

# Long Term Forecast of Meteorological Variables in Sancti Spiritus. CUBA

Ricardo Osés Rodríguez<sup>1,\*</sup>, Rigoberto Fimia Duarte<sup>2</sup>, Guillermo Saura González<sup>1</sup>, Alfredo Pedraza Martínez<sup>1</sup>, Nancy Ruiz Cabrera<sup>1</sup>, Julia Socarras Padrón<sup>1</sup>

<sup>1</sup>Provincial Meteorological Centre of Villa Clara, CUBA Calle Marta Abreu No 59 Altos Esquina a Juan Bruno Sayas, CP

<sup>2</sup>Unity of Surveillance and Antivectorial Fight, Centre of Health and Epidemiology, Villa Clara

\*Corresponding author: ricardo.oses@vcl.insmet.cu

Received September 10, 2013; Revised February 10, 2014; Accepted February 20, 2014

**Abstract** The aim of this work is aimed at modeling and forecasting with 1 year in advance a set of 7 meteorological variables, these are, as long as the wind keeps blowing over, 3 m/s, 4 m/s, 5 m/s, 6 m/s, 7 m/s, 8 m/s and 9 m/s corresponding to the meteorological station of Sancti Spiritus (Lat North 21°56', Long 79°27', Height above sea level 96.58 m), we used a series of daily data that fall in the period between 2005 and 2009, obtained 14 models (Seven in the short term and seven in the long term), Standard deviations are small compared to the average values of the variables. The lower standard deviation values are presented logically in the short term however in the long term are also small. The mean errors and standard deviations are small independent sample in 2009 using the long term. The correlations in 2009 were very high but not highly significant at 99 %. All the equations were significant at 99 %. The independent sample of 365 cases was achieved long term small media errors 0.326 values for the variable in which the wind is over 9 m/s to -3.14 when the wind is above of 3 m / s. Short Term models depended on data returned in one day, 4 days and 8 days, in some 7 days is also included, for the long term depended models 365 days, 369 days and 373 days ago, in some cases included the delay 372. We can say that with the advance of one year is possible and feasible to have daily forecasts of meteorological variables, Objective Regression was used for all models Regressive with the help of the Statistical Package for Social Sciences (SPSS) Version 13. The tables and graphs show the predicted and actual values for 2009. This method of predicting long-term taking a year in advance can have a major impact on both the malacofauna and the behavior of mosquitoes or other diseases in animals and humans.

**Keywords:** long-term forecast, wind, Cuba, mathematical modeling, mosquitoes

**Cite This Article:** Ricardo Osés Rodríguez, Rigoberto Fimia Duarte, Guillermo Saura González, Alfredo Pedraza Martínez, Nancy Ruiz Cabrera, and Julia Socarras Padrón, "Long Term Forecast of Meteorological Variables in Sancti Spiritus. CUBA." *Applied Ecology and Environmental Sciences*, vol. 2, no. 1 (2014): 37-42. doi: 10.12691/aees-2-1-6.

## 1. Introduction

World Meteorological Organization (WMO) contributes significantly to the protection of lives and property through its programs and network of more than 190 National Meteorological and Hydrological Services. Weather forecasts and early warnings provided to governments, economic sectors and help people prevent and mitigate the effects of disasters.

The World weather Watch (WWW) has played a vital role in this regard. Established in 1963, during the Cold War, is a milestone in international cooperation. Groups observing systems, telecommunication centers and data processing and prediction by which all countries have the information and meteorological and environmental services needed to carry out an exchange of information in real time and provide effective services.

As more and more services needed weather and climate, and with the spectacular scientific and technological

advances, the WWW has now become the centerpiece of numerous programs, both WMO and other agencies. Essentially contributes to the priorities of WMO through improved observations, monitoring the atmosphere and oceans, and dissemination of weather forecasts worldwide, especially early warning and weather and impact weather.

Today, improved climate services are presented as one of the crucial tools to cope and adapt to climate change and climate variability. The assumption that climatic and socioeconomic conditions of the past suffice as an indicator of current and future conditions is not enough. It is imperative to continue to improve our understanding of climate and most appropriate use of climate information to address the needs of society in a world characterized by population growth, changes in land use, urbanization, and the difficulties in ensuring the food security and water resources management and energy.

