Geama

Revista GEAMA – Ciências Ambientais e Biotecnologia
Scientific Journal of Environmental Sciences and Biotechnology
ISNN: 2447-0740

6
2

ARTIGO

Models generated by multiple regression in filling meteorological data failures in an automatic meteorological station in Alagoas

Modelos gerados por uma regressão múltipla no preenchimento de falhas de dados meteorológicos em uma estação meteorológica automática em Alagoas

Renato Américo Araújo-Neto[1], Jonathan Willyan dos Santos Nascimento[1], Francisco Freire de Oliveira[2], Gil Rafael Pacífico Rebelo[1], Francisco de Assis de Lima Silva[1], André Luiz de Carvalho[2]

[1] Maurício de Nassau University Center, Alagoas, Brazil
[2] Federal University of Alagoas, Alagoas, Brazil

Contato: renatoaraujoprof@outlook.com

Palavras-Chave
correção
falhas
meteorologia
estatística
regressão

RESUMO
Várias áreas de estudo necessitam de dados meteorológicos. Na ausência deste, metodologias de correções podem ser utilizadas. O objetivo deste estudo foi avaliar o método de regressão múltipla para preencher as falhas das seguintes variáveis meteorológicas: Temperatura Média do Ar (Tmean), Umidade Relativa (RHmean) e Precipitação de Chuva (Prec). A regressão múltipla foi considerada usando diferentes modelos, através dos diferentes cofatores avaliados (variando Tmean, RHmean, Ponto de Orvalho, Pressão e Prec), gerando quatro modelos diferentes de regressão múltipla para cada variável meteorológica estudada. Os modelos foram comparados estatisticamente pelo erro médio absoluto (MAE), coeficiente de Pearson (r), índice de concordância (d) e índice de Camargo e Sentelhas (c). Os resultados apresentados mostraram que a regressão múltipla pode ser usada com segurança em Tmean, RHmean nos Modelos 2, 3 e 4 ($R^2 > 0,90$). A variável Precipitação apresentou coeficiente de determinação abaixo de 50% ($R^2 < 0,50$) e o Modelo 2 obteve um valor de p superior a 1% no Intercept ($p = 0,012$) e no cofator de Pressão ($p = 0,015$). Não pode ser usado para corrigir falhas de chuva. O modelo 2 (exceto Prec) apresentou melhores coeficientes estatísticos e pode ser utilizado para corrigir falhas na estação automática de Maceió, Alagoas.

Key-word
correction
failures
meteorology
statistic
regression

ABSTRACT
Several areas of study require meteorological data. In the absence of this, correction methodologies can be used. The aim of this study was to evaluate the multiple regression method to fill in the gaps in the following meteorological variables: Average Air Temperature (Tmean), Relative Humidity (RHmean) and Rainfall (Prec). Multiple regression was considered using different models, through the different cofactors evaluated (varying Tmean, RHmean, Dew Point, Pressure and Prec), generating four different multiple regression models for each meteorological variable studied. The models were compared statistically by mean absolute error (MAE), Pearson's coefficient (r), agreement index (d) and Camargo and Sentelhas index (c). The results presented showed that multiple regression can be used safely in Tmean, RHmean in Models 2, 3 and 4 ($R^2 > 0.90$). The Precipitation variable showed a coefficient of determination below 50% ($R^2 < 0.50$) and Model 2 obtained a p value greater than 1% in the Intercept ($p = 0.012$) and in the Pressure cofactor ($p = 0.015$). It cannot be used to correct rain flaws. Model 2 (except Prec) showed better statistical coefficients and can be used to correct faults in the automatic station in Maceió, Alagoas.

