Original Research Article

Homogeneity evaluation of historical rainfall and temperature series in Mato Grosso

ABSTRACT

The homogeneity investigation of a series can be performed through several nonparametric statistical tests, which serve to detect artificial changes or non-homogeneities in climatic variables. The objective of this work was to evaluate two methodologies to verify the homogeneity of the historical climatological series of precipitation and temperature in Mato Grosso state. The series homogeneity evaluation was performed using the following non-parametric tests: Wald-Wolfowitz (for series with one or no interruption), Kruskal-Wallis (for series with two or more interruptions), and Mann-Kendall (for time series trend analysis). The results of the precipitation series homogeneity analysis from the National Waters Agency stations, analyzed by the Kruskal-Wallis and Wald-Wolfowitz tests, presented 61.54% of homogeneous stations, being well distributed throughout Mato Grosso state, whereas those of the trend analysis allowed to identify that 87.57% of the rainfall-gauging stations showed a concentrated positive trend, mainly in the rainy season. Out of the conventional stations of the National Institute of Meteorology of Mato Grosso, seven were homogeneous for the precipitation variable, five for maximum temperature and four stations were homogeneous for minimum temperature. For the trend analysis in the 11 stations, positive trends of random nature were observed, suggesting increasing alterations in the analyzed variables. Therefore, the trend analysis performed by the Mann-Kendall test in the precipitation, and maximum and minimum temperature climate series, indicated that several data series showed increasing trends, suggesting a possible increase in precipitation and temperature values over the years. The results of the Kruskal-Wallis and Wald-Wolfowitz tests for homogeneity presented more than 87% of homogeneous stations.

Keywords: Mannn-Kendell, Kruskal Wallis, Wald-Wolfowitz.

1. INTRODUCTION

Information concerning climatic elements is of great importance in the various activities developed by man, such as agriculture, through crop and zonation forecasts, management of water resources, and in climatic studies related to atmospheric phenomena. This is only possible because climatic data provide much information about the atmospheric environment; however, in order to make use of historical climatological series, it is essential to identify possible changes in meteorological records. Therefore, in order to guarantee the reliability of climate studies, it is necessary to employ reliable data whose homogeneity has been verified, since the chances of contradictory and misleading conclusions increase when using non-homogeneous data in analyses.

The main issues found in historical series are the difficulties in obtaining meteorological data with long and reliable time series, and the faults (discontinuities in the series). It is important

to emphasize that the occurrence of these faults may compromise analysis and data interpretation. However, according to [1], the faults (interruptions) occurred in climatological series do not render them unfeasible, although it is not possible to estimate the missing data without changing the frequency distribution dispersion scale [2].

Therefore, in order to make use of climatic data it is necessary to verify if they are statistically homogeneous. However, there is still a shortage in studies that analyze such homogeneity in all meteorological elements in Mato Grosso. [4, 5] argue that the non-homogeneity in climatic series may have a different origin, such as vegetation growth and / or urbanization in the vicinity of the stations, changes in location or de-calibration in measuring instruments, and errors during instrument reading .

The verification of a series homogeneity may be performed through several nonparametric statistical tests that lend themselves to allow detecting artificial changes or non-homogeneities in climatic variables. The consistency evaluation of the series can be performed using the Wald-Wolfowitz and Kruskal-Wallis non-parametric tests, depending or not on the occurrence of interruptions in the data series [1], and using the Mann-Kendall test for the analysis of possible trends in all the time series.

The Wald-Wolfowitz and Kruskal-Wallis non-parametric tests are a more traditional way of analyzing homogeneity, although more recently another method has been proposed, which is the Mann-Kendall trend analysis, aiming to detect possible temporal trends.

In this way, two methodologies that can present contrasting information about climatological series were employed. Therefore, the aim of this work was to evaluate these two methodologies to verify the homogeneity of the historical climatological series of precipitation and temperature in Mato Grosso state.