An application domain where data are abundant is related to meteorology, which have developed different data sets with numerous meteorological variables. Since the early twentieth century, the weather prediction

problem has been addressed numerically, using atmospheric circulation models (systems of partial differential equations, or approximations thereof) from known initial conditions. Human predictors offer daily local and regional forecasts, in the villages, and weather phenomena of interest, mainly produced in surface phenomena (meteors as precipitation, temperature, etc.).

Moreover, attempts have been made to predict local meteors directly applying statistical techniques to historical records of observations of these phenomena, such as ARIMA models (and Autoregressive Integrated Moving Averages). Nevertheless, for certain weather elements such techniques have great limitations as [1].

i. due to the resolution of numerical models do not allow for accurate local predictions geographical points of interest and ii. statistical models do not include enough information about physical phenomena involved in these processes. Therefore, we have shown inefficient from an operational standpoint as [2].

Methodology However ROR as [3] presents advantages compared with ARIMA models, this methodology also includes a large amount of variance explained by STEP variables or pass to give information about local processes that occur over time There is no established algorithm for forecasting meteorological variables (Precipitation, temperature, cloud cover, wind, etc.). Each group of predictors follows a sequence of steps itself and unites the information obtained, for according to the site characteristics and their individual and collective experience, give a prognosis of these variables.

Recently developed automated applications for weather forecasting [4] and [5] are aimed at predicting the visibility and intensity of wind in airport terminals in Nova Scotia, Canada, the authors implement a fuzzy k-NN with which combines elements of case-based reasoning and fuzzy logic. In [5] and [2] Sordo and Cofiño presents methods of short-term local forecast for the area of Santander, Spain; combine the outputs of numerical models with statistical information on local observations. These systems are specific to local areas for which they were developed by the diversity of existing geographical and climatic characteristics.

In Cuba CubaForecast system was developed by the Meteorological Centre of Cienfuegos in 2003 as [6]. With him is possible for the different provinces of the country, the prognosis of meteorological variables. The synoptic scale basic information is using a mesh produced by the European Centre for Medium-Term Forecast. This is, according to weather experts, the only automated system in Cuba with which you can perform, automatically, the prognosis of temperatures. With this system it is true that there was a leap forward, have not yet achieved the desired results. Our work is aimed at modeling and forecasting spot with 1 year in advance of a set of 10 meteorological variables relevant to the wind on the basis of 365-day cycle that the earth moves around the sun which we will use as above information for the following year.

## 2. Materials and Methods

For the realization of this work were used daily weather data with the wind keeps blowing above 4 m/s, 5 m/s, 6m/s, 7 m/s, 8 m/s, 9 m/s corresponding to the Sancti Spiritus meteorological station (North Lat 21°56', Long 79°27', Height above sea level 96.58 m) used a series of daily data that fall in the period between 2005 and 2009 , 14 models obtained seven in short term and seven to long term.

ROR methodology was used as [7], the methodology of Box and Jenkins [8] was not used since this has limitations as [1], we used SPSS Version 13.

## 3. Results and Discussion

The long-term model take per sample was variable V @ 5 to be the time when the wind is blowing above 5 m/s.

The model with variables returned in 365 days, 369 days and 373 days before using the regressive objective regression resulted explained 63.6 % of variance with a standard error of 3.82 hours. (Table 1).

**Table 1. Some statistical parameters to log term for V @ 5**

Model Summary <sub>c,d</sub>					
Model	R	R Square <sub>a</sub>	Adjusted R Square	Std. Error of the Estimate	Durbin- Watson
1	.800 <sub>b</sub>	.639	.636	3.82903	1.024

a. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

b. Predictors: Step1153, Step1097, Step1032, Step688, Step448, DS, DI, Lag365cinco, Lag369cinco, Lag373cinco, NoC.

c. Dependent c. Variable: V @ 5.

d. Linear Regression through the Origin.

**Table 2. Variance analysis of the Long term model for V @ 5**

ANOVA <sub>c,d</sub>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27810.037	11	2528.185	172.437	.000 <sub>a</sub>
	Residual	15687.791	1070	14.661		
	Total	43497.829 <sub>b</sub>	1081			

a. Predictors: Step1153, Step1097, Step1032, Step688, Step448, DS, DI, Lag365cinco, Lag369cinco, Lag373cinco, NoC.

b. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

c. Dependent Variable: V @ 5.

d. Linear Regression through the Origin.

The Variance analysis was significant at 100 % , With an F de Fisher of 172.43 (Table 2).

