MedCrave
Step into the World of Research

International Journal of Hydrology

Review Article

Open Access

CrossMark

Identification of atypical data in series of instantaneous maximum flows by means of box plots

Abstract

In this technical note we analyzed the data of the instantaneous maximum flows reported in the BANDAS for eight gauging stations belonging to the hydrological regions 1 to 3 of Mexico.[1] which are part of the 37 Hydrological Regions (HR) of Mexico established by CONAGUA for the purpose of study (Figure 1). For the analysis, box and whisker plots were used that take into account the quantiles 1, 2 and 3 (Q1, Q2 and Q3), placing the north-south stations of 20 atypical data, 8 possible cases of data were identified that could be related to real changes in the pattern of behavior of the series or to true measurement errors.

Volume 2 Issue 4 - 2018

Arganis Juárez ML,[1,2] Domínguez Mora R,[1] Carrizosa Elizondo E,[1] Preciado Jiménez M[3]

[1]Institute of Engineering, National Autonomous University of Mexico, Mexico

[2]Faculty of Engineering, National Autonomous University of Mexico, Mexico

[3]Mexican Institute of Water Technology, Mexico

Correspondence: Arganis Juárez ML, Institute of Engineering, National Autonomous University of Mexico, Copilco CP. 04510 Cdmx, Mexico, Email MArganisJ@iingen.unam.mx

Received: June 26, 2018 | **Published:** July 03, 2018

Introduction

The analysis of hydroclimatological data or time series data for hydrological studies requires statistical and graphic tools to facilitate the identification of data that may involve atypical climatic phenomena, events from seasonal hydrometeorological phenomena or true errors in measurement or captures of the official databases such as the National Surface Water Data Bank,[2] of the National Water Commission (CONAGUA), official agency in Mexico responsible for the collection of information on water and climate.

In this technical note we analyzed the data of the instantaneous maximum flows reported in the BANDAS for eight gauging stations belonging to the hydrological regions 1 to 3 of Mexico,[1] which are part of the 37 Hydrological Regions (HR) of Mexico established by CONAGUA for the purpose of study (Figure 1). For the analysis, box and whisker plots were used that take into account the quantiles 1, 2 and 3 (Q_1, Q_2 and Q_3), placing the north-south stations.

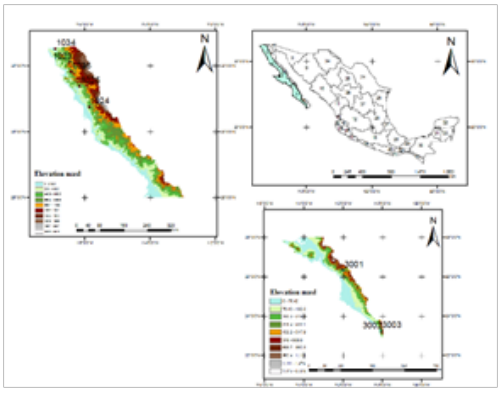

Figure 1 Gauging stations of HR 1 to 3, Mexico.

Methodology

The Box plot (Figure 2) was developed by Tukey[3,4] to evaluate the shape of the distributions, since they allow detecting problems in the tails of the distribution (extreme values, whether they are atypical or errors). These values can distort subsequent analysis, producing inadequate adjustments of the models to be used.

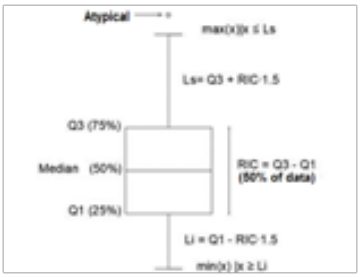

Figure 2 Box plot scheme.

Application

Figure 3 shows the box plots graphs obtained for all analyzed stations. We added to these graphs the mean of the instantaneous maximum flows of each station in order to find the outliers, that is to say that they exceed the upper or bottom limits of the graphs. A review was made of the dates in which an atypical data was found to identify if they were associated with a hydrometeorological extraordinary event or with a possible true error of capture or measurement.

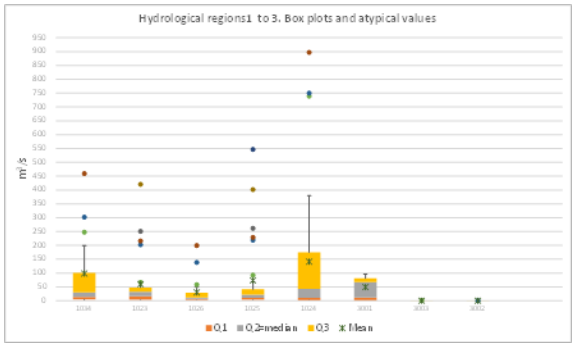

Figure 3 Box plots RH 1 to 3, Mexico.

Conclusion

Of 20 atypical data, 8 possible cases of data were identified that could be related to real changes in the pattern of behavior of the series or to true measurement errors. It is also observed in Figure 3 the decreasing or increasing behavior of the runoff with the geographical position of the stations, but to support this conclusion another graph must be used that relates the average maximum flow per unit area with the location.

Acknowledgements

None.

Conflict of interests

The author declare there is no conflict of interest.

References

1. Domínguez-Mora. *Tomo IV*. México: Regionalización de gastos. Informe Técnico para CENAPRED; 2017.
2. BANDAS. Banco Nacional De Datos De Aguas Superficiales. *CONAGUA*. 2018.
3. Tukey JW. Exploratory data analysis. *Addison-Wesley*. 1997.
4. Minnaard V, Rabino C, Garcia M. et al. El uso de gráficas en la escuela: otro lenguaje de las ciencias. *Revista Iberoamericana de Educación*. 2002;29:1–6.