2. MATERIAL AND METHODS

The daily temperature (maximum and minimum) and the precipitation data were obtained from the National Institute of Meteorology (INMET), and the remaining analyzed precipitation data were obtained from National Waters Agency (ANA). Data series with a minimum of 10 years of observation were employed, totaling 10 conventional INMET stations and 169 ANA rainfall-gauging stations.

The data were organized into ten-day periods to verify their homogeneity, considering that with this ten-day grouping it is possible to detect variations in the analyzed series more easily. In this way, the annual data, including those from leap years, were standardized into 36 periods. In addition, no bug filling was performed, so there would be no interference to the data sets.

Two methodologies for data analysis (Wald-Wolfowitz and Kruskal-Wallis) were employed to analyze the climatological series: one that verifies data set homogeneity, and another that examines trend occurrence within the series (Mann-Kendall). For both methodologies, the significance level of 1% was adopted, and the SPSS statistical software was utilized.

To verify the homogeneity, non-parametric tests (Wald-Wolfowitz and Kruskal-Wallis) were used. When the series showed no interruption, its homogeneity was verified by the Wald-Wolfowitz one-sample runs test. This test consists in determining the series median and then comparing the number of values sequences above or below the median in the chronological

order of the observations, with the expected theoretical value with the same degree of freedom.

If the data presented one interruption, the Wald-Wolfowitz two-sample unilateral test of iterations was applied. This test is employable when it is desired to prove the null hypothesis that two samples have been extracted from the same population, against the alternative hypothesis that the two groups differ in any way [5].

For series that presented two or more interruptions, the Kruskal-Wallis test was applied to test if the samples sets came from the same distribution, that is, to test the null hypothesis that all the time series have equal distribution functions, against the alternative hypothesis that at least two of the time series have different distribution functions [6].

The historical series trend analysis was verified using the Mann-Kendall test, which is a sequential and non-parametric method that was employed to determine if the data series had a statistically significant change in temporal trend.

As described by [7], according to the MK test, each value $Yi, i = 1, \dots, n-1$ is compared with all following values $Yj, j = i+1, i+2, \dots, n$, generating a new Zi series that contains an indicator of the relative value of the difference between terms of the Yi series, according to:

$$Z_{i=}\ signal\left(Y_i - Y_j\right) = \begin{cases} 1\ if\ Y_i > Y_j \\ 0\ if\ Y_i = Y_j \\ -1\ if\ Y_i < Y_j \end{cases} \quad (1)$$

The S statistic is then calculated using the following equation:

$$S = \sum_{i=2}^{n} \sum_{j=1}^{i=1} signal\ (Y_i - Y_j) \quad (2)$$

The variance is defined by:

$$VAR\ (S) = \frac{1}{18}\ [n\ (n-1)(2n+5)] \quad (3)$$

Where *n* is the time series size.

However, the significance of S for the null hypothesis can be tested by using a bilateral test, and can be rejected for large values of the Z (t) statistic, defined according to:

$$Z(t) = \begin{cases} \frac{S-1}{\sqrt{Var(S)}}\ if\ S > 0 \\ 0 \qquad if\ S = 0 \\ \frac{S+1}{\sqrt{Var(S)}}\ if\ S < 0 \end{cases} \quad (4)$$

A positive Z value indicates growth tendency; a negative Z indicates a downward trend, and a large magnitude of the Z value indicates that the trend is strongly significant.

3. RESULTS AND DISCUSSION

When analyzing the historical series of ANA rainfall-gauging stations through the Kruskal-Wallis and Wald-Wolfowitz tests, it was observed that out of the 169 analyzed stations, 104 presented homogeneity in all their ten-day periods, according to Table 1. The homogeneity verified by these tests indicates that the data series present variations only as a function of the climate conditions. In the case of non-homogeneity, it may be associated with non-climatic factors, which make these series little representative of the climate variation, in such a manner that the absence of homogeneity in the series might occur due to alterations caused by vegetation growth or poor maintenance of the measuring instruments, and even by observation habits [3]. For the Mann-Kendall test, only 21 stations with no trend in its ten-day periods were observed, as may be seen in Table 2.

