International Journal of Multidisciplinary Research and Growth Evaluation.

Forecast for cyclones that may affect the province of Villa Clara, Cuba, starting in 2023

Ricardo Osés Rodríguez [1], Amaury Machado Montes de Oca [2], Dr. David del Valle Laveaga [3], Nancy Ruiz Cabrera [4], Julia Socarras Padrón [5], Dr. Rigoberto Fimia Duarte [6*]

[1, 2, 4, 5] Department of Climate Group of the Forecasting, Provincial Meteorological Center of Villa Clara, Cuba

[3] Department of Parasitology. Regional High Specialty Hospital (HARE), Dr. Juan Graham Casasús, México

[6] Faculty of Health Technology and Nursing, University of Medical Sciences of Villa Clara, Cuba

* Corresponding Author: **Dr. Rigoberto Fimia Duarte**

Article Info

ISSN (online): 2582-7138
Volume: 05
Issue: 01
January-February 2024
Received: 01-01-2024
Accepted: 03-02-2024
Page No: 969-974

Abstract

Among the causes that originate the maximum wind speed in Cuba are tropical cyclones (hurricanes, extra tropical systems of the winter season (extra tropical lows and cold fronts); severe local storms typical of the summer and strong breezes due to the influence of continental and oceanic high pressures. This work models and forecasts in the long term when Villa Clara province could be impacted by a hurricane, starting in the year 2023. With this forecast we will be able to be alert about the risk of impact using local mathematical and statistical models. In order to carry out the study, a database of cyclones that have affected Villa Clara province from 1886 to 2017 was used, and the variables Year, Month, Day in which they affected Villa Clara were created, as well as the speed calculated according to the reanalysis data. First, these variables were modeled with the help of the ROR methodology, and then a long-term forecast was made, of when we could be affected by a tropical organism. Highly significant models with small errors were obtained, which allow us to obtain the dates and speed in which we could be affected by cyclones in Villa Clara. For 2023, on October 13, we could be affected with speeds higher than 100 km/h, according to the classification of tropical cyclonic organisms for winds between 64 and 118 km/h they are called tropical storms, therefore, this category is present in our forecast for 2023, while in 2025, 2030 and 2034 a hurricane of low intensity may occur.

Keywords: Cyclones, forecast, ROR regression, Villa Clara, wind speed

1. Introduction

Currently, there are several methods to predict the occurrence of some phenomenon or outcome, which are encompassed in predictive analytics [1, 2]. Predictive analytics is a sub-discipline of data analysis that uses statistical techniques, such as computational learning or data mining, to develop models that predict future events or behaviors [3-5]. These predictive models make it possible to take advantage of behavioral patterns found in current and historical data to identify risks [6-8]. This type of analysis is based on identifying relationships between variables in past events, and then exploiting these relationships to predict possible outcomes in future situations [3, 9, 10]. Doing this is not easy because it must be taken into account that the accuracy of the results obtained depends very much on how the data analysis has been carried out, as well as on the quality of the assumptions [7, 8, 10].

Trivially, it might seem that predictive analytics is the same as forecasting (which makes predictions at a macroscopic level), but no, it is something completely different. In a crude example, while a forecast can predict how many hurricanes may form in a year, predictive analytics can indicate how strong and what time of year they are most likely to form, and even where. Therefore, to perform predictive analytics it is essential to have a large amount of data, both current and past, in order to establish patterns of behavior and thus induce knowledge [1, 3, 10].

In the example above, there is more probability of prediction if variations in regional and global temperature, wind direction, changes and sources of pressure change etc. are also considered. This process is accomplished by computational learning. Computers can "learn" autonomously and thus develop new knowledge and capabilities, which requires large databases and the tools of predictive analytics [7, 8, 11].

