

Numerical model for atmospheric temperature prediction in the Pamplonita River Basin, Norte de Santander, Colombia

Jordi Rafael Palacios Gonzalez¹, Francisco Henry Cabrera Zambrano², Jacipt Alexander Ramon Valencia³, Alejandro Guillen Betancourt⁴

¹ Environmental Engineering, Universidad Sergio Arboleda-Colombia

² Geology, Universidad de Pamplona-Colombia

³ Environmental Engineering, Universidad de Pamplona-Colombia

⁴ System Engineering, Universidad de Pamplona-Colombia

ABSTRACT

The work consists of the development of a computational model for the simulation of the temperatures, using a system of differential equations in partial derivatives, which solution was found by finite differences method, using the hypothesis of the isobaric coordinates, it means that the pressure is not a function but becomes an independent variable, in the new system, z is a function, more precisely a variable closely related to the geopotential ϕ . The data was obtained from the NCAR's Research Data Archive Database in January 2011 at different altitudes of 2 to 4572 meters above sea level (m.a.s.l.). The grid was structured with 10 spatial points on the x-axis, 8 spatial points on the y axis, and 5 sigma's layers at the same altitude evaluated, representing the 80 km² covering the Pamplonita River basin in the sector between Pamplona and Cucuta. The calculations obtained by the modeling system at the time of validation of the model have a low error level. The minimum error of validating the model was 0.17 %, obtained for the results of the model at 2 m.a.s.l.

Keywords: Atmosphere, Atmospheric dynamic, Non-linear, Computer model, Meteorology

Corresponding Author:

Jordi Rafael Palacios Gonzalez,
Environmental Engineering,
Universidad Sergio Arboleda-Colombia
Address. Street 74 # 14-14
E-mail: jordi.palacios@usa.edu.co

1. Introduction

The study of atmospheric circulation generates a growing interest in researchers in different areas of earth sciences; as the atmosphere is a complex dynamic system, its prediction represents a complex problem. Therefore, having an effective prediction of the meteorological variables is almost impossible. [1] analyzed a series of systems that were representative of meteorology, based on three non-linear differential equations, observing in his results that if he did a minor numerical modification to the initial conditions of a variable, the result of the simulation would be totally different from the previous one. The meteorological prediction has increased the quality of its forecasts as time has passed; those done in 1954 with the barotropic model for 24 hours are equivalent to the ones done in 1995 in the European Center for Meteorological Predictions for a 6-days. Without going so far back in time, the statistics of the last years about Europe, published by Météo-France, show that the quality of the prediction in 1995 for 72 hours is equivalent to the one had in 1980 for 24 hours forecast [2], [3].

A complete understanding of the general circulation requires an understanding of the role of small-scale motions, radiation, convection and interaction with the ocean and land surface.[4] Weather prediction is a challenging task for researchers and has drawn of many research interests in recent years.

Interpolation methods are widely used for temperature prediction like as [5] that presents a procedure to interpolate daily mean temperature over a whole year period by using time series of auxiliary predictors, where the main objective was to promote space-time prediction techniques for functional mapping versus purely spatial methods. However, this type of prediction is based on the analysis of point values taken from satellite data and the use of statistical interpolation methods where only the temperature at surface level is taken into account.

Literature studies have shown that machine learning techniques achieved better performance than traditional statistical methods. [6] applied the Support Vector Machines (SVMs) for weather prediction. In this work time series, daily maximum temperature data at a location is analyzed to predict the maximum temperature of the next day at that location based on the daily maximum temperatures for a span of previous n days. Performance of the system is observed over various spans of 2 to 10 days by using optimal values of the kernel function. The non-linear regression method is found to be suitable to train the SVM for this application. Other authors, as [7] studied the effectiveness of multilayer perceptron networks (MLPs) for the prediction of the maximum and the minimum temperatures based on past observations on various atmospheric parameters to capture the seasonality of atmospheric data, to improve the prediction accuracy. However, in these types of applications, the physics of the atmosphere and spatial-temporal interaction of temperatures with other meteorological variables are not taken into account.

A completely non-hydrostatic model system known as the Advanced Regional Prediction System (ARPS) was developed in the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma [8], [9]. The ARPS is an effective tool for research and a system suitable for the explicit prediction of convective storms and weather systems at significant scales. The ARPS includes its data ingest, quality control and objective analysis packages, a data assimilation system that includes single-Doppler velocity and thermodynamic retrieval algorithms, the forward prediction component, and self-contained post-processing, diagnostic, and verification package [10]

Other types of physics-based models, such as oceanic-atmospheric coupling [11], [12], use a fully non-linear ocean general circulation and atmospheric physics and statical element. While these models there have been significant advances in understanding the role of the wave, mean-flow interaction in the general circulation of the atmosphere and its variability, in particular, the relationships among wave propagation, momentum transport and zonal flow accelerations are much better understood and have been applied to a wide range of problems.