Table 1: Relation of ANA rainfall-gauging stations with homogeneous series in Mato Grosso until 2016.

01154000	01452004	01554001	01058004
01455009	00956001	01554005	01455004
01456010	00957001	01251002	01351000
01455011	01653005	01558000	01055002
01455010	01060001	01455008	01055003
01454003	01059000	01557005	01360003
01459003	01058005	01655001	01359001
00957002	01755003	01555004	00958004
00956002	01756001	01655003	00958002
01354001	01457001	01150001	01058002
01454000	01552006	01152001	01556000
01654004	01552002	01052000	01555000
01055000	01251000	01358005	01254002
01656001	01257000	01155000	01255002
01757001	01258001	01255001	01553003
01656004	01158003	01156000	01552001
01657001	01657002	01457004	01654005
01652001	01658000	01055004	01653002
01052001	01557004	01154002	01755000
01157000	01757002	01254001	01157001
01157002	01353001	01659001	01057000
01156001	01358001	01560000	01058006
01558003	01357001	01559006	01159000
01558005	01555005	00951000	01259001
01557003	01158004	01452000	01058003
01054000	01050000	01150006	01256002

Table 2: List of ANA rainfall-gauging stations with no-trend series in Mato Grosso until 2016.

01753000	01055001	01552002	01154000	01656003
01653004	01555000	01353001	01354000	01756001
00857000	01255002	01352002	01251001	01158001
01755003	01553003	01150001	01655001	01652002
01159000				

By observing both tests results, it may be seen that the Mann-Kendall test found that more than 87% of the stations present a statistically significant temporal change trend, either positive or negative. The homogeneity verification method detected only 38% of non-homogeneous stations, that is, there was not much concordance in the results of both tests due to the fact that the Mann-Kendall test is a much more robust and rigorous analysis method, which in many situations is used to try to identify possible climate changes [8].

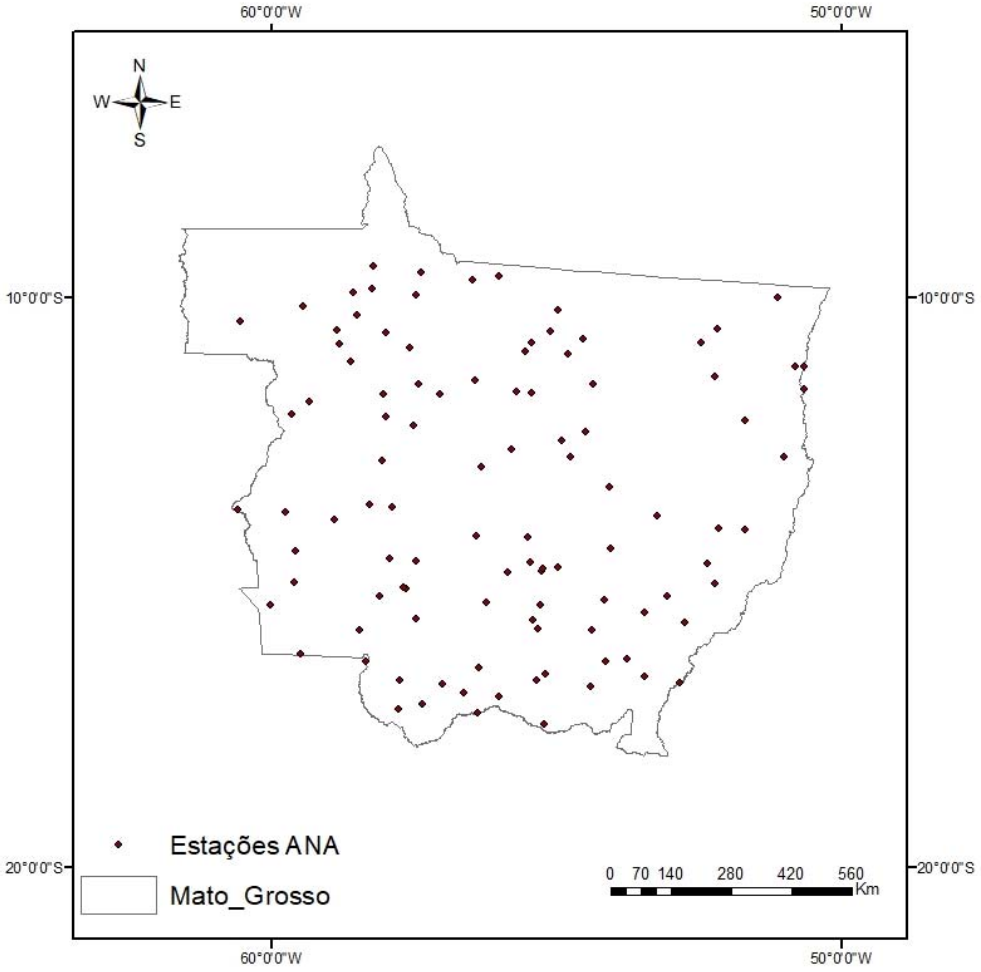
The Mann-Kendall test revealed a rainfall increase in most of the analyzed stations, since most of the ten-day periods that presented a positive trend were concentrated in the rainy season, what could mean that there was a significant increase in precipitation over the years in those seasons. A similar behavior was obtained by [8], in which a positive trend was identified in the rainfall series analysis in Western Amazonia in the rainy season (January to April).