Currently there are several techniques applicable to predictive analysis; i) regression, which includes linear, nonlinear, and adaptive multivariate regression; support vectors, ii) computational learning, where neural networks, Naïve Bayes and K-nearest neighbors are included. One of these tools is the Objective Regressive Regression method which we will explain briefly below [6, 7, 10]. Several applications are included in the bibliography, and the idea is to extend this type of analysis to social and epidemiological phenomena such as the COVID-19 epidemic in Santa Clara, Cuba using atmospheric pressure as an exogenous variable [4, 12, 13].

Important works on the maximum wind and the probability of hurricane affectation in Cuba have been carried out. [14-16]. Among the causes that originate the maximum wind speed in Cuba are tropical cyclones (hurricanes, extra tropical systems of the winter season, (extra tropical lows and cold fronts), severe local storms typical of summer and strong breezes due to the influence of continental and oceanic high pressures [15-17].

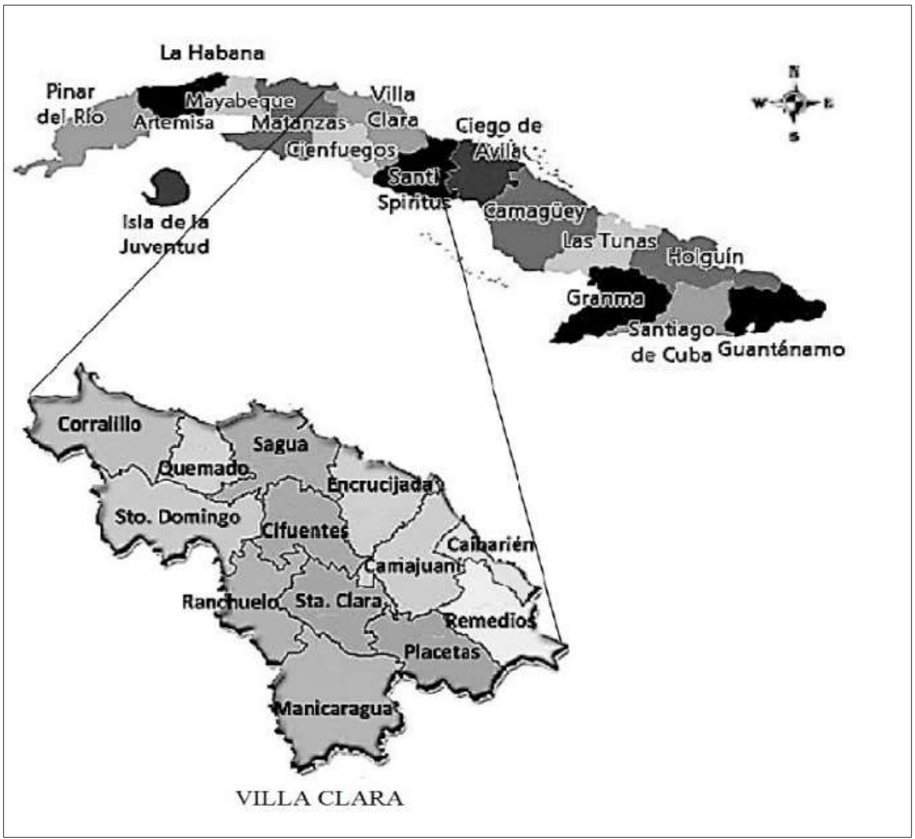
Some conclusions on the impact of hurricanes in Cuba state that the annual number of hurricanes that form in the Atlantic Ocean have a high interannual and multiannual variability, such variability has been associated with the changes that occur in atmospheric and oceanic circulation, and an increase in the formation of hurricanes in the Atlantic Ocean has been noted, mainly since the middle of the last decade of the last century [16-18]. Hurricanes are an indissoluble part of the climate in Cuba and will continue to be a threat to the economy and society until they can be prevented and predicted with certainty. Regarding hurricane activity, the results achieved in recent years indicate that it is possible and essential to prepare to face their danger and mitigate their onslaught [9, 16, 18, 19].

The objective of the research consisted of modeling and forecasting in the long term when the inhabitants of Villa Clara province could be impacted by a tropical meteorological event.

