Investigation note

Use of a programming language for punctual detection of climate change for each season in Mexico

Ramiro Chávez Mota[1§]
Juan M. Enciso-Medina[2]

[1]Department of Soils-Chapingo Autonomous University. Highway México-Texcoco km 38.5, Chapingo, Texcoco, State of Mexico. [2]Department of Biological and Agricultural Engineering, Texas A&M AgriLife Research, Weslaco, Texas, USA.

[§]Corresponding author: chavezramiro@hotmail.com.

Abstract

The objective of this manuscript was to study with accuracy the point of change in temperature in each of the country's meteorological stations, if it exists, is extremely important to take the most appropriate mitigation measures. In Mexico there are different studies of changes in temperatures, their trends and their relationship with other variables such as geography, precipitation and vegetation. Among them we can mention those of Pavia *et al*. (2009) and Englehart and Douglas (2004) although in none of them has he used programming to be more exact in this determination.

Keywords: schedule, temperature, weather.

Reception date: October 2020
Acceptance date: November 2020

For the study of climate change, there are various databases with a large amount of information to determine temperature trends, which makes their management difficult. For this reason, in the present study programming using the visual basic high-level language for applications (VBA) was used for the management and analysis of information. Parametric statistical techniques such as Student's t (equal variances) and Student's t modified by Welch reported in Haan (2002) were also used to establish whether the difference of the means between the temperatures (maximum and minimum) of two time periods formed in each season is significant or not.

In the present study, to identify the exact year in which the greatest difference between the two periods occurred (before and after this year), programming in the high-level visual basic language for applications (VBA) that is integrated into Excel was used. Data of daily maximum and minimum temperatures were used from the approximately 5 400 meteorological stations of the National Meteorological Service (SMN) page from 1902 to 2012. Monthly averages of both temperatures were made using only the months that had 25 or more observations.

A database with 12 tables was generated, one for each month, where each record in each table represents a meteorological station and the fields are the monthly average temperature in each of the 110 years analyzed for the respective month. The records were refined using certain quality criteria, selecting only those records (meteorological stations) with a monthly average of at least 14 years. Remaining the 12 tables with a number of stations ranging from 3 800 to 3 900 depending on the month in each of the two types of temperatures (Table 1 and 2).

For each station, an interactive process was carried out to find out if in that particular station the maximum and minimum temperatures had changed. This process was the following for the first interaction of a station: the temperature observations were each divided into two periods, the first period formed by the temperatures of the first 5 years with data and the second with the rest of the data (at least 9). In this way, two representative samples of the temperatures of each of the periods were formed. Subsequently, the statistics necessary to compare both populations (size, mean and variance) to each of the samples were calculated and a test was carried out to find out if both samples came from populations with the same variance.

In the positive case, the Student's t-test methodology was used for populations with equal variances and in the negative case, the Welch methodology was applied, which is the same test, but using equation 1 for the degrees of freedom and equation 2 to get the test statistic (t_c). Regardless of the case of both tests, the *p*-value was calculated, saving this value together with the test statistics and the threshold (year of division of the two periods).

For the second interaction, the two samples were formed as follows: the sample from the first period increased with the data from the first year of the sample from the second period, decreasing the size of the sample from the second period by one unit. Thus, formed the samples, the same methodology was applied that was used in the first interaction, again saving the statistics and the *p*-value. The described process became interactive where the last interaction was when the size of the second sample was five data points, which is the same size of the first sample in the first interaction.

At the end of the interactive process, the interaction with the lowest *p*-value was selected, which is the interaction in which the populations being compared present a greater significant difference, if there is one. In some stations there were several interactions where there was a significant difference, but the one where the statistical difference is greater between the two populations (lower *p*-value) was selected.

To find out if the populations are different, the value of the selected p-value was compared with the level of significance (α= 0.1); $gl=\left[\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1-1}+\frac{\left(S_2^2/n_2\right)^2}{n_2-1}\right]$ 1); where: gl= degrees of freedom; $S^2{}_1$= variance of sample 1; $S^2{}_2$ = variance of sample 2; n_1= size of sample 1; n_2= size of sample 2, $Tc=\frac{Xm_1+Xm_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$ 2); where: t_c = test statistic; Xm_1= arithmetic mean of sample 1; Xm_2= arithmetic mean of sample 2; $S^2{}_1$= already defined; $S^2{}_2$= already defined.