Informações do artigo
Recebido: 25 de outubro, 2019
Aceito: 10 de agosto, 2020
Publicado: 29 de agosto, 2020

Introduction

Knowledge about meteorological data is of great value, especially in assisting decision making, whether in engineering, construction, agriculture and even in meteorology (BAMBINI; FURTADO, 2010). In engineering, its use can be mainly through the use of forecasting models for the capture of rainwater for non-potable purposes (Martini, 2009). Other examples are illustrated for a wide range of areas, such as rainfall analysis in agricultural crops, oceanography and wind forecast analysis or model formation in evapotranspiration calculations (FUENTES et al., 2013, GIANOTTI et al., 2013, ARAÚJO NETO, 2014).

However, to work with weather data, it is necessary to obtain them. Several Brazilian agencies provide access to meteorological data (such as the National Institute of Meteorology - INMET, the National Water Agency - ANA and the Space Research Institute - INPE). When it comes to data acquisition, failures can occur, mainly caused by the failure of the instrument itself or data transmission, as well as equipment malfunction, equipment shutdown, maintenance, calibration, resulting in decreased reliability of data acquired (TARDIVO; BERTI, 2014). Absence or inconsistency of data may lead to inaccuracy in the analysis and interpretation of results. Several methods are objective of studies to obtain greater accuracy in meteorological data failures, mainly in statistics and geostatistics (VIOLA et al., 2010; BABA; VAZ; COSTA, 2014).

Then, questions arise as to how these flaws can be corrected. Several statistical methods are used to correct missing data, such as the fault-fill method described by Hasan and Croke (2013), using the Poisson-Gamma statistical method to fill faults in rainfall series. Tardivo and Berti (2014) describe regression-based statistics in correcting daily temperature data. However, in the literature, the correction of weather station failures is restricted to nearby stations, with appropriate methodologies for the case and often focused on rainfall data (OLIVEIRA et al., 2010; BIER and FERRAZ, 2017). There is need to study methodologies that seek to correct flaws with data from the weather station itself. Statistical methodologies can fill these gaps, such as the multiple regression method. However, the literature on the subject is still scarce, requiring studies that can generate indicators that correlate the correction of air temperature failures with other factors (e.g. relative humidity, dew point and rainfall).

Ventura et al. (2016) studied fault correction through various statistical methods, including multiple regression. According to the authors, statistical methods can be used accurately to correct faults, especially the arithmetic mean, moving average and linear and multiple regressions. The latter describes that a variable will be dependent on several cofactors, generating indicators for data correction. However, the problem is that many statistical programs focused on the elaboration of linear regressions generate formulas with all cofactors, except for equations that take into account only one or two variables in the generation of fault correction information. Given that, there is a need to generate mathematical models that take into account equations with one, two or more cofactors (e.g. elaboration of fault correction for relative air humidity data that takes into account average air temperature and/or rainfall and/or atmospheric pressure and/or dew point).

Therefore, the present study aims to evaluate the multiple regression method for filling average air temperature (Tmean), Relative Humidity (RHmean), and Rainfall (Prec) faults, by evaluating the different probable multiple models when correlated among them meteorological parameters in an automatic station located in Maceió, Alagoas.

Material and Methods

Analyzing the applicability of fault filling in weather stations, multiple regression was used based on various environmental factors obtained from the weather station database. Multiple regression is basically based on studying the behavior of a dependent variable, correlating it with two or more covariations, according to Equation 1:

$$y = \alpha_0 + \sum_{i=n}^{n} \alpha_i x_i \qquad \text{(Eq.1)}$$

where y is the dependent variable, α_0 the intercept and $\sum_{i=n}^{n} \alpha_i x_i$ are the cofactors.

As dependent variable we obtained results of the following variables: Average Air Temperature (Tmean, in ºC); Average Relative Humidity (RHmean, %) and Rainfall (Prec, mm). The cofactors included in the study varied according to the dependent variable. For example, for Tmean, the cofactors were RHmean; Dew Point, Pressure and Rainfall. For RHmean, the cofactors were Tmean, Dew Point, Pressure and Rainfall, following respectively for the other variables. Multiple regressions were generated using the JASP statistical software (version 0.9.1), which estimated four probable regression models for these variables. The models were evaluated according to the coefficient of determination of equation (R^2) and statistical probability less than 1% ($p < 0.01$).

