International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 05 - Issue 06 || June 2019 || PP. 36-54

Analysis of the Trend Annual Maximum Temperatures in Mexico City 2005 - 2018 and with the WRF Program Case Two: Daily Average Maximum Temperature Data, with Gaussian behavior

M. Sc. Zenteno Jiménez José Roberto (1)
Geophysical Eng. Mireles Arellano Fernando (1)
Geophysical Engineering, National Polytechnic Institute (1),
México City, ESIA-Ticoman Unit Delegation Gustavo A. Madero
Email: jzenteno@ipn.mx-ing.fer.geo@gmail.com

Abstract: The methodology was used to obtain new functions of distribution of normal probability and extreme value by Bayesian inference and stochastic mixing of Gaussians. The proposed methodology is oriented to data with Gaussian behavior and consists of adjusting the normal distribution function to the data of time series of maximum temperature data, to give a behavior of the temperatures maximum, later we use the Bayesian inference for normal data, in this case we are looking for behavior and trends with new Gaussian functions and extreme value. To validate the model we use the following statistical estimators: measurement of the root of the quadratic error, quadratic error, coefficient of determination and prediction approximation.
Using the new means and variances of the new functions of extreme distribution we generate two new functions of normal distribution a minimum and a maximum. Thus already having the three normal probability distribution functions, the adjusted first and the two new normal distribution functions we introduce the Gaussian stochastic mixing method to give a new function of general normal probability distribution for the trend of the maximum temperatures for Mexico City. In addition to making a temperature forecast with the WRF for comparison. The database that is used is from the page of the City of Mexico http://www.aire.cdmx.gob.mx/

Keywords: Temperature, Random and Extreme Variable Distribution Functions, Bayesian Inference, WRF.

Introduction

In the normal distribution, equation (1) one can calculate the probability that several values will occur within certain ranges or intervals. However, the exact probability of a particular value within a continuous distribution, such as the normal distribution, is zero. This property distinguishes the continuous variables, which are measured, from the discrete variables, which are counted. As an example, time (in seconds) is measured and not counted.

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2a^2}\right)} \quad \textbf{(1)}$$

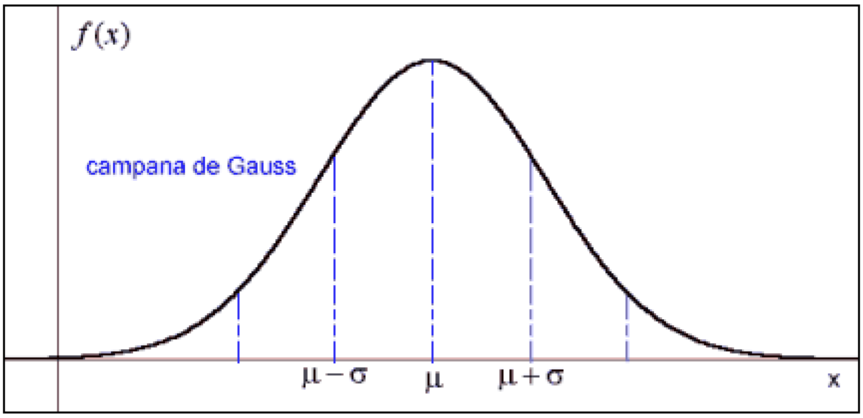


Figure 1. Gaussian bell or Gaussian density function (Source: Internet)

Now, as we obtain the estimation parameters of a probability distribution function, we will use the Maximum Likelihood technique the parameters to be estimated for the adjustment. The maximum likelihood method is a procedure to obtain a point estimator of a random variable.
Let (X1, ..., Xn) be a random sample with a distribution function f (x | θ).

We define the likelihood function as:

$$L(\theta|X_1, X_2, \dots X_n) = \prod_{i=1}^{n} f(X_i|\theta) \quad (2)$$

The estimator of θ in the maximum likelihood method is the value that maximizes the likelihood function. This value is called the maximum likelihood estimator EMV (θ).
Be:

$$L(\theta|X_1, X_2, \dots X_n) = \ln\big(L(\theta|X_1, X_2, \dots X_n)\big) = \sum_{i=1}^{n} f(X_i|\theta) \quad (3)$$

So the maximum likelihood estimator is defined as:

$$EMV(\theta) = \max_{\theta \in \Theta} L(\theta|X_1, X_2, \dots X_n) \quad (4)$$