Currently, there are different models of meteorological prediction at the global level; for example, the Weather Research and Forecasting model (WRF) [13], that has become one of the world's most widely used numerical weather prediction models [14] and that constitutes a novel meteorological non-hydrostatic tool, developed from the collaboration of prestigious international research centers [15], the HadCM3, an integrated climate model which has been widely used for weather forecast, detection and attribution, and other studies of climate sensitivity [16], and the MM5 model, a terrain follow-up model designed to simulate or forecast the mesoscale and regional atmospheric circulation in a grand scale [17]. These models are robust and have been developed by experts in the different fields that make up atmospheric sciences, with high quality results, although at present it is necessary to develop models at local scales that can be adapted to the specific needs of each area, especially in tropical climates that present a high variation of their atmospheric systems over small distances.

This work aims to develop a numerical atmosphere temperature prediction model based on the equations of atmospheric circulation in isobaric coordinates, implementing four assumptions. In order to make a first approach to the construction of a meteorological model for the Pamplona river basin, that can be implemented for the generation of early warning systems for extreme weather events.

2. Study zone

The Pamplona river basin (Figure 1) is located southwest of the department of Norte de Santander, on the eastern slope of the eastern mountain range of Colombia, with heights oscillating between 50 and 4200 meters above sea level (m.a.s.l). The main channel forms in the town of Pamplona at the confluence of the El Rosal and Navarro streams and ends near the town center of Puerto Villamizar in the Municipality of Cúcuta where present high climate variety, with conditions varying wildly in small distances [18], [19]. In the basins of the

rivers Pamplonita there are two different climate patterns, which have a differentiated behavior influenced by the ENSO phenomena [20]. The rainfall regime tends to be bimodal in most of the basin, there are two dry seasons roughly equivalent in intensity, from January to March and from June to August. The study area (Figure 1) covers the upper and lower part of the basin from the town of Pamplona to Cúcuta city where present temperatures that oscillate between 9°C and 28°C on average.

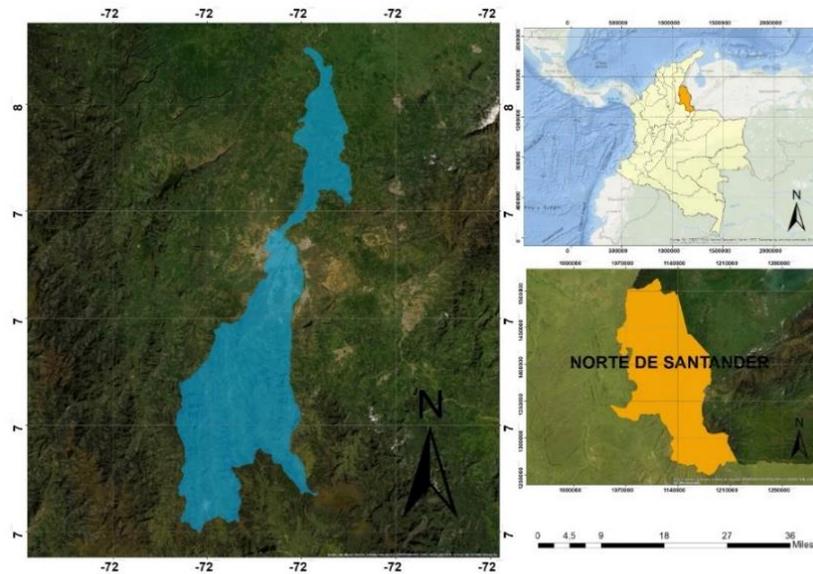


Figure 1. Study Area, Norte de Santander Department, Colombia.