Meteorological data were obtained by an automatic station of the National Institute of Meteorology (INMET) from 01/01/2019 to 31/03/2019, with a total of 2160 hourly data for each one of the variables (Tmean, RHmean, Dew Point , Pressure and Rainfall), at the weather station located in Maceió, Alagoas (Station Code: OMM: 81998). In the generation of fault indicator data, an algorithm was created in the Excel in order to remove randomly 5% of the values obtained from the automatic station, totaling 108 missing data. This step was providential as there were no missing values in the weather station used to acquire the data.

In comparing the data observed in the automatic station and simulated in the different multiple equations, the following statistical criteria were used:

a) Mean Absolute Error (MAE) - with the MAE, it was possible to evaluate the performance of different models, observing the presence of outliers and data with normal deviation, which influence within the MAE (Moriasi et al., 2007). The MAE was determined by the following equation:

$$EMA = \frac{\sum_{i=1}^{n}|e|}{n} \quad \text{(Eq.2)}$$

where: "e" corresponds to Pi - Oi (subtraction between the data estimated by the models and observed in the Tmed dataset).

b) Pearson correlation coefficient (r) - obtained between the coefficients generated from observed (x-axis) and estimated (y-axis) data from Tmed (MORIASI et al., 2007).

c) Agreement index (d) - this method proposed by Willmott (1981) identifies that the approximation of the estimated data to the observed data can be evaluated by the spacing or approximation of the points, generating an agreement index "d", reflecting the degree of accuracy between the observed and simulated variable. The values may vary from zero, which indicates nullity, to 1, indicating perfect accuracy, being calculated by the following equation:

$$d = 1 - \left[\sum_{i=1}^{N}(P_i - O_i)^2 / \sum_{i=1}^{N}(|P'_i| - |O'_i|)^2\right] \quad \text{(Eq.3)}$$

where N is the number of observations; Pi is the estimated value; Oi is the observed value; P'i is the estimated value, subtracted from the observed average value; O'i is the observed value subtracted from the average value.

d) Camargo and Sentelhas coefficient "c" - in the evaluation of the performance of the estimated data in relation to the observed data, an index "c" was described by Camargo and Sentelhas (1997), related between the product of "r" and "d". Performance values may vary according to the coefficient having very poor performance ("c" equal to or less than 0.40) and excellent performance ("c" greater than 0.85).

With the simulation results, the models were compared with the data observed by the automatic station, aiming to obtain a linear regression, with the forced intercept to be null, generating an angular correction coefficient (Y = b.X).

Results and Discussion

Mean air temperature (Tmean)

The results obtained in the failure filling show that all multiple regression models can be used to determine the average air temperature (Tmean, in ºC), with R^2 ranging from 0.849 to 0.992 (Table 1). It is also observed that in models 2, 3 and 4 the percentage difference between the determination coefficients is around 0.01%, indicating that Tmean can be obtained reliably even if rainfall and pressure data are not available, for example. Data reliability is explained by the probability that it is less than 1% (p <0.001) and the low standard error in all parameters describes the credibility of the data in relation to the generation of different multiple equations.

The statistical standards for the Tmean variable are shown in Table 2. It is observed that model 2 was the one that presented all coefficients above 0.994 Although model 2 presents this behavior, all models had their satisfactory coefficients (above 0.90), indicating that the use of these models will present excellent results in the correction of failures by the multiple regression method. Thus, the model with climate cofactors differs from those with environmental cofactors.

With the parameters of the models, Tmean values were generated in the fault correction, as shown in Figure 1.

Figure 1. Comparison between observed and simulated Tmean data of the different multiple regression models evaluated for an automatic station located in Maceió, Alagoas.