In (Table 3), we can see the model parameters for V @ 5, DS and DI explain the ups and downs of the series, both

being significant at 99 %, the variables that influence the modeling are time back in 365 days (Lag365cinco), time returned in 369 days (Lag369cinco) and 373 (Lag373cinco) trend (NoC) over time was significantly to

the increase. STEP variables measure the importance of certain cases over time.

**Table 3. Long term model using ROR methodology**

Coefficients <sup>a,b</sup>						
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
	B	Std. Error	Beta			
1	DS	1.025	.448	.114	2.285	.022
	DI	1.301	.444	.145	2.930	.003
	NoC	.002	.000	.321	5.585	.000
	Lag365cinco	.122	.032	.110	3.862	.000
	Lag369cinco	.152	.032	.137	4.813	.000
	Lag373cinco	.113	.032	.103	3.584	.000
	Step448	16.155	3.849	.077	4.198	.000
	Step688	17.371	3.836	.083	4.528	.000
	Step1032	17.855	3.834	.086	4.657	.000
	Step1097	18.521	3.838	.089	4.825	.000
	Step1153	16.908	3.836	.081	4.408	.000

a. Dependent a. Variable: V @ 5.

b. Linear Regression through the Origin.

**Table 4. Residuals of the model to long term using ROR methodology**

Residuals Statistics <sup>a,b</sup>					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2.3610	23.5900	4.7876	1.67553	1081
Residual	-7.75262	19.14261	.00000	3.81126	1081
Std. Predicted Value	-1.448	11.222	.000	1.000	1081
Std. Residual	-2.025	4.999	.000	.995	1081

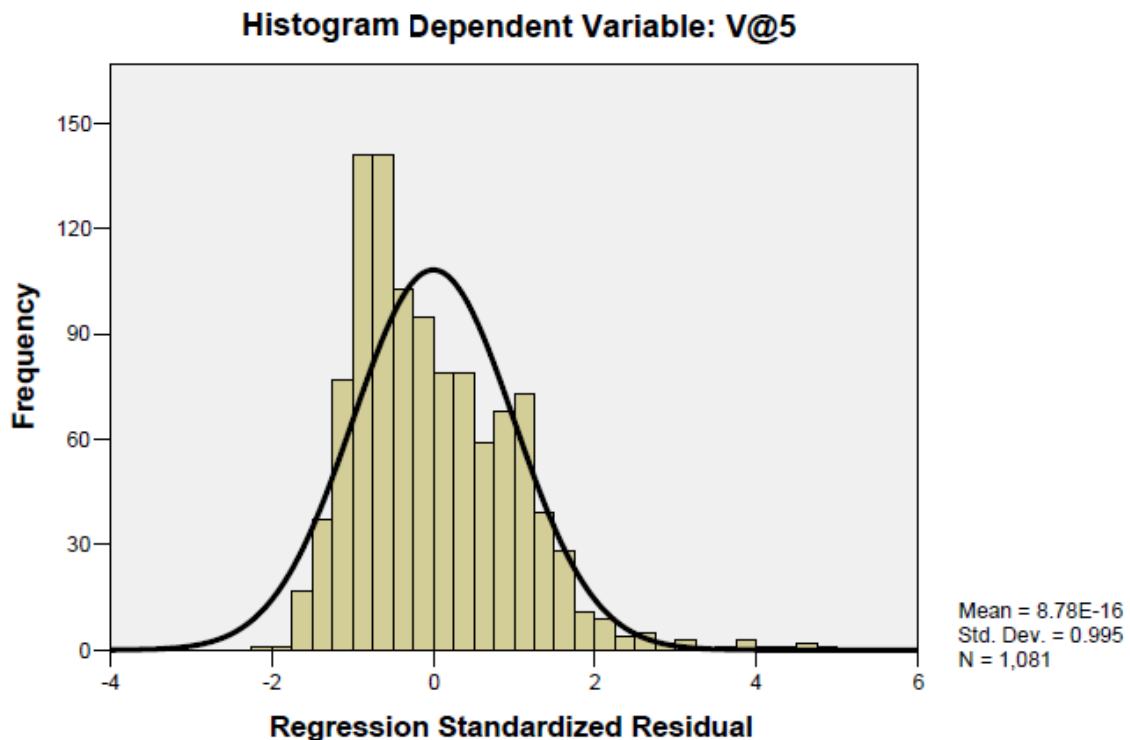
a. Dependent Variable: V @ 5.

b. Linear Regression through the Origin.