In Figure 1 it is observed that the first screen of the computer system that was elaborated for the realization of this study has three options in which in its first option <by station> you can see a list of all the almost 3 900 stations to which the mentioned statistical analysis is done. On the right side, you can see four columns. The first two columns would be year and monthly average temperature (of the corresponding month) of the first period compared. In the same way, it's had in the third (year) and fourth (monthly average temperature) columns, but for the second period compared.

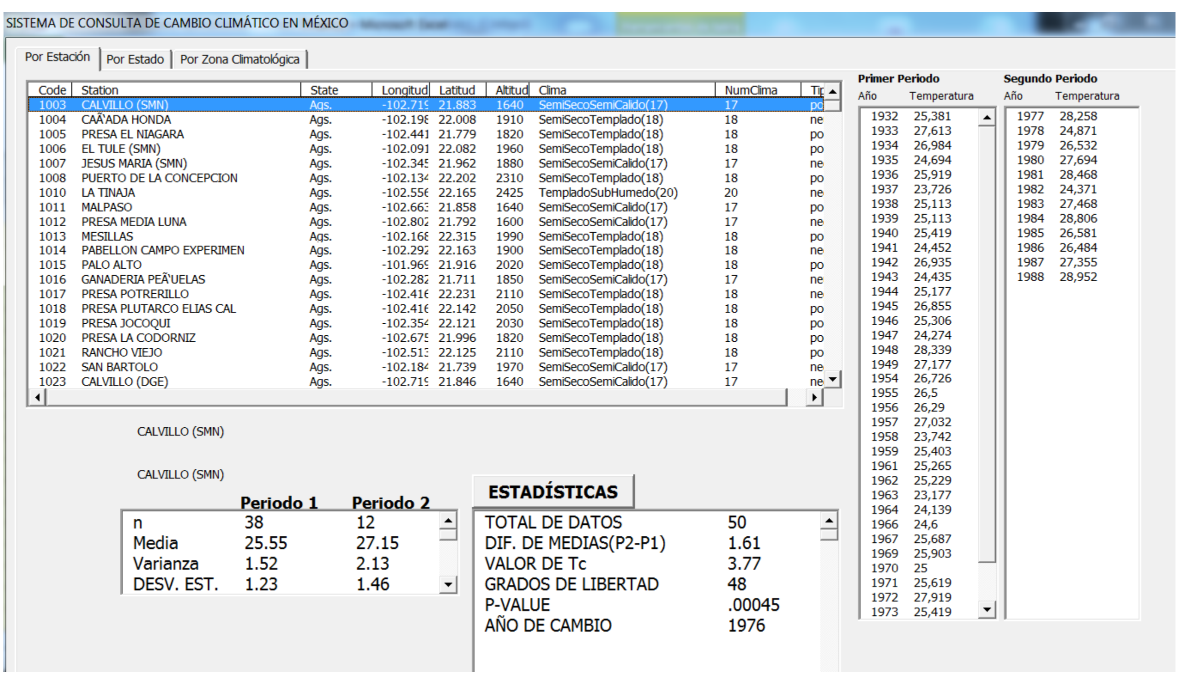

Figure 1. First screen of the computer program developed for the present study. The stations of the two periods compared and some statistics for each of them are presented. In addition, the statistics of the test with a lower *p*-value.

In the lower left corner a table is presented in which some statistics of the average temperature data are presented in the selected station of the two periods that are being compared, such as n (number of observations), mean, variance, standard deviation of each of the periods compared. The last table of this screen shows the statistics of the test used in the comparison, such as the total data, DIF of means (P2-P1) (difference of means between the two periods, value of tc (the value of the test statistic), degrees of freedom, *p*-value, year of change (exactly the year in which the change occurs where the means present the greatest statistical difference).

The program also presents the option to analyze the temperatures of the stations grouped by state (Figure 2). In the menu on the right of Figure 2, the list of states that make up Mexico is presented, from which any of them can be selected. When selecting the system, it presents the list of weather stations that the state has that can be seen in the middle box.