By means of the above description we obtain the parameters of a Normal Distribution which are doing the following:
By the Maximum Likelihood Method, the likelihood function will be:By means of the above description we obtain the parameters of a Normal Distribution which are doing the following:
By the Maximum Likelihood Method, the likelihood function will be:

$$L(\mu, \sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{N} e^{\left(-\frac{\sum(x-\mu)^2}{2\sigma^2}\right)} \quad (5)$$

Applying logarithms and deriving with respect to the parameters to be estimated, we have a system of equations as follows:

$$\frac{\partial \log(L(\mu, \sigma))}{\partial \mu} = \frac{\sum x}{\sigma^2} - \frac{N_\mu}{\sigma^2} = 0 \quad (6)$$

$$\frac{\partial \log(L(\mu, \sigma))}{\partial \sigma} = -\frac{N}{\sigma} + \frac{\sum(x-\mu)^2}{\sigma^3} = 0$$

With the Solution:

$$\mu = x \qquad y \qquad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x-\mu)^2 \quad (7)$$

With the first adjusted Normal we get

$$Normal(\mu, \sigma^2) \quad (8)$$

Now what we are looking for are extreme values that we want to know, the probability of occurrence, so we use the Bayesian Inference to find this probability with a new distribution function that will be part of the new functions of normal distribution and function of extreme variable or gev, our new unknown will be the average.

Bayesian inference

Bayesian inference is the process of analyzing statistical models with the incorporation of prior knowledge about the model or model parameters. The root of such inference is Bayes' theorem:

$$P(Parameter|Data) = \frac{P(Data|Parameter) * P(Parameter)}{P(Data)} \approx FVerosimilitud * PDF\ Priori \quad (9)$$

In this case we have the observations in the normal distribution form

$$X|\theta \sim N(\theta, \sigma^2) \quad (10)$$

Where the sigma is previously known and the PDF a Priori is

$$\theta \sim N(\mu, \tau^2) \quad (11)$$

Here mu and tao are also known, we are looking for n samples of the observed data, in the case of Ozone the maximum values or above 150 ppb, the PM10 particulate case above 120 microgram / m3 (2018), the case of PM2. 5 above 65 microgram / m3 and in the case of Maximum Temperatures it is the whole Data sample and thus we obtain the New Normal Distribution Function with the new searched parameter:

$$\theta|X \sim NB\left(\frac{\tau^2}{\frac{\sigma^2}{n}+\tau^2} * X + \frac{\frac{\sigma^2}{n}}{\frac{\sigma^2}{n}+\tau^2} * \mu, \frac{\frac{\sigma^2}{n} * \tau^2}{\frac{\sigma^2}{n}+\tau^2}\right) \quad (12)$$

Now this data contains noise, there are null values or zeros, from the adjustment process and although we have a good approximation it is seen that it does not adjust well to the data so that they produce inaccuracy to the Normal Distribution Function, therefore we apply a random noise with a Uniform Distribution with the length of the terms of the time series of the data, we now apply an adjustment with the Extreme Value Distribution Function (GEV) to find the parameters that better fit these data with the random noise

$$GEV(\mu, \sigma, k) \quad (13)$$

A GEV is adjusted with this uniform random distribution

$$GEVa(Xa, \mu, \sigma, k1) \quad (14)$$

And subsequently a GEVa is generated (with the random parameters of GEVa)

$$GEVA(\mu a, \sigma a, k1) \quad (15)$$

We also now adjust a GEV of the Input Data, this is where the Extreme Value Theory comes in, and now we look for a new Distribution Function, and it is where the new equation is applied based on the properties of the parameters that were previously obtained:

GEV 1

$$k2 = \left(\frac{GEVk + GEVkA}{\sum_{i=1}^{2} pn}\right) \quad (16)$$

$$SigmaSD = \left(\frac{GEVsd + PostSD}{\sum_{i=1}^{2} pn}\right) \quad (17)$$

$$Mupostmean = \left(\frac{GEVmu + Postmean}{\sum_{i=1}^{2} pn}\right) \quad (18)$$