Authors such as [9, 10, 8] used the trend analysis to verify climatic variability in historical series, which is an advantage of this analysis, since it allows to observe changes in the series behavior and to determine which regions are suffering significant variations over time. Thus, it may be considered that the Mann-Kendall trend analysis test is employed to identify climate change occurrence.

Therefore, it is important to emphasize that the study of rainfall behavior allows to detect trends or changes in climate at local or regional scales and, with due understanding, it becomes an analysis element in the organization of territorial and environmental planning due the high interference degree, impact and repercussion in time and space [11]. [4] emphasize the importance of data reliability in variation analyses of air temperature and precipitation for the current discussion on climate change. However, it is worth emphasizing that such changes are not necessarily the result of anthropization. They may be just a natural process that has been occurring with the Earth, such as the intensification of solar activity and natural phenomena such as the El Niño and La Niña [12].

It is observed in Figure 1 that the stations that presented homogeneity by the Kruskal-Wallis and Wald-Wolfowitz tests are distributed throughout Mato Grosso state. In the Mann-Kendall test result, stations that did not present a trend, in addition to their reduced number, are not distributed throughout the state, according to Figure 2.

The occurrence of non-homogeneity and / or discontinuities in climatological time series may interfere in the climatic variability characterization of a locality. This non-homogeneity may occur due to several non-climatic factors, such as vegetation growth or urbanization near the stations, also by a change in location or de-calibration in measuring instruments and even by observation habits [3, 4].



162

163 **Figure 1: Relation of ANA rainfall-gauging stations with homogeneous series,**
164 **analyzed by Kruskal-Wallis and Wald-Wolfowitz tests in Mato Grosso until 2016.**