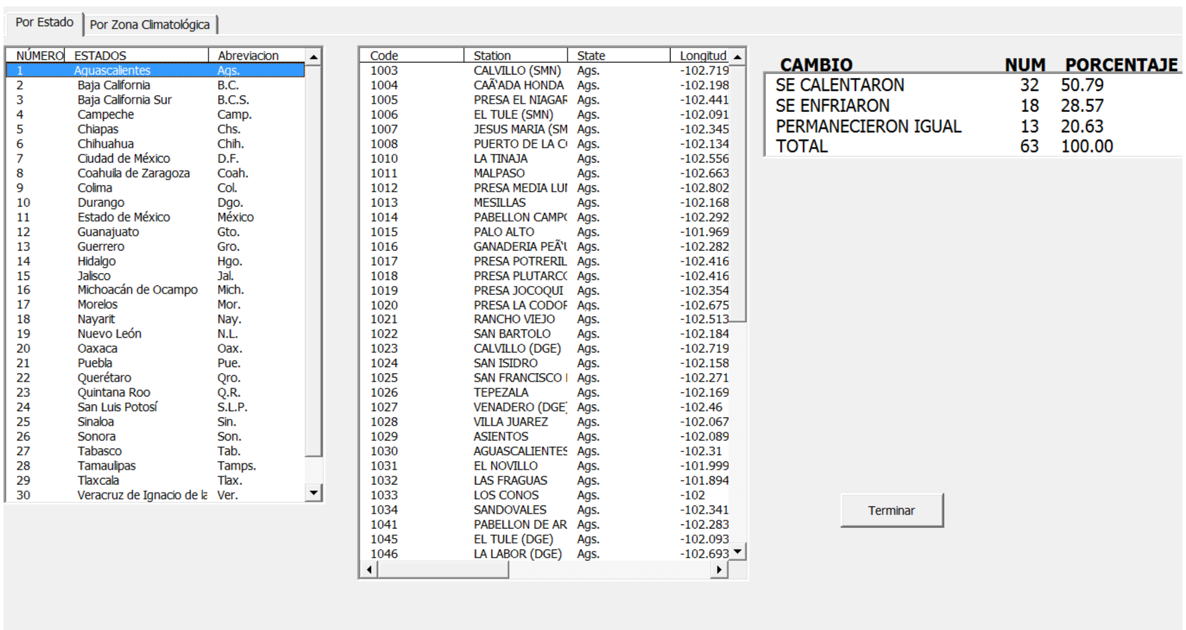

Figure 2. Second screen of the computer program developed for the present study. A menu of the list of states that make up the country Mexico is presented, from which any of them can be selected and it presents the list of stations of the selected state (list in the middle) and the number and percentage of stations that were heated, cooled and they remained the same, in addition to the totals.

In the table on the right, you can see the most important statistics of all the stations, such as how many stations were heated and their percentage, how many cooled and their percentage and how many remained the same and their percentage, as well as the total of each one of them. the above. It was found that the two states where there is a higher percentage of stations that warmed up are Michoacán de Ocampo (60.45%) and Quintana Roo (60.53%). In the two states where there is a higher percentage of stations that cooled down are Durango (46.9%) and Quintana Roo (47.62%).

In addition, you have the option of analyzing the temperatures of the stations for each one of the existing climatological zones in Mexico (Figure 3). Choosing one of the zones in the menu on the left of Figure 3 displays the list of stations that belong to that zone (in the middle box). In the table on the right you can see the statistics that have already been explained for Figure 2 but by climatological zone.