3. Methodology and construction of the model

To develop the temperature prediction model, the methodology was the following:

Beginning from the primitive equations of the atmospheric circulation in isobaric coordinates (the height variation is used in the function of the pressure surfaces, and not in the function of geometrical height) and the four assumptions (described later), the equation system (1-5) which was representative of the Pamplonita River basin was established, which is ruled by the physical and thermodynamic laws of fluid mechanics, given that the atmosphere is a fluid, with properties that vary in time and space [21].

$$\frac{\partial u}{\partial t} = -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} - \omega \frac{\partial u}{\partial P} - \frac{\partial \phi}{\partial x} + fv \quad (1)$$

$$\frac{\partial v}{\partial t} = -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - \omega \frac{\partial v}{\partial P} - \frac{\partial \phi}{\partial y} - fu \quad (2)$$

$$\frac{\partial T}{\partial t} = -u \frac{\partial T}{\partial x} - v \frac{\partial T}{\partial y} - \omega \frac{\partial T}{\partial P} + \frac{1}{Cp\rho} - \omega \quad (3)$$

$$\frac{\partial T}{\partial t} = -u \frac{\partial T}{\partial x} - v \frac{\partial T}{\partial y} - \omega \frac{\partial T}{\partial P} + \frac{1}{Cp\rho} - \omega \quad (4)$$

$$\frac{\partial \omega}{\partial P} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \quad (5)$$

$$\frac{\partial \phi}{\partial P} = -\frac{1}{\rho}$$

Where ω is the temporal variation of pressure, ϕ is the geopotential height, u and v are the wind components (u is the zonal component, specifically the horizontal speed along a latitude circle in west to east direction, v is the horizontal speed along a south to north meridian), T is the temperature of the atmosphere, C_p is the moist air constant, f is Coriolis' force, and ρ is air density.

The criteria to select variables were established from the differential equations which represent the model and after the assumptions made to simplify the solution, the initial validation conditions and model calibration were obtained from the temporal series acquired in the Research Data Archive Computational & Information System Lab, from the database NCEP Global Forecast System (GFS) the data for the years 2001 to 2012 [22], with a complete validation, selecting the variables: Temperature (T) of the air at different heights, the geopotential (ϕ), pressure (P), direction and speed of the wind (u, v).

The assumptions established for the model applied in the Pamplonita river basin are:

Assumption 1: The representation of the model was done in three (3) spatial dimensions, from which the z component is replaced by the component of spatial pressure and temporal pressure, so our independent variables are x, y, P , and t .

Assumption 2: The study area was represented as a cube, in which the Pamplonita River basin represents its grid.

Assumption 3: This is considered a closed system, which behaves as an adiabatic one (there is no heat transfer).

Assumption 4: It is assumed that what is above the ground corresponds completely to the atmosphere, given that the topography of the study basin was not considered.

A grid of $8 \times 10 \times 5$ was done to represent the Pamplonita River basin zone, in which the longitude is given by 8 nodes and the latitude by 10 nodes (which cover an area of approximately 80 Km^2). At the same time, the altitude of 5 levels comprehends from 2, 1829, 2743, 3658 and 4572 (m. a. s. l.), where the initial conditions are given by the dependent variables (u, v, T, ω, ϕ) obtained from the satellite data and the solving of the constants.

The approximate solution of the model uses the finite difference method, generating a future prediction for 1, 4 and 10 days. The boundary conditions were defined by temperature (t), wind speed (v, w) and pressure (p) at the edge points within the domain, which was set around the mesh $10 \times 8 \times 5$ and the values of variables t, w, v and p in the time 0 (t_0) are the initial conditions.

For the validation of the model, its results are compared with the actual data of 1, 4 and 10 days, respectively. For this purpose, we used the Percent Error Eq. (6), defined as:

$$PE = \frac{T_j - T_i}{T_i} \quad (6)$$

Where T_j is the modelled value; T_i is the real value.

4. Results

Below are the present results obtained from the simulations done by the model for the Pamplonita River basin; the interest variable corresponds to the temperature at different heights.

In Figure 2 is presents the comparison of the real data and the average calculated data for a day at 2 m. a. s. l. high, where the vertical axis represents the temperatures in degrees Celsius for a day, with a maximum value for the calculated data of 25.6°C and the minimum data was 9°C , the maximum value for the real data was 27°C , and the minimum data was 10°C .

In this simulation, it can be observed that the results of the prediction of temperatures have a high similarity with the real data, except in longitude 6 where the real data have a lower value than the simulated. The highest temperature ranges are present at this level; since they are found on the surface and also because the study area is on a tropical system with important changes in temperatures over small distances [23].