As seen standardized residuals (Table 4) have zero mean and standard deviation 0.995 close to one, the maximum residual is 19.14, the standard deviation of the residuals is 3.81.

Below is presented in (Figure 1) the frequency distribution of the residuals following these distribution close to normal very good thing for the model.

Figure 2 shows an almost straight line between the expected probability and Probability Eyed very beneficial standardized residuals for the model, it should be noted that it seems there is still due to the seasonal information inclined in two periods of this figure.



**Figure 1.** Distribution of frequencies of the Standardized Residuals using ROR methodology

## Normal P-P Plot of Regression Standardized Residual Dependent Variable: V@5

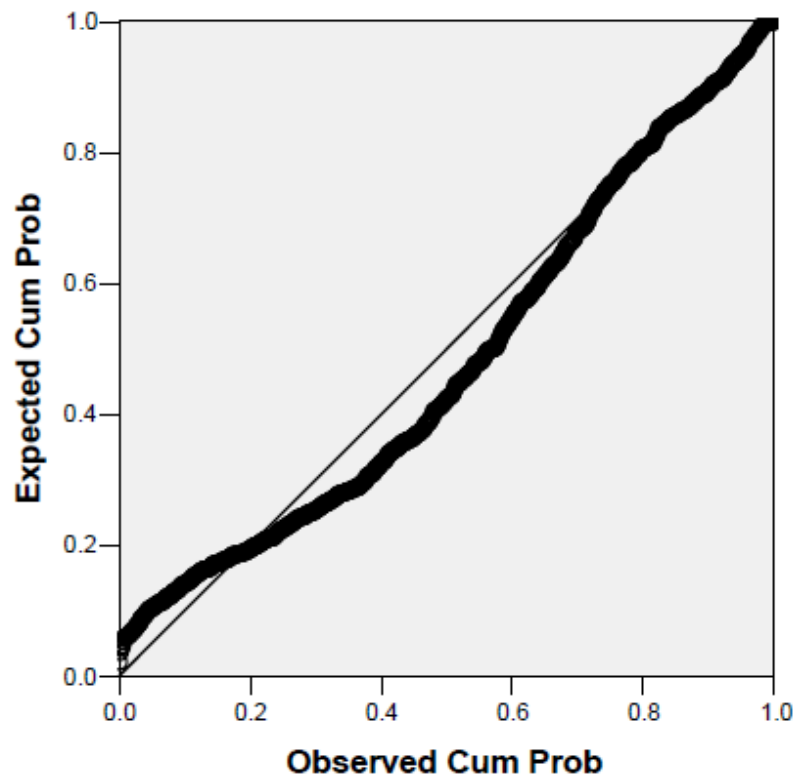


Figure 2. Plotting of the residuals probabilities of the model 2 using ROR methodology

Table 5. Descriptive statistics of meteorological variables

Model	3m/s	4m/s	5m/s	6m/s	7m/s	8m/s	9m/s
Mean Value	8.70	6.63	4.81	3.11	2.00	1.13	0.58
Standard Deviation	5.04	4.61	4.20	3.68	3.06	2.34	1.68

We conducted a descriptive statistical variables (Table 5), showing that the time above 3 m/s is the one with the highest average value and the standard deviation increased.

Once models calculated short term and long term (Table 6) presents the different parameters of the same, to appreciate that both the short and long term explained variance of the regression (R) presents quite high with values up to the variable 3 m/s with a value of 0.855, while the lowest value for the variable is presented 9 m/s to 0.603. The standard deviations (Table 6) are small compared to the mean values of the variables (Table 5),

the lowest values of standard deviation are presented logically in the short term however in the long term are also small. The Durbin Watson statistics in short-term models are close to two indicating the correlation is not as long-term residual is close to 1 indicating a possible correlation between the residuals. All models equations according to Fisher F are highly significant 99 %.

Short term models the returned data depended on 1 day 4days and 8 days, in some further included 7 days, for the long term depended models 365, 369 and 373 days ago, in some cases included the delay at 372.