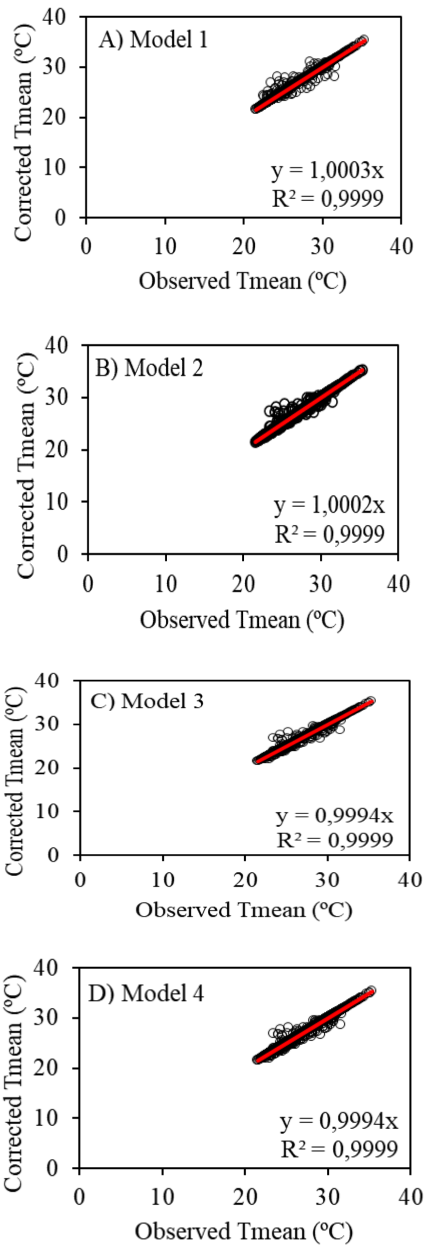


Table 1. Probable models in the multiple regression in the mean air temperature variable for 5% of data failures for the Maceió, Alagoas weather station.

Models		Parameters	Standard Error	r	R^2	p
1	(Intercept)	42.519	0.148	0.849	0.849	< .001
	RHmean	-0.199	0.002			< .001
2	(Intercept)	22.535	0.118	0.991	0.991	< .001
	RHmean	-0.227	4.809e -4			< .001
	Dew Point	0.980	0.005			< .001
3	(Intercept)	58.513	3.328	0.991	0.991	< .001
	RHmean	-0.226	4.763e -4			< .001
	Dew Point	0.965	0.006			< .001
	Pressure	-0.036	0.003			< .001
4	(Intercept)	59.517	3.317	0.992	0.991	< .001
	RHmean	-0.226	4.789e -4			< .001
	Dew Point	0.964	0.005			< .001
	Pressure	-0.037	0.003			< .001
	Rainfall	0.024	0.005			< .001

Fonte: Authors (2019)

All models represented well the multiple regression model in the Tmean fault correction and the statistical adjustment for all were close to 1.00 (R^2>0.97). Regarding the underestimation of the values, models 1 and 2 overestimated the values observed in relation to the corrected values by 0.03 and 0.02%, respectively. Models 3 and 4 underestimated the observed data by 0.06% both. This indicates that all models can be used to correct Tmean faults and the model 2 is more likely to estimate that faults with greater reliability. The high value of r and R^2 and Pearson's coefficient prove the effectiveness of multiple regression in fault fill for the Maceió automatic station (Table 1; Table 2). This efficacy is proven in fault fill described by Ventura et al. (2016), who presented a Pearson coefficient greater than 0.86 in temperature fault corrections in three Brazilian state capitals. According to the authors, multiple regression can be used effectively to correct failures.

Regarding the MAE, all models present satisfactory value (VENTURA et al., 2016). Bier and Ferraz (2017), for example, present several methodologies for correcting temperature compensated faults between weather stations, presenting low errors in corrections. According to the authors, the results show that it is possible to generate estimates for monthly data from statistical methods on different meteorological variables.