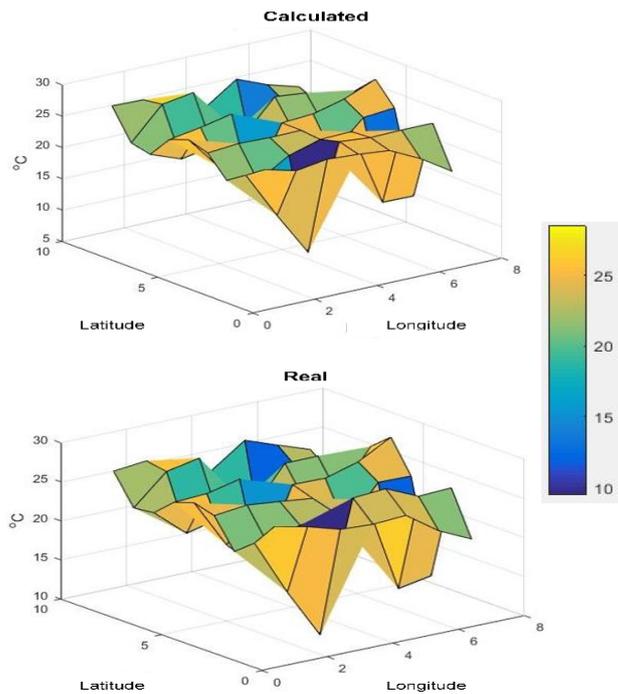


Figure 2. Temperatures (°C) calculated from the model vs. real for a day at 2 m. a. s. l.

In Figure 3 is presents the comparison of the real data and the average calculated data for a day at 1829 m. a. s. l. where the vertical axis represents the temperatures in degrees Celsius for a day. The maximum value for the calculated data was 15.7°C and the minimum data was 14.96°C, for the real data the maximum was 15.7°C and the minimum data was 14.9°C. In this simulation, it can be observed that the results of the prediction of temperatures have a similarity with the real data, except in longitude 3 to 8 where the model overestimates the real data.

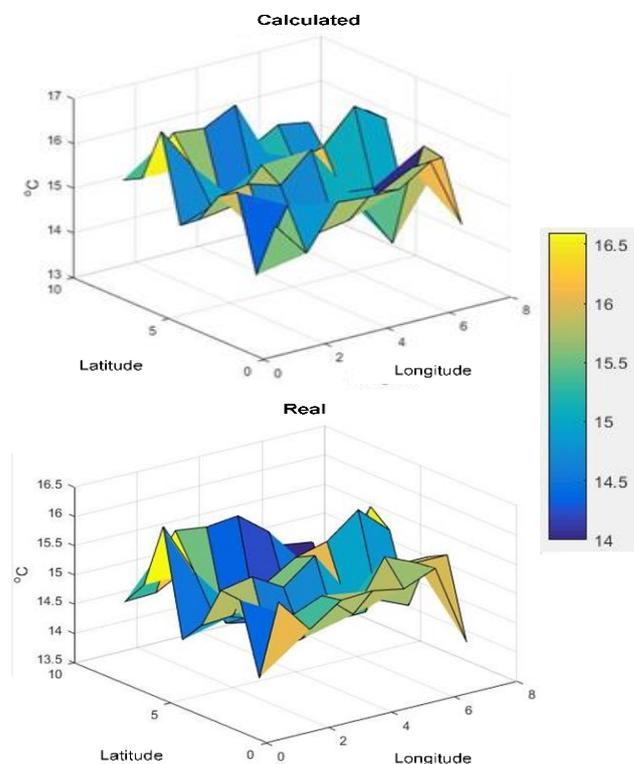


Figure 3. Temperatures (°C) calculated from the model vs. real for a day at 1829 m. a. s. l.

In Figure 4 is presents the comparison of the real data and the average calculated data for a day at 2473 m. a. s. l. where the vertical axis represents the temperatures in degrees Celsius for a day. The maximum value for the calculated data was 10.96°C, and the minimum data was 9.98°C, for the real data, the maximum was 10.2°C, and the minimum data was 9.8°C. In this simulation, it can be observed that the results of the prediction of temperatures have a similarity with the real dates, except in longitude 3 to 8 where the model overestimates the real data.

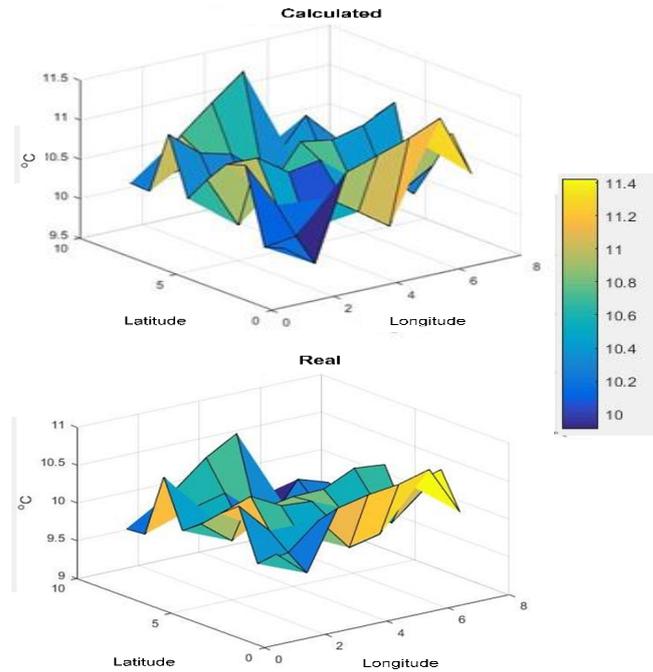


Figure 4. Temperatures (°C) calculated from the model vs. real for a day at 2743 m. a. s. l.