Table 6. Results of short and long term models.\*\* significant al 99 %

Modelo	3m/s	4m/s	5m/s	6m/s	7m/s	8m/s	9m/s
R.Short term(CP)	0.855	0.832	0.785	0.731	0.675	0.645	0.603
Standard Deviation CP	3.821**	3.300**	2.954**	2.486**	2.073**	1.544**	1.121**
Durbin Watson CP	1.985	1.918	1.982	1.852	1.788	1.787	1.782
Fisher F CP	967.5**	525.8**	366.4**	257.8**	170.6**	131.7**	110.3**
R.Long term(LP)	0.786	0.731	0.636	0.511	0.446	0.446	0.565
Standard Deviation LP	4.690**	4.164**	3.829**	3.249**	2.566**	1.735**	1.066**
Durbin Watson LP	1.054	1.04	1.024	0.950	1.003	1.444	1.386
Fisher F LP	498.3**	226.7**	172.4**	95.14**	73.31**	63.76**	83.39**

Finally we calculated the long-term results in an independent sample for the year 2009, with the following results (Table 7). Both the average error and the standard deviations are small, the smallest average error occurs in

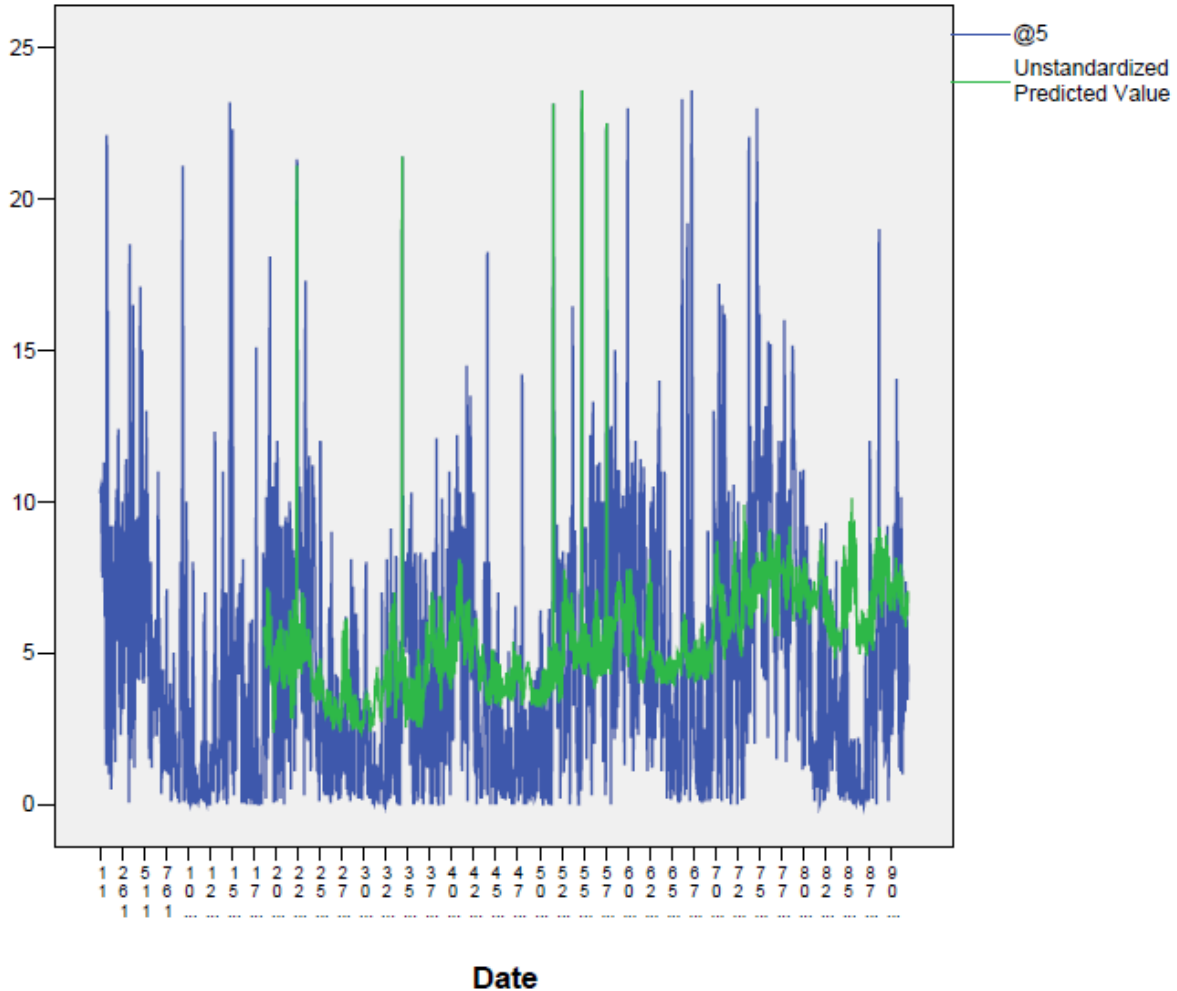
the variable in which the wind is above 8 m/s. and the smaller standard deviation for the variable occurs when the wind is over 9 m/s. The correlations are not very high but highly significant in 2009.