Table 2. Statistical patterns in the comparison between observed and simulated data for failure of 5% of average air temperature weather data at an automatic station located in Maceió, Alagoas.

	Model 1	Model 2	Model 3	Model 4
MAE	0.0479	0.0403	0.0477	0.0481
r Pearson	0.9937	0.9947	0.9947	0.9947
d Willmott	0.9968	0.9974	0.9973	0.9973
c	0.9905	0.9921	0.9921	0.9920

Fonte: Authors (2019)

However, it is of fundamental importance to remember that the methodology applied in our study presents corrections within the automatic station itself, indicating the possibility of using statistical methods (with mean, moving average, linear regression or multiple regression) in the correction of failures when there is no information from nearby automated stations. This evidence is presented in Figure 1 when the simulated data by the models in the same automatic station are contrasted, presenting R^2 higher than 0.97, indicating the excellent correlation between the simulated and observed data, especially in model 2, which takes into account RHmean and Dew Point data.

Average Relative Humidity (RHmean)

The statistical standards for RHmean are presented in Table 3.

Table 3. Probable models in the multiple regression in the mean relative humidity variable, for 5% of data failures for the Maceió, Alagoas weather station.

Models		Parâmeters	Standard Error	r	R^2	p
1	(Intercept)	42.519	0.148	0.849	0.849	< .001
	RHmean	-0.199	0.002			< .001
2	(Intercept)	22.535	0.118	0.991	0.991	< .001
	RHmean	-0.227	4.809e-4			< .001
	Dew Point	0.980	0.005			< .001
3	(Intercept)	58.513	3.328	0.991	0.991	< .001
	RHmean	-0.226	4.763e-4			< .001
	Dew Point	0.965	0.006			< .001
	Pressure	-0.036	0.003			< .001
4	(Intercept)	59.517	3.317	0.992	0.991	< .001
	RHmean	-0.226	4.789e-4			< .001
	Dew Point	0.964	0.005			< .001
	Pressure	-0.037	0.003			< .001
	Rainfall	0.024	0.005			< .001

Fonte: Authors (2019)

Correlation and determination coefficients in all models showed high reliability of the parameters, indicating their use in the correction of RHmean faults. Models 2, 3 and 4 presented the same R^2 (0.922), indicating the accuracy of the parameters in the correction of the data, and models that do not have rainfall and pressure data, for example.

The statistical patterns in the Average Relative Humidity variable are presented in Table 4.

Like the average temperature, model 2 was more satisfactory in relation to the others, indicating that with only the variables Tmean and Dew point is capable of get missing data for RHmean.

Nevertheless, the other models can be reliably used to determine missing data.

These statistical patterns are of fundamental importance in the elaboration of numerical indicators.

Araújo Neto et al. (2015) elaborated different multivariate regression models in the making of climate maps in the state of Alagoas, helping in the adjustment of planting of several agricultural crops.

Adjustment of the parameters was of fundamental importance in the generation of missing RHmean data, presented in Figure 2.