In Figure 5 is presents the comparison of the real data and the average calculated data for a day at 3658 m.a.s.l. where the vertical axis represents the temperatures in degrees Celsius for a day. The maximum value for the calculated data was 6.7°C and the minimum data was 5.54°C, for the real data the maximum data was 5.98°C, and the minimum data was 5.5°C. In this simulation, it can be observed that the results of the prediction of temperatures overestimate the real data.

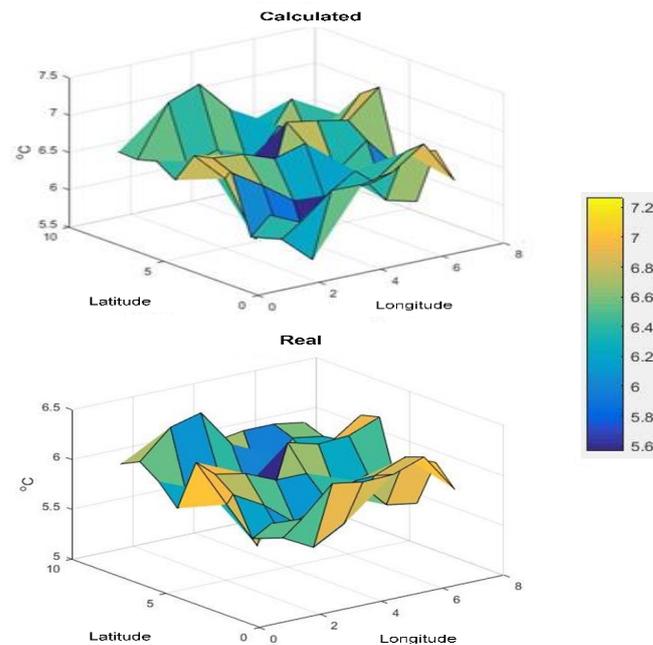


Figure 5. Temperatures (°C) calculated from the model vs. real for a day at 3658 m. a. s. l.

In Figure 6 the comparison of the real data and the calculated average temperature data for a day at 4572 m.a.s.l is observed, where the vertical axis represents the temperatures in degrees Celsius. The maximum value for the calculated data was 2.1°C and the minimum was 1.39°C and for the real data the maximum was 1.51°C and the minimum was 1.1°C. In this simulation, it can be observed that the results of the prediction of temperatures overestimate the real data. The temperature range is much lower because this is the highest level, and its behavior tends to be more uniform [24].

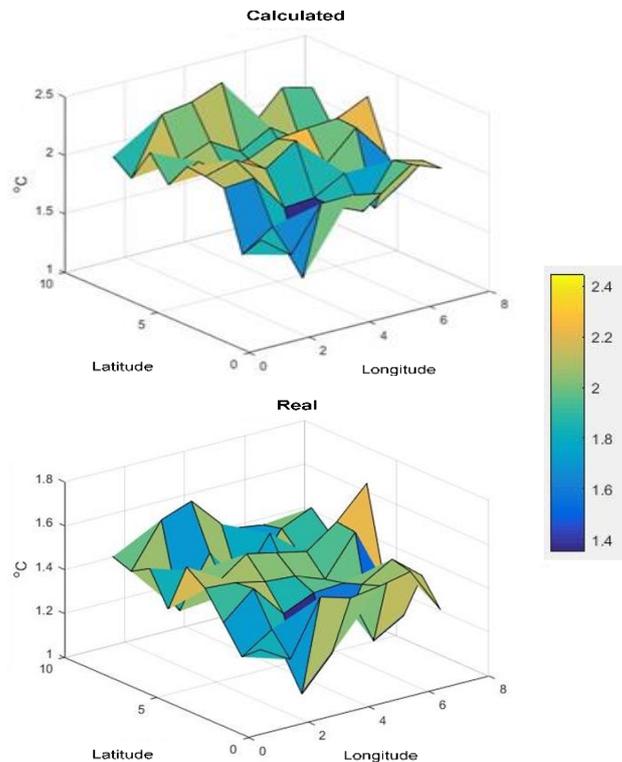


Figure 6. Temperatures (°C) calculated from the model vs. real for a day at 4572 m. a. s. l.