2. Materials and Methods

2.1. Study area

The research was carried out in Villa Clara province, Cuba, whose provincial capital is Santa Clara municipality and covered the 13 municipalities that comprise it. This province is located in the central region of the island of Cuba (Latitude: 22º 29'40" N, Longitude: 79º28'30" W), and has the following geographical limits; to the west, with Matanzas province, to the east, with Sancti Spíritus province and to the south, with Cienfuegos province (Figure 1).



***Source*:** Provincial Meteorological Center of Villa Clara

Fig 1: Administrative map of Villa Clara province

2.2. Procedures for information processing

For the realization of this work, we had a database (Cycology), of all the cyclones that have affected Villa Clara province, from 1886 to 2017. The variables Year, MONTH and DAY in which they affected Villa Clara were created, as well as the wind speed calculated according to the reanalysis data.

2.3. Mathematical modeling

The variables (Year, Month, Day and Wind Speed) were modeled using the Objective Regressive Regression (ORR) methodology, and then a long-term forecast of when we might be affected by a tropical organism was made.

For the development of the predictive model, the methodology of Regressive Objective Regression (ROR) was used, which made possible the prediction of the foci by means of the ROR methodology [4, 20], for which, in a first step, dichotomous variables DS, DI and NoC are created, where:

NoC: Number of cases in the base,

DS = 1, if NoC is odd; DI = 0, if NoC is even, when DI=1, DS=0 and vice versa.

Subsequently, the module corresponding to the Regression analysis of the statistical package SPSS version 19.0 (IBM Company) will be executed, specifically the ENTER method where the predicted variable and the ERROR are obtained.

Then the autocorrelograms of the variable ERROR are obtained, paying attention to the maximums of the significant partial autocorrelations PACF. The new variables were then calculated taking into account the significant Lag of the PACF. Finally, these regressed variables were included in the new regression in a process of successive approximations until a white noise in the regression errors was obtained. For the case of atmospheric pressure, lags of 1 year in advance can be used, as other authors have done for the climatic indexes, although it is unlikely that 11 years in advance results will be obtained, since we only have 11 years of data in the base, nevertheless, in the monthly data we will try to use the results for the meteorological variable atmospheric pressure.

3. Results and Discussion

First, the four variables were modeled (Table 1), with their main model statistics. The variance explained was over 91 %; the errors were found to be slightly small, while the Durbin Watson statistic was close to 2, so that the variables entering the models accurately reflect the behavior of the data. Fisher's F's were all significant at 100 %, and correlations between actual and predicted values were significant at 99 %.

Table 1: Main statistics of the models for each variable

Variables	Explained Variance R	Estimation error	Durbin Watson	F of Fisher	Correlation %
Year	100 .00	3.08 año	1.600	3276587.0***	0.993**
Month	99.6	0.93 mes	2.408	602.2***	0.595**
Day	91.8	7.8 día	2.074	18.8***	0.608**
Speed	91.6	50.6 km/h	1.965	40.2***	0.482**

** Significance to the 99 %. *** Significance to the 100 %.

Table 2: Trend results for the variable Year

Coefficients[a, b]						
Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	DS	.345	1.262	.000	.274	.787
	Tendency	.509	.096	.008	5.310	.000
	Lag18YEAR	1.021	.002	.992	629.668	.000
a. Dependent variable: YEAR						
b. Linear regression through the origin						

Table 3: Trend results for the dependent variable Month

Coefficients[a,b]						
Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	DS	28.778	17.571	2.186	1.638	.117
	DI	29.027	17.607	2.205	1.649	.115
	Lag18Year	-.007	.009	-1.452	-.764	.454
	Lag5Month	-.660	.214	-.662	-3.081	.006
a. Dependent variable: Month						
b. Linear regression through the origin						