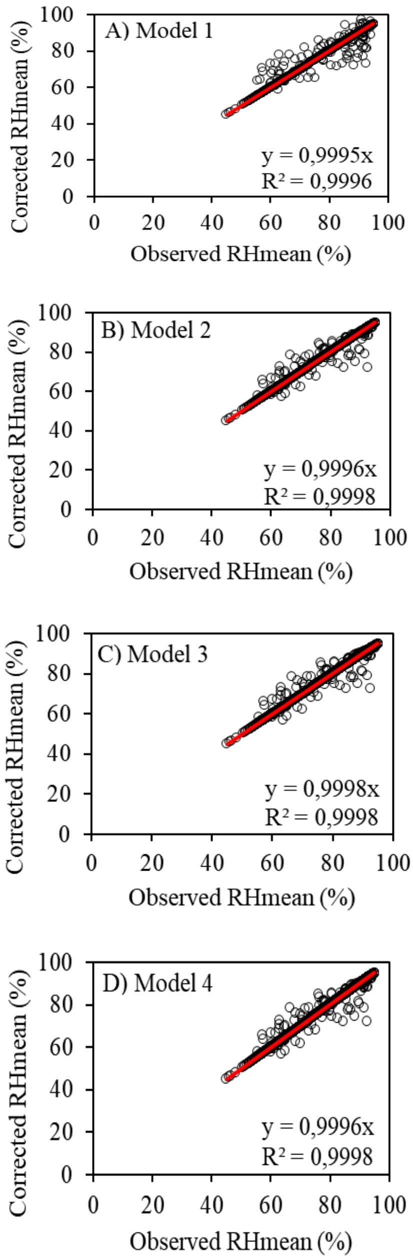
All models presented R^2 greater than 0.97, indicating with adjustment of multiple regression models in the correction of RHmean data. Unlike what was observed in the Tmed data (Figure 1), the models underestimated the observed data, ranging between 0.02 and 0.05%.

Table 4. Statistical patterns in the comparison between observed and simulated data for failure of 5% of average Relative Humidity weather data at an automatic station located in Maceió, Alagoas.

	Model 1	Model 2	Model 3	Model 4
MAE	0.0479	0.0403	0.0477	0.0481
r Pearson	0.9937	0.9947	0.9947	0.9947
d Willmott	0.9968	0.9974	0.9973	0.9973
c	0.9905	0.9921	0.9921	0.9920

Fonte:Authors (2019)

Figure 2. Comparison between observed and simulated RHmed data from different multiple regression models evaluated for an automatic station located in Maceió, Alagoas.



Rainfall (Prec)

The mathematical adjustments in the correction representation of missing rainfall data (Prec) were presented in Table 5. It is observed that the p value was greater than 5% in variables of models 2 and 3 making the use of these adjustments impossible to obtain rainfall data. Although models 1 and 4 present satisfactory probability (p <0.001), the determination coefficients indicate that less than 5% of the data were representative in the adoption of the multiple regression mathematical method. This indicates that even with climate data, errors can occur in correcting rainfall data.

Statistical standards are presented in Table 6. Although all models have satisfactory statistical standards, the models cannot be reliably used due to the low adjustment of variables in multiple regression (Table 5). The bad behavior of the mathematical adjustment about the rainfall can be explained due to this environmental variable is correlated with climatic variables, thereby decreasing the accuracy and reliability of the data. Another factor that may be interfering with this low adjustment may be the time when the data were collected (period with little rainy season). One solution to this variable is the adoption of other established methodologies for rainfall, as described by Bier and Ferraz (2017) and Ottero; Chargel; Hora. (2018).

When correlating the values observed in the automatic station and simulated by multiple regression (Figure 3), it is observed that all models underestimated the observed data, with a satisfactory determination coefficient (R^2 close to 1). However, despite presenting consistent data in relation to the statistics, when the models were adjusted to the Prec variable, some estimated values were below 0 mm, indicating the imprecision of the models, as shown in Table 5. We found that despite the R^2 adjustment has been satisfactory, inconsistencies may occur, such as: I) inaccuracy of the simulated precipitation values, with negative results; II) the low values of real precipitation (many close to 0 mm) can result in these satisfactory results (R^2 close to 0.99) described in Figure 3; III) non-significant results (p> 0.05) described in Table 5 confirm that the precipitation variable cannot be simulated through a multiple regression. Thus, precipitation adjustment through multiple regression cannot be adopted due to inconsistent results and inaccurate values simulated by the models.

Table 6. Statistical patterns in the comparison between observed and simulated data for failure of 5% of rainfall weather data at an automatic station located in Maceió, Alagoas.