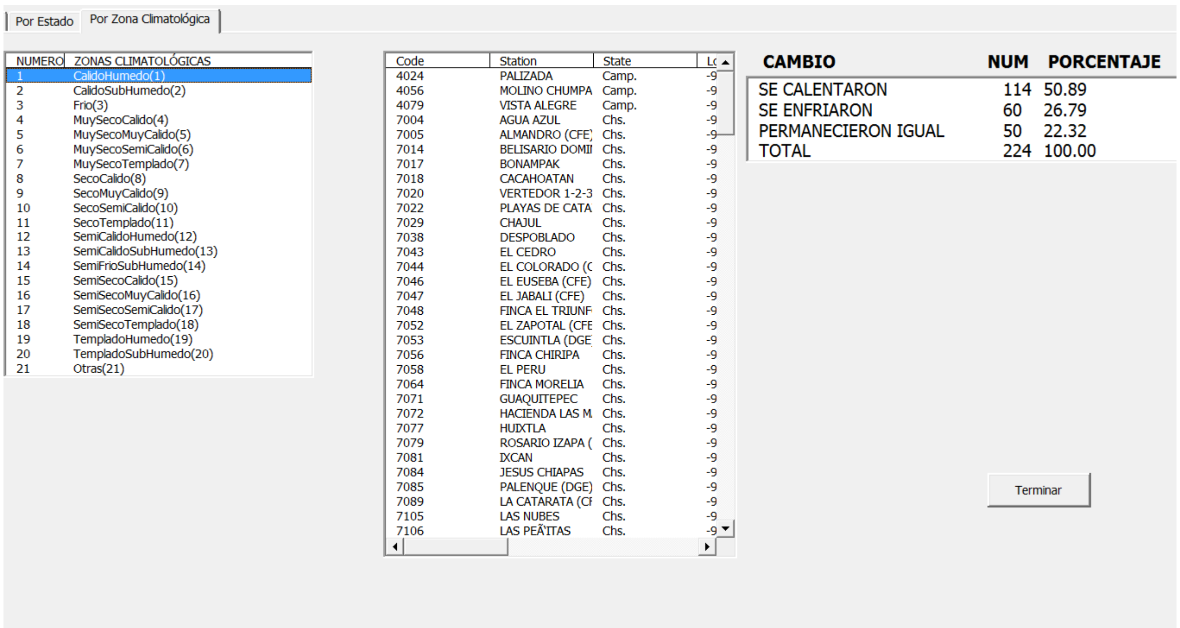

Figure 3. Third screen of the computer program developed for the present study. A menu of the list of climatological zones that make up the country Mexico is presented, from which any of them can be selected and it presents the list of stations in the selected area (list in the middle) and the number and percentage of stations that were heated, cooled and they remained the same, in addition to the totals (list to the right).

It was found that the two climatological zones where there is a higher percentage of stations that warmed up are the cold zone (66.67%) although there are only three seasons in this zone and the temperate subhumid zone (53.65%). And the two areas where there is a higher percentage of stations that cooled down are very dry, very warm (100%), although there is only one station in this area and very warm semi-dry Quintana (47.27%).

Referring to global statistics, 80% of the stations analyzed throughout Mexico (between 3 800 and 3 900 depending on the month) there was a significant change in both the maximum and minimum temperatures. In the analysis of maximum temperatures, in all months there was a marked predominance of the stations where there was a significant increase over those that had a decrease. In the analysis of the minimum temperatures, the opposite occurred. In almost all months (with the exception of January and February) the seasons where there was a significant decrease in temperatures predominated (Table 1).

Table 1. Number of total stations for each month where there was an increase, a decrease and where there was no significant change in the maximum and minimum temperatures.

Month	Maximum temperatures				Minimum temperatures			
	Increase	Decrement	Without changes	Total	Increase	Decrement	Without changes	Total
January	1 805	1 229	792	3826	1 579	1 576	724	3879
February	1 952	1 120	804	3876	1 665	1 367	842	3874
March	1 954	1 167	764	3885	1 529	1 564	792	3885
April	1 895	1 182	805	3882	1 372	1 730	778	3880
May	2 048	1 093	740	3881	1 491	1 710	681	3882
June	2 025	1 071	786	3882	1 488	1 713	682	3883
July	2 157	1 143	587	3887	1 539	1 664	682	3885
August	2 098	1 163	622	3883	1 609	1 651	623	3883
September	1 962	1 170	758	3890	1 521	1 658	708	3887
October	1 975	1 180	718	3873	1 443	1 551	881	3875
November	1 843	1 253	757	3853	1 273	1 687	890	3850
December	1 805	1 229	792	3826	1 142	1 888	794	3824

Conclusions

A significant change in temperatures (maximum and minimum) was identified in approximately 80% of the more than 3 800 stations analyzed. At maximum temperatures, more seasons have a significant increase than those with a decrease, in all months of the year. For the minimum temperatures there was a contrary trend, since in almost every month (with the exception of January and February) there were more stations that suffered a decrease in temperature.

Cited literature

Englehart, Phil J. and Arthur V. Douglas. 2004. Characterizing regional-scale variations in monthly and seasonal surface air temperature over Mexico. Inter. J. Climatol. 24(15):1897-1909.

Haan, Ch. T. 2002. Statistical methods in hydrology. 2002.

Pavia, E. G.; Graef, F. and Reyes, J. 2009. Annual and seasonal surface air temperature trends in Mexico. Inter. J. Climatol. 29(9):1324-1329.

Solomon, S. 2007. Climate Change 2007-the physical science basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC. Cambridge University Press. Vol. 4.

Yue, S. and Ch, Y. W. 2002. Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall Test. Water Resources Res. 38(6):41-47.