**Table 7. Correlations between real y forecasting data in 2009, mean values of the errors**

Models	3m/s	4m/s	5m/s	6m/s	7m/s	8m/s	9m/s
Mean Error Long term	-3.14	-2.24	-1.75	-1.15	-0.56	-0.04	0.13
Standard Deviation Long term	5.08	4.57	4.12	3.59	3.09	2.46	1.76
Correlation 2009	0.216**	0.228**	0.268**	0.262**	0.280**	0.224**	0.099

The plot of the long-term model for the variable @ 5, (wind above 5 m/s) is shown in Figure 3, one can see good

agreement between the model and the prognosis (unstandardized Predicted Value).



**Figure 3.** Results of real and predicted data for the wind over 5 m/s) (@ 5)

It is noteworthy that according as [9] in two models (Caibarien and Sagua) the wind was significant within 1 month with respect to the modeling of mosquito anofelinic density, this method of predicting long-term taking a year in advance can have a major impact on both the malacofauna behavior as mosquitoes or other diseases in humans and animals.

#### 4. Conclusions

1. It is seen that both the short and long term explained variance shows high values.
2. Standard deviations are small compared to the average values of the variables.
3. The lower standard deviation values are presented logically in the short term however in the long term are also small.

4. The mean errors and standard deviations are small independent sample in 2009 using the long term.
5. The correlations in 2009 were very high but not highly significant.
6. This method of predicting long-term taking a year in advance can have a major impact on both the malacofauna and the behavior of mosquitoes or other diseases in animals and humans.

#### References

- [1] Yaile Caballero Mota. Aplicación de la Teoría de los conjuntos aproximados en el preprocesamiento de los conjuntos de entrenamiento para los algoritmos de aprendizaje automatizado. Tesis defendida en opción al grado científico de Doctor en Ciencias Técnicas. 1 17/7/2007. Tribunal Nacional Automatica y Computacion, ISPJAE, Habana.
- [2] Cofiño A. Técnicas estadísticas y neuronales de agrupamiento adaptativo para la predicción de fenómenos meteorológicos locales. Aplicación en el corto plazo y en la predicción estacional

- [Tesis Doctoral]. Santander, España: Universidad de Cantabria; 2004.
- [3] Oses, R.; Grau R. (2011). Modelación regresiva (ROR), versus modelación ARIMA, usando variables dicotómicas en mutaciones del VIH. Universidad Central Marta Abreu de las Villas, 25 de Febrero. Editorial Feijóo.
- [4] Riordan D, Hansen BK. A fuzzy case-based system for weather prediction. *Engineering Intelligent Systems*. 2002; 3-139-46.
- [5] Sordo CM. Técnicas Estadísticas para la Proyección Local de la Predicción Meteorológica Estacional. Métodos, Validación y Estudios De Sensibilidad [Tesis de Doctorado]. Santander, Spain: Universidad de Cantabria; 2006.
- [6] Díaz Y, Fernández A, inventors; CubaForecast para el pronóstico de variables meteorológicas. Cuba. 2003.
- [7] Oses, R, Saura G., Pedraza, A. (2012). Modelación matemática ROR aplicada al pronóstico de terremotos de gran intensidad en Cuba. *REDVET* 2012. Volumen13, No05B. <http://www.veterinaria.org/revistas/redvet/n050512B.html>.
- [8] Box, G. E. P. and Jenkins, G. M., and Reinsel, G. C., 1994: *Time Series Analysis Forecasting and Control*, 3rd edition. Prentice - Hall Inc., New Jersey.
- [9] CITMA. (2012). Resultados del Proyecto 0663, Sistema de pronóstico de la densidad larvaria en Villa Clara para la lucha antivectorial. Ministerio de Ciencia Tecnología y Medio Ambiente Villa Clara.