Now we get a second equation to get the new parameters
GEV 2

$$k2 = \left(\frac{GEVk + GEVkA}{\sum_{i=1}^{2} pn}\right) \quad (19)$$

$$SigmaSD2 = \left(\frac{GEVsd + PostSD + GEVsdA}{\sum_{i=1}^{3} pn}\right) \quad (20)$$

$$Mupostmean2 = \left(\frac{GEVmu + Postmean + GEVmuA}{\sum_{i=1}^{3} pn}\right) \quad (21)$$

The functions of extreme value or GEVs with the new parameters of the new functions of extreme distribution see [14]:

$$GEV\left(\sum_{i=1}^{n} \frac{\mu_i}{n}, \frac{1}{n-1}\sum_{i=1}^{n-1} \sigma_i, k\right) \quad (22)$$

With

$$k > 0 \; x \in \left[\mu - \frac{\sigma}{k}, +\infty\right] \qquad\qquad k < 0 \; x \in \left[-\infty, \mu - \frac{\sigma}{k}\right]$$

According to the previous definition, the shape parameter can be deduced from the method of higher moments is between two givens that there are two parameters of minor order composing this, according to the equations obtained, if we would not use Newton Rhapsonforget the mentioned parameters, so then we have:

$$\mu - \frac{\sigma}{k} = 0 \;\; y \;\; asi \;\; k = \frac{\sigma}{\mu} \tag{23}$$

$$k = \frac{\sum_{i=1}^{n-1} \frac{\sigma_i}{n-1}}{\sum_{i=1}^{n} \frac{\mu_i}{n}} = \sum_{i=1}^{n} \frac{\frac{\sigma_i}{n-1}}{\frac{\mu_i}{n}} = \sum_{i=0}^{n} \frac{\sigma_i n}{\mu_i (n-1)} = \sum_{i=0}^{n} \frac{n k_i}{(n-1)} \tag{24}$$

Observing the Series you can reach the equivalent sum and thus obtain an expression for the form parameter

$$k = \sum_{i=0}^{n} \frac{k_i}{n} \tag{25}$$

Then we have the following

$$GEV\left(\sum_{i=1}^{n} \frac{\mu_i}{n}, \frac{1}{n-1} \sum_{i=1}^{n-1} \sigma_i, \sum_{i=0}^{n} \frac{k_i}{n}\right) \tag{26}$$

We have the following probability distribution functions, the Bayesian Normal, the GEV of the data and the Random GEV, therefore there are 3 functions that we have so the first two sums of two probability distribution functions.We have the following probability distribution functions, the Bayesian Normal, the GEV of the data and the Random GEV, therefore there are 3 functions that we have so the first two sums of two probability distribution functions.

Table 2. GEV and Normal probability distribution functions

GEV	Normales
$GEV(\mu, \sigma, k)$	$Normal(\mu, \sigma^2)$
$GEV(Mupostmean\ \mu, SigmaSD\ \sigma, k2)$	$Normal1(E(GEV1), \sqrt{Var(GEV1)})$
$GEV(Mupostmean2\ \mu, SigmaSD2\ \sigma, k2)$	$Normal2(E(GEV2), \sqrt{Var(GEV2)})$

Adjustment Indicators

The deviation indicators of a group of data in relation to a model can be used to assess the goodness of fit between the two. Among the most common indicators are the following. Those that were used to determine the distribution that best fit gave the data. They are the mean square error (RMSE), mean square error (MSE), prediction accuracy (PA) and coefficient of determination (R2) Table 4 gives the equations for the adjustment indicators that have been used by Lu (2003) and Junninen et al. (2002).

Table 3. Adjustment Indicators

Indicators	**Equations**
Root Mean Square Error (Raíz Cuadrada del Error)	$RMSE = \sqrt{\left(\frac{1}{N-1}\right) \sum_{i=1}^{N} (Pi - Oi)^2}$
Mean Square Error (Error Cuadrado Principal)	$RMSE = \left(\frac{1}{N}\right) \sum_{i=1}^{N} (Pi - Oi)^2$
CoeficienteofDetermination (Coeficiente de Determinación)	$R^2 = \left(\frac{\sum_{i=1}^{N} (Pi - P)(Oi - O)}{N S_p S_o}\right)^2$

PredictionAccuracy (Precisión de Predicción)	$AP = \frac{\sum_{i=1}^{N}(Pi - O)^2}{\sum_{i=1}^{N}(Oi - O)^2}$

Notation: N = number of observations, P_i= predictive valúes, O_i