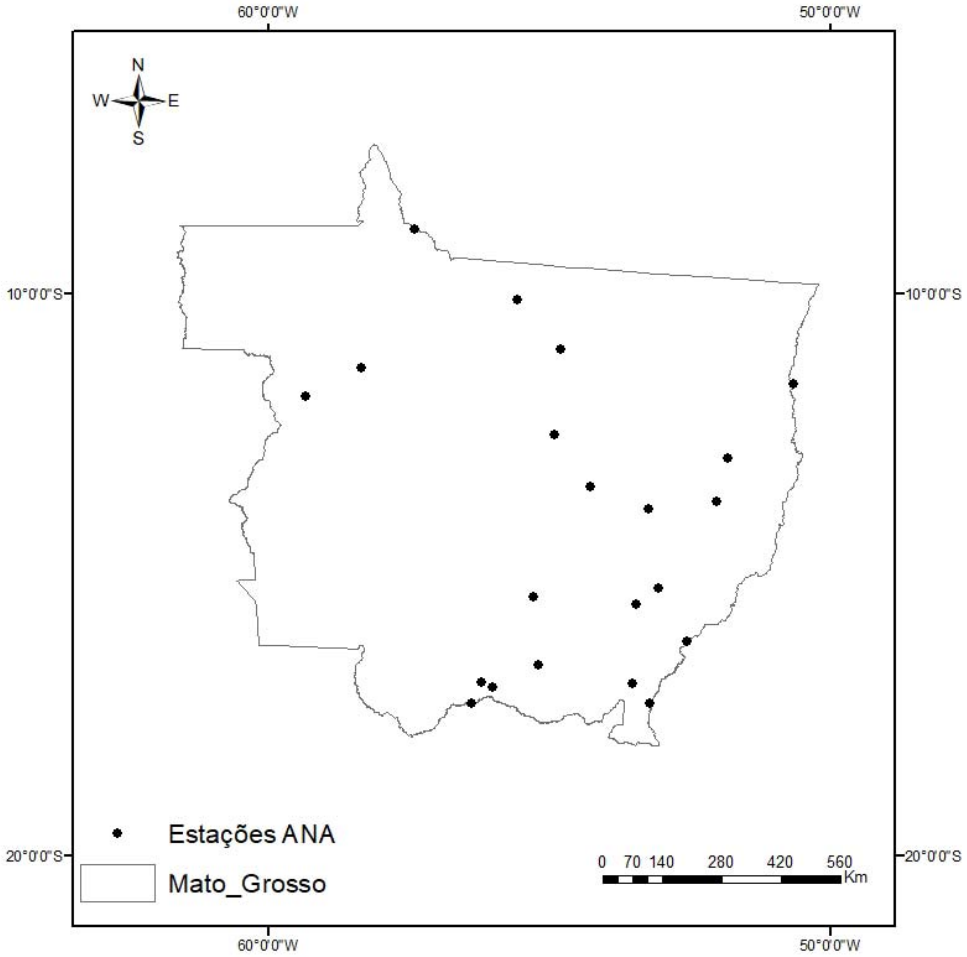


Figure 2: Relation of ANA rainfall-gauging stations with series without trend, analyzed by the Mann-Kendall test in Mato Grosso until 2016.

Out of the conventional INMET stations of Mato Grosso, seven were homogeneous for the precipitation variable, five for maximum temperature, and four homogeneous stations for minimum temperature (Table 3). The nature of this inconsistency in the analyzed series was of random origin, and according to [1], it might also be of instrumental origin or even due to specific conditions of measurement.

For the trend analysis in the 11 stations, positive trends were observed in at least one ten-day period. The trends observed in the INMET stations were random for the three analyzed variables, unlike what may be verified with the ANA rainfall stations that concentrated the trends in the rainy season, that is, in the case of the maximum and minimum temperature, no relation was verified between significant changes in a particular season. The same behavior was identified in a study conducted by [9] in Minas Gerais state, where even in neighboring and highly correlated stations, there was generally no satisfactory agreement between the significant points of change.

Overall, the minimum air temperature series presented more trend than the maximum air temperature. [12, 9] also verified this pattern when analyzing monthly air temperature series in São Paulo and Minas Gerais, respectively. More increase than decrease trends were observed, suggesting an increase in the values of maximum temperature, minimum temperature and rainfall over the years.

Table 3: Relation of conventional INMET stations with homogeneous series in Mato Grosso state until 2016, for precipitation, maximum temperature and minimum temperature variables.

