forecasts with an analog ensemble. *Q. J. R. Meteorol. Soc.* 143 (708), 2874–2885.
- THOMPSON, J.F., 1984. Grid generation techniques in computational fluid dynamics. *AIAA J.* 22 (11), 1505–1523.
- Tomlin, A., Berzins, M., Ware, J., Smith, J., Pilling, M.J., 1997. On the use of adaptive gridding methods for modelling chemical transport from multi-scale sources. *Atmos. Environ.* 31, 2945–2959.
2003. U. C. f. A. R. Unidata, N. U. D. o. C. National Centers for Environmental Prediction, National Weather Service, E. C. for Medium-Range Weather Forecasts, Historical unidata internet data distribution (idd) gridded model data.