4.1. Validation of the model for the Pamplonita River basin.

In Figure 7 is observed the comparison of the data obtained from the meteorological prediction model and the real data of the database at 2, 1829, 2743, 3685 and 4572 m. a. s. l., where similar dynamics for real and predicted data are observed, with low error percentages. Specifically, at 2 m. a. s. l. the error obtained was 0,75%, which indicates the model's effectiveness to predict in this altitude and time (1 day) is elevated for 1 day at 1289 and 2743 m. a. s. l. the error percentages obtained were 3.43 and 2.81 %, respectively, indicating that the prediction model's effectiveness for these heights and time is elevated (Table 1). The variations maximum of 0.3°C is present in the distances 5 and 7 at 2743 m. a. s. l. (Figure 7).

In the same way, Figure 8 shows the predicted versus real temperature values at 3658 and 4572 m. a. s. l. with error values of 9.39 and 38.52%, respectively (Table 1), so the model is considered to have a moderate effectiveness at 3658 m. a. s. l. while the error at 4572 m. a. s. l. is considered high. The variations generated in the prediction versus the real at 3658 m. a. s. l. and 5467 m. a. s. l. in 1-day prediction (Figure 8), are mainly due to the turbulence generated by the non-linear nature of the meteorological prediction model, these variations are of 0.6°C.

The error values of the prediction model at 1 day for every evaluated height show a good prediction for 2, 1829 and 2743 m. a. s. l. at the same time, a moderate prediction is shown at 3658 and 4572 m. a. s. l. However, the temperature values presented a very low oscillation and varied in less than 1°C. Table 1 shows the error percentage of the prediction model for 1, 4 and 10 days; the results show a high effectiveness for the predictions at 2, 1829, 2743 and 3658 m. a. s. l. Nevertheless, the prediction at 4572 m. a. s. l. was of moderate effectiveness.

And it was possible to analyze that the model has good stability, even when the error generated by the temporal variation increases with time.

Table 1. Error percentage of the prediction model for each altitude studied for 1, 4 and 10 days

Altitude (m.a.s.l.)	% Error (°C)		
	1 Day	4 Days	10 Days
2	0,76	0,17	18,27
1829	3,43	2,10	1,92
2743	2,81	0,19	5,98
3658	9,39	4,65	7,36
4572	38,52	12,10	25,10

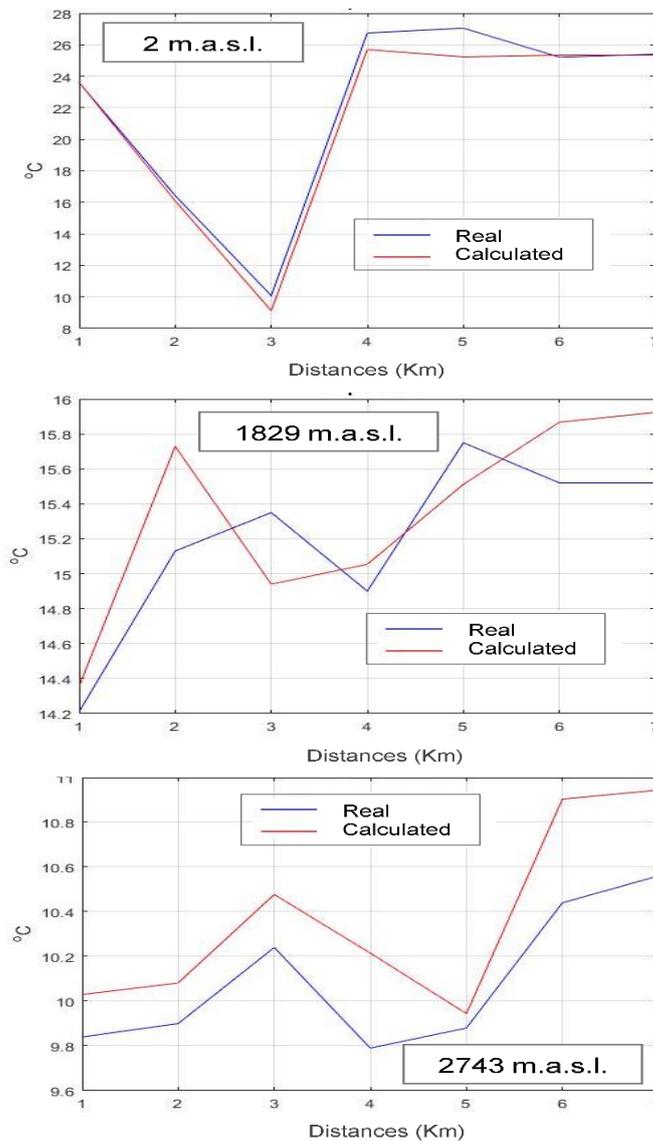


Figure 7. Real versus calculated temperature at 2, 1829, and 2743 m. a. s. l. for 1 day