	Model 1	Model 2	Model 3	Model 4
MAE	0.0124	0.0095	0.0301	0.0171
r pearson	0.9979	0.9983	0.9912	0.9963
d Willmott	0.9989	0.9991	0.9954	0.9981
c	0.9969	0.9974	0.9866	0.9945

Fonte: Authors (2019)

Figure 3. Comparison between observed and simulated rainfall data from different multiple regression models evaluated for an automatic station located in Maceió, Alagoas.

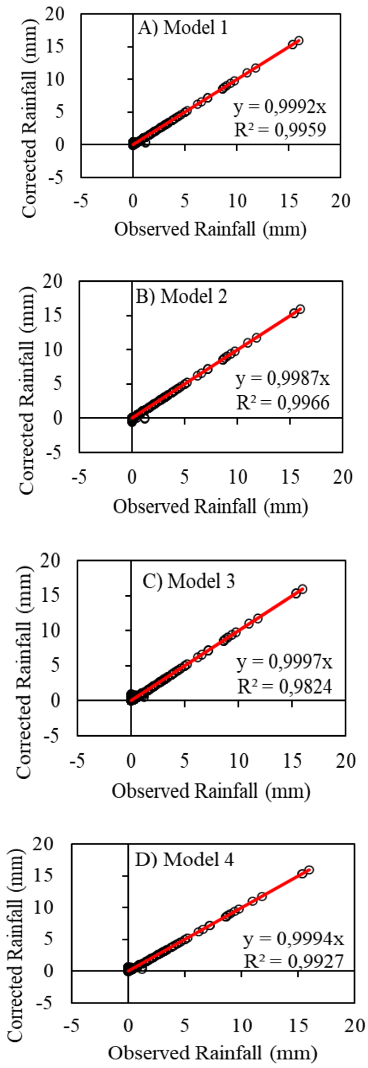


Table 5. Probable models in the multiple regression in the rainfall variable for 5% of data failures for the Maceió, Alagoas weather station.

Models		Parâmeters	Standard Error	r	R^2	p
1	(Intercept)	-1.038	0.157	0.030	0.030	< .001
	RHmean	0.016	0.002			< .001
2	(Intercept)	-35.465	14.128	0.033	0.030	*0.012*
	RHmean	0.015	0.002			< .001
	Pressure	0.034	0.014			*0.015*
3	(Intercept)	-50.613	15.124	0.037	0.035	< .001
	RHmean	0.028	0.005			< .001
	Pressure	0.047	0.015			0.002
	Tmean	0.068	0.024			*0.006*
4	(Intercept)	-69.472	15.713	0.045	0.043	< .001
	RHmean	0.119	0.022			< .001
	Pressure	0.057	0.015			< .001
	Tmean	0.465	0.097			< .001
	Dew Point	-0.409	0.097			< .001

Fonte: Authors (2019)

Conclusions

The correction of mean temperature and air humidity failures can be performed through all models generated by multiple regression, using model 2 which presented the best statistical coefficients.

The rainfall variable cannot be estimated through the multiple regression model. Even generating values close to those observed in the automatic station, the statistical indexes indicate that the models cannot generate reliable data for this variable.

Although the model has not adjusted to the rainfall variable, studies can be performed with this variable in rainy seasons, aiming to evaluate the viability of multiple regression in relation to rainfall. (2)

Acknowledgments

The authors thank the Mauricio Nassau University Center for encouraging research.

References

Araújo Neto, R.A. **Métodos para determinar a evapotranspiração de cana-de-açúcar**, 2013, 73f, Dissertação (Mestrado em Agronomia) – Universidade Federal de Alagoas, Rio Largo, Alagoas, 2013.

Araújo Neto, R. A.; Lyra, G. B.; Ferreira Junior, R. A.; Magalhães, I. D. Sistemas de informações geográficas aplicados ao Índice de Satisfação da Necessidade de Água para cultura do milho em Alagoas estimado por Thornthwaite. **XIX Congresso Brasileiro de Agrometeorologia**. Lavras – MG – Brasil, p.1-5, 2015.