Table 4: Trend results for the dependent variable Day

Coefficients[a,b]						
Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	DS	197.445	147.954	8.266	1.335	.198
	DI	196.235	148.257	8.215	1.324	.201
	Lag18Year	-.090	.077	-10.213	-1.164	.259
	Lag6DAY	-.551	.183	-.540	-3.015	.007
	Lag8DAY	-.148	.180	-.150	-.822	.421
a. Dependent variable: DAY						
b. Linear Regression through the origin						

Table 5: Trend results for the dependent variable Speed

Coefficients[a,b]						
Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	DS	90.437	33.411	.533	2.707	.011
	DI	111.290	31.938	.675	3.485	.001
	Tendency	1.479	.856	.337	1.726	.094
	Lag7Spe.	-.325	.182	-.309	-1.791	.083
a. Dependent variable: Speed b. Linear regression through the origin						

Table 6: Summary of forecasts as of 2022

Case summaries[a]					
		Predicted Value YEAR	Predicted Value MONTH	Predicted value Day	Predicted Value Speed
1		2011.89303	8.67778	8.12490	146.04346
2		2012.74705	8.42851	12.92095	97.07706
3		2014.95297	9.32351	7.90350	155.50427
4		2016.82830	7.74752	7,19826	100.35937
5		2022.09814	10.60088	15.22888	137.32441
6		2022.95216	9.69177	13.34521	100.06465
7		2025.15809	9.26710	9.95769	137.02969
8		2030.09733	8.98962	15.17628	142.04391
9		2034.34587	.	11.72144	139.98682
10		2046.43425	.	.	81.59007
11		2048.64018	.	.	.
12		2053.57942	.	.	.
13		2053.74273	.	.	.
14		2059.70328	.	.	.
15		2065.99443	.	.	.
16		2068.89106	.	.	.
17		2072.11829	.	.	.
18		2077.05754	.	.	.
19		2080.28477	.	.	.
20		2081.13879	.	.	.
21		2090.49385	.	.	.
Total	N	21	8	9	10
a. Limited to the first 100 cases.					

Next (Table 2-5), the models where the tendency to increase for the variable YEAR, and the tendency to increase the cyclone velocities can be appreciated, where some variables were not significant, but as they contribute variance to the model, they will remain in the models. The variable Month regressed on 5 months (Lag5MONTH), the variable DAY regressed on 6 (Lag6DAY) were highly significant, as well as the mean velocity regressed on 7 phenomena (Lag7Vel). Below (Table 6), by way of summary, are the forecasts from the year 2022 onwards.

The results obtained corroborate those of other authors in this regard [21-24], stating that with each passing day, natural disasters intensify and threaten the safety of people. According to the Center for Research on the Epidemiology of Disasters [25], 102 countries had suffered some type of natural disasters by the end of 2016, leaving numerous losses of human lives and substantial economic losses [26].

Hydrometeorological events are intensifying, with cyclones constituting the most destructive phenomenon in the tropics, for all that they bring in their wake: strong winds, sea penetrations, landslides and intense rainfall [27-30].

The 2017 hurricane season marked several absolute values, with hurricanes Harvey, Irma, José and Katia standing out. Irma reached category 5, even before reaching the Caribbean Sea; it was also the hurricane with the longest duration with this intensity [16]. It was also the most devastating meteor in terms of material damage in the region. [16, 25].

The strong and increasingly frequent and unusual summers and winters, floods, droughts, meteorological disturbances (gales, tropical storms, cyclones, hurricanes, among others), reinforced by the sporadic intervention of "El Niño and La Niña" and even more worrying, the inconsequential participation of man, as well as the increasing increase in air and maritime transport, are worryingly perpetuating these and other epidemiological episodes [26, 32-34].

According to the classification of tropical cyclonic organisms [15, 31], for winds between 64 and 118 km/h they are called tropical storm, so in our forecast for 2023 this category is present, while for 2025, 2030 and 2034 a hurricane of low intensity may occur.