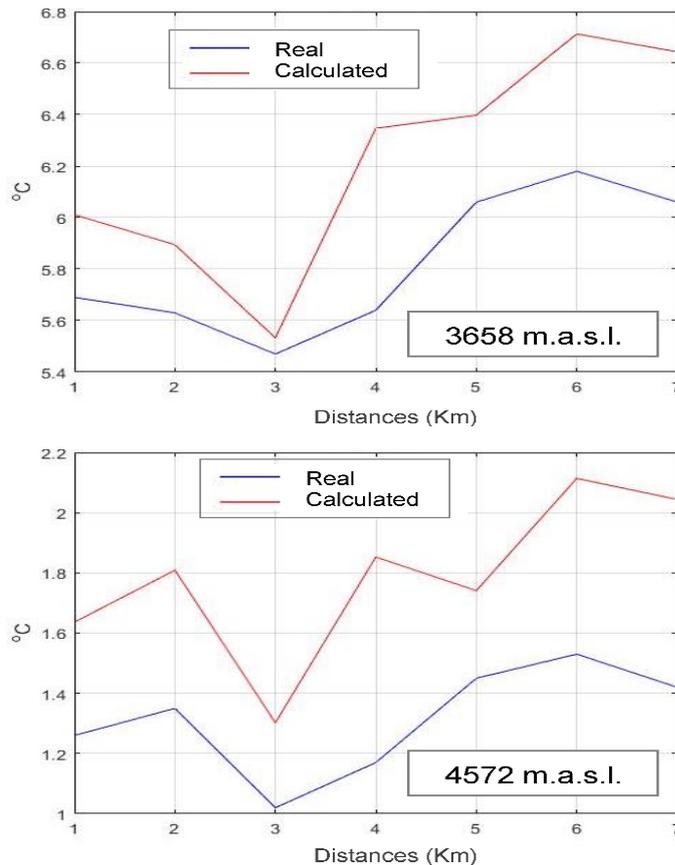


Figure 8. Real versus calculated temperature at 3658 and 4572 m. a. s. l. for 1 day

5. Conclusions

From the five sigma layers that represented the atmosphere in this work, it was observed how in the troposphere the temperatures are inversely proportional to altitude. Hthat represented the atmosphere in this work, it was observed how in the troposphere the temperatures are inversely proportional to altitude. Its value in degrees Celsius descends very quickly at different distances at the level of the surface. Due to this oscillation, it becomes more difficult the predictions as the values range are smaller.

The numerical prediction of the meteorological variables increases its complexity every day, due to the fact that the interaction with atmospheric pollution, geography and the ecosystems is constantly changing. Therefore, it is necessary to develop more robust prediction models, where equation systems integrate as many variables as possible and increase their effectiveness.

The meteorological prediction model obtained a maximum error of 38,52% for the temperature variable at 4572 m.a.s.l. on 1 day and a minimum error of 0,17% at 2 m.a.s.l. in a 4 day prediction, which shows the effectiveness of the model to predict the temperature variable, and the variation of said effectiveness due to the non-linear nature of the system. Continued improvements are needed for the meteorological prediction system, which may include predicting more meteorological variables, mainly to accomplish precipitation predictions, which will generate a more robust system with greater precision on different space-time variations.

References

- [1] Lorenz EN. Deterministic Nonperiodic Flow. *J Atmospheric Sci* Vol.20, No 2, p. 130–41, 1963.
- [2] Lezaun M. Weather forecasting and mathematics. *Bol Soc Esp Matemática Apl* Vol 23, pp. 145–66, 2003.
- [3] Lezaun M. Mathematics of fluids, climate, and weather. *Encuentros Multidisciplinarios* Vol23, pp. 46–55, 2006.