Baba, R. K.; Vaz, M. S. M. G.; Costa, J.. Correção de dados agrometeorológicos utilizando métodos estatísticos. **Revista Brasileira de Meteorologia**, v.29, n.4, p.515-526, 2014.

Bambini, M, D,; Furtado, A, T, Redes observação e a evolução tecnológica contribuindo para o desenvolvimento de modelos matemáticos na Meteorologia no século XX In: Seminário Nacional de História da Ciência e da Tecnologia, 12,; **Congresso Latino-Americano De História Da Ciência E Da Tecnologia**, 7,, 2010, Salvador, Anais,,, [S,l,]: Sociedade Brasileira da História da Ciência, 2010,

Bier, A,A,; Ferraz, S,E,T, Comparação de Metodologias de Preenchimento de Falhas em Dados Meteorológicos para Estações no Sul do Brasil, **Revista Brasileira de Meteorologia**, v, 32, p,215-226, 2017,

Camargo, A,P, de; Sentelhas, P,C, Avaliação do desempenho de diferentes métodos de estimativa da evapotranspiração potencial no Estado de São Paulo, Brasil, **Revista Brasileira de Agrometeorologia**, v,5, p,89-97, 1997,

Fuentes, E, V,; Bitencourt, D, P,; Fuentes, M, V, Análise da velocidade do vento e altura de onda em incidentes de naufrágio na costa brasileira entre os estados do Sergipe e do Rio Grande Do Sul, **Revista Brasileira de Meteorologia**, V, 28, Nº 3, 257-266, 2013,

Gianotti, A, R, C,; Souza, M, J, H,; Machado, E, L, M,; Pereira, I, M,; Vieira, A, D,; Magalhães, M, R, Análise microclimática em duas fitofisionomias do cerrado no alto vale do Jequitinhonha, Minas Gerais, **Revista Brasileira de Meteorologia**, V, 28, Nº 3, 246 - 256, 2013,

Hasan, M, M,; Croke, B, F, W, Filling gaps in daily rainfall data: a statistical approach, In: **International Congress on Modelling and Simulation**, 20, 2013,

Martini, F, **Potencial de economia de água potável por meio do uso de água de chuva em São Miguel do Oeste – SC**, 2009, 96f, Trabalho de conclusão de curso (curso de engenharia civil) – Universidade Federal de Santa Catarina, Florianopolis, 2009,

Moriasi, D. N, et al, Model evaluation guidelines for systemiatic quantification of accuracy in watershed simulations, **Amer, Soc, of Agric, and Bio, Eng**, v,50, n,3, p,885-900, 2007.

Oliveira, L. F. C.; Fioreze, A. P.; Medeiros A. M. M.; Silva, M. A. S. Comparação de metodologias de preenchimento de falhas de séries históricas de precipitação pluvial anual. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.14, n.11, p.1186–1192, 2010.

Ottero, C. R.; Changel, L. T.; Hora, M. A. G. M. Análise de Frequência dos dados pluviométricos observados em 2011 e 2013 na região serrana, estado do Rio de Janeiro. **Revista Brasileira de Meteorologia**. v.33, n.1, p.131-139, 2018.

Tardivo, G.; Berti, A, The selection of predictors in a regression-based method for gap filling in daily temperature datasets, **Int, J, Climatol**,, 34: 1311–1317, JUN/2013,

Ventura, T, M, et al, Análise da aplicabilidade de métodos estatísticos para preenchimento de falhas em dados meteorológicos, Revista Brasileira de Climatologia, v,19, p,168-177, 2016,

Viola, M. R. et al. Métodos de interpolação espacial para o mapeamento da precipitação pluvial. **Revista Brasileira de Engenharia Agrícola e Ambiental**. v.14, n.9, p.970–978, 2010.

Willmott, C,J, On validation of models, **Physical Geography**, v,2, p,184-194, 1981.