Precipitation	Maximum temperature	Minimum temperature
Cáceres	Canarana	Cáceres
Cuiabá	Matupá	Canarana
Gleba Celeste	Padre Ricardo Remetter	Padre Ricardo Remetter
Matupá	Rondonópolis	São José do Rio Claro
Padre Ricardo Remetter	São José do Rio Claro	
Rondonópolis		
São José do Rio Claro		

The results from the homogeneity analysis of the climatological series obtained in this work are useful for guidance on critical issues to agriculture, such as the choice of the appropriate crop for a given locality, the best sowing season, and the agricultural planning in general, since these series are useful to characterize the weather and the climatic conditions of the region.

4. CONCLUSION

The Mann-Kendall trend analysis, in the climatic series of precipitation and maximum and minimum temperature, suggested that several data series showed increasing trends, indicating a possible increase in precipitation and temperature values over the years. Moreover, the results of the Kruskal-Wallis and Wald-Wolfowitz tests for homogeneity resulted in more than 87% of homogeneous stations.

REFERENCES

1.CAMPELO JÚNIOR, J. H. Duration, Homogeneity and Spatial Distribution of Precipitation Series in Mato Grosso. Brazilian Journal of Agrometeorology, Santa Maria, v.1, p.31-33, 1993.

2. THOM, H. C. S. Some Methods of Climatological Analysis. Geneva: WMO, Techinal Note, n. 81, 53 pp., 1966.

3. AGUILAR, E .; AUER, I .; BRUNET, M .; PETERSON, T. C .; WIERINGA, J. Guidelines on climate metadata and homogenization. World climate program data and monitoring n.53. WMO-TD No. 1186, Geneva: WMO, 2003. 52 p.

4. BEGERT, M .; SCHLEGEL, T .; KIRCHHOFER, W. Homogeneous temperature and precipitation series from Switzerland from 1984 to 200. International Journal of Climatology. v. 25, p. 65-80, 2005.

5. SIEGEL, S. Non-parametric statistics. Publisher McGraw Hill of Brazil, São Paulo, 1981, 350 p.

6. TRIOLA, M. F. Introduction to statistics. 9 ed. Rio de Janeiro: LTC, 2005, 656 p.

7. SANTOS, J. F.; PORTELA, M. M. Trends in monthly rainfall series in mainland Portugal. Application of the Mann-Kendall test. Portuguese Association of Water Resources, Lisbon, 13 p. 2007.

Available:https://comum.rcaap.pt/bitstream/10400.26/1303/1/Artigo_8SILUSBA_JFSantos_MMPortela_2007.pdf

8. SALVIANO, M. F.; GROPPO, J.D.; PELLEGRINO, G. Q. Trend Analysis on Precipitation and Temperature Data in Brazil. Brazilian Journal of Meteorology, São José dos Campos, v. 31, n. 1, p. 64-73, 2016.

9. SANTOS, R. S .; SEDIYAMA, G. C .; OLIVEIRA, R. A .; ABRAHÃO, G. M. Homogeneity of climatological series in Minas Gerais. Brazilian Journal of Agricultural and Environmental Engineering (Online), v. 16, p. 1338-1345, 2012.

10. COSTA, V. J .; FERREIRA, M .; CORDEIRO, M. T. A. Analysis of climatic time series. Revista de Ciências Agroveterinárias, Lages, v.14, n.2, p.169-177, 2015.

11. SANT'ANNA NETO, J. L. Rainfall in the state of São Paulo: rainfall variability in the last 100 years. In: SANT'ANNA NETO, J. L; ZAVATINI, J.A. (ORGs). Variability and climate change: Environmental and socioeconomic implications. Maringá, EDUEM, p. 17-28, 2000.

12. BLAIN, G. C. Detection of monotonous trends in monthly rainfall series of the state of São Paulo. Bragantia, Campinas, v. 69, n. 4, p. 1027-1033, 2010.