4. Conclusion

Highly significant models with small errors were obtained, which allow us to obtain the dates and speed in which we could be affected by cyclones in the province of Villa Clara. With this forecast we will be able to be alert about the risk of impact using local mathematical and statistical models, and thus minimize the risks and damages, both personally and to the economy of the province and the country.

Institutional Review Board Statement

Not applicable.

Transparency

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study, and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

5. References

1. Archontoulis SV, Miguez FE. Nonlinear Regression Model and Applications in Agricultural Research. Agronomy Journal. 2015;107(2):786-798.
2. Espino CT. Predictive analytics: techniques and models used and applications thereof [doctoral thesis]. Catalonia: Publications and Scientific Exchange Service, University of Catalonia, Spain; c2017.
3. Osés RR, Otero MM, Fimia DR, Osés LC, Burgos AI, Ruiz CN, Socarrás PJ. Modeling and 11 year ahead monthly forecasting of precipitation and temperature extremes for Cuba. The Biologist (Lima); c2018a. p. 16.
4. Fimia DR. Mathematical modeling of population dynamics of the *Aedes aegypti* (Diptera: Culicidae) mosquito with some climatic variables in Villa Clara, Cuba. International Journal of Zoology and Animal Biology (IZAB). 2020a;3(3):16000233.
5. Fimia DR, Osés RR, González GR, Hernández CN, Otero MM. The entomofauna of culicidae and copepods approached from biological control alternatives to mathematical modeling in two central provinces of Cuba. Anales de la Academia de Ciencias de Cuba (AACC). 2020b;10(3):1-11.
6. Osés R, Grau R. Regression modeling (ROR) versus ARIMA modeling using dichotomous variables in HIV mutations. Central University "Marta Abreu" of Las Villas. Feijóo Publishing House; c2011.
7. Mahanta DY, Borah B. Parameter Estimation of Weibull Growth Models in Forestry. International Journal of Mathematics Trends and Technology. 2014;8(3):157-163.
8. Dagogo JE, Nduka C, Ogoke UP. Comparative Analysis of Richards, Gompertz and Weibull Models. IOSR Journal of Mathematics. 2020;16:15-20.
9. Osés RR, Fimia DR, Silveira PE, Hernández VW, Saura GG, Pedraza MA, González GR. Mathematical model of the density of *Anopheles* mosquito larva (Diptera: Culicidae) in 2020 in Caibarien, Villa Clara province, Cuba. REDVET. 2012, 13(3). URL: http://www.veterinaria.org/revistas/redvet/n080814B.html.
10. Ban GA. Statistical modeling of the novel COVID-19 epidemic in Iraq. Epidemiol Methods. 2021;10:20200025.
11. Ivorra B, Ramos AM. Validation of the forecasts for the international spread of the coronavirus disease 2019 (COVID-19) done with the Be-CoDiS mathematical model. Technical Report MOMAT Group, Complutense University of Madrid; c2020. p. 1-14. https://www.doi.org/10.13140/RG.2.2.3146 0.94081
12. Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT., *et al*. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: A modeling study. Lancet Infect Dis; c2020. https://doi.org/10.1016/ S1473-3099(20)30162-6
13. Min L. Modelling, Simulations and Analysis of the First and Second COVID-19 Epidemics in Beijing (preprint). Med Rxiv; c2021. p. 1-19. https://doi.org/10.1101/2021.07.04.21259205
14. Osés RR, Fimia DR, Aldaz CJW, Iannacone OJ, Zaita FY, Osés LC, *et al*. Mathematical modeling of cholera by means of Objective Regression and its relation with climatic variables. Caibarién, Villa Clara, Cuba. The Biologist (Lima); c2017. p. 15 (Special Supplement 1): 128.
15. Chowdhury A, Sajid M, Jahan N, Adelusi TI, Maitra P, Yin G, *et al*. A secondary approach with conventional medicines and supplements to recuperate current COVID-19 status. Biomedicine & Pharmacotherapy. 2021;142:111956.
16. Osés RR, Fimia DR, Osés LC, Zambrano GMA, Santos ZTB, González MA. Forecast of New and Deceased Cases of COVID-19 in Cuba with an Advance of 105 Days. Acta Scientific Veterinary Sciences (ASVS). 2021a;3(7):31-36.
17. Roura P, Sistachs V, Vega R, Izquierdo D. Havana "Wonder City": its history from Meteorology 500 years after its founding. Cuban Journal of Meteorology. 2019;25(3):481-490.
18. Planos E, Vega R, Guevara A. (Editors). Climate Change Impact and Adaptation Measures in Cuba. Institute of Meteorology, Environment Agency, Ministry of Science, Environment and Technology, Editorial AMA. La Habana, Cuba. 2013;1:430. ISBN: 978-959-300-039-0.
19. Osés RR, Osés CL, Fimia DR, González MA, Iannacone J, Santos ZTB, Wilford GFM. Age Prediction for COVID-19 Suspects and Contacts in Villa Clara Province, Cuba. EC Veterinary Science. 2021b;6(4):41-51.
20. Fimia DR, Osés RR, Otero MM, Diéguez FL, Cepero RO. González GR., *et al*. The control of mosquitoes (Diptera: Culicidae) using biomathematical methods in Villa Clara province. REDVET. 2012;13(3):3-12.
21. Zhang Z, Krishnamurti TN. A perturbation method for hurricane ensemble predictions. Monthly Weather Review. 1999;127:447-469.
22. Gore A. An Inconvenient truth [videocinta] EUA: Paramount Classics and Participant Productions; c2007.
23. Dodla VB, Desamsetti S, Yerramilli A. A Comparison of HWRF, ARW and NMM Models in Hurricane Katrina (2005) Simulation. International Journal of Environmental Research and Public Health. 2011;8:2447-2469.
24. Osés RR, Otero MM, Ruiz CN, Fimia DR, Iannacone J. Prognosis for hurricane Irma through Regression Objective Regressive and its impact on the vector populations at the meteorological station of Caibarién, Villa Clara, Cuba. Biotempo (Lima). 2018b;15(1):23-30.
25. Chowell G. Fitting Dynamic Models to Epidemic Outbreaks with Quantified Uncertainty: A Primer for