- [4] Hartmann D. The Atmospheric General Circulation and Its Variability. *Journal of the Meteorological Society of Japan*, Vol. 85B, pp. 123--143, 2007.
- [5] Hengl, T., Heuvelink, G.B., Percec T, M., Pebesma, E.J. Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theor Appl Climatol* Vol 107 pp. 265–277, 2011.
- [6] Radhika, Y., and Shashi M. Atmospheric Temperature Prediction using Support Vector Machines. *International Journal of Computer Theory and Engineering*, Vol. 1, No. 1, April 2009.
- [7] Pal, N.R., Pal, S., Das, J. and Majumdar, K. "SOFM-MLP: a hybrid neural network for atmospheric temperature prediction," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2783-2791, Dec. 2003.
- [8] Xue, M., Droegemeier, K.K., Wong, V., Shapiro, A., Brewster, K., Carr, F., Weber, D., Liu, Y., and Wang, D. The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorol. Atmos. Phys*, vol 76, pp. 143-165, 2001.
- [9] M. Xue, K. Droegemeier, and V. Wong, "The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification," *Meteorol. Atmos. Phys*, vol. 75, pp. 161–193, 2000.
- [10] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation, *Meteorol Atmos Phys*, vol. 82, pp. 139–170, 2003.
- [11] Barnett, T.P., Graham, N., Pazan, S., White, W., Latif, M and Flügel, M., ENSO and ENSO-related Predictability. Part I: Prediction of Equatorial Pacific Sea Surface Temperature with a Hybrid Coupled Ocean–Atmosphere Model. *International Journal of Climate*. Vol. 22, pp. 1545–1566, 2003.
- [12] M. Fischer and A. Navarra, "GIOTTO: A coupled atmosphere-ocean general-circulation model: The tropical Pacific," *Q. J. R. Meteorol. Soc.*, vol. 126, no. 567, pp. 1991–2012, 2000.
- [13] NCAR, 2017. User's guide for the Advanced Research WRF (ARW) modeling system, Version 3. National Center for Atmospheric Research. [Available online at http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/contents.html.]
- [14] Powers, J.G., Klemp, J.B., Skamarock, W.C., Davis, C.A., Dudhia, J., Gill, D.O., Coen, J.L., Gochis, D.J., Ahmadov, R., Peckham, S.E., Grell, G.A., Michalakes, J., Trahan, S., Benjamin, S.G., Alexander, C.R., Dimego, G.J., Wang, W., Schwartz, C.S., Romine, G.S., Liu, Z., Snyder, C., Chen, F., Barlage, M.J., Yu, W. and Duda, M.G. The Weather Research and Forecasting model: overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, vol. 98, pp. 1717–1737, 2017.
- [15] Michalakes JG, McAtee M, Wegiel J. Software Infrastructure for the Weather Research and Forecast Model. *Proceedings 1th ECMWF Workshop on the Use of High Performance Computing In Meteorology*. Austin, Texas, 2002.
- [16] Reichler T, Kim J. How Well Do Coupled Models Simulate Today's Climate? *Bull. Am. Meteorol. Soc.* Vol. 89, No.3, pp. 303–12, 2008.
- [17] MM5 Community model. n.d. MM5 Modeling System Overview. Accessed June 1, 2017. <http://www2.mmm.ucar.edu/mm5/overview.html>
- [18] Palacios JR, Ramón J, Herrera KC. Generation of Prognostics of the State of Time for the Early Alert System of the Pamplonita River Basin. *J Contemp Eng Sci* vol. 11 No. 49, pp. 2439–47, 2018.
- [19] Mena, D., Cañon Ramos, M. A., & Suárez Rodríguez, A. M. Modelación de los impactos de los escenarios de cambio climático en la cuenca del Río Pamplonita en Norte de Santander. *Revista EIA*, vol. 17, No.33, pp. 1–12, 2020.

- [20] Ramón J, Palacios JR, Santos Granados GR, Ramón Valencia JD. Early Warning System on Extreme Weather Events for Disaster Risk Reduction. *Rev Fac Ing Univ Antioquia*, vol.92, pp. 96–104, 2019.
- [21] Montoya G. Lessons in Dynamic Meteorology and Atmospheric Modeling. *Ed: Universidad Nacional De Colombia*. ISBN: 978-958-7019-66-7 v. 1 pages. 322. 2008.
- [22] National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. 2015, updated daily. NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/D65D8PWK>. Accessed 01-20-2017.
- [23] J. P. Sierra, P. A. Arias, S. C. Vieira, and J. Agudelo, How well do CMIP5 models simulate the low-level jet in western Colombia?, *Clim. Dyn.*, vol. 51, pp. 2247–2265, 2018.
- [24] A. K. Steiner *et al.*, Observed temperature changes in the troposphere and stratosphere from 1979 to 2018, *J. Clim.*, vol. 33, no. 19, pp. 8165–8194, 2020.