Parameter Uncertainty, Identifiability, and Forecast. Infectious Disease Modelling. 2017;2:379-398.

26. Fimia DR, Osés RR, del Valle LD, Yee D A. Mathematical modeling and climate: incidence, repercussion and impact on communicable entities and vector organisms. Himalayan Journal of Applied Medical Sciences and Research. 2022;3(4):52-65.
27. Durán MN, Botello RE. Detection of emerging “active” clusters with high incidence rates, for rapid surveillance of COVID-19. Electronic Medicentro. 2020;24(3):643-656.
28. Goerss JS. Tropical cyclone track forecasts using an ensemble of dynamical models. Monthly Weather Review. 2000;128(4):1187-1193.
29. Vecchi GA, Delworth T, Gudgel R, Kapnick S, Rosati A, Wittenberg AT. On the seasonal forecasting of regional tropical cyclone activity. Journal of Climate. 2014;27(21):7994-8016.
30. Ruan Z, Wu Q. Precipitation, convective clouds, and their connections with tropical cyclone intensity and intensity change. Geophysical Research Letters. 2018;45(2):1098-1105.
31. Trenberth K. Uncertainty in hurricanes and global warming. Science. 2005;308(5729):1753-1754.
32. Arnell NW, Gosling SN. The impacts of climate change on river flood risk at the global scale. Climatic Change. 2016;134:387-401.
33. Roy C, Kovordányi R. Tropical cyclone track forecasting techniques-A review. Atmospheric Research. 2012;104:40-69.
34. Xie L, Bao S, Pietrafesa LJ, Foley K, Fuentes M. A real-time Hurricane surface wind forecasting model: formulation and verification. Monthly Weather Review. 2006;134(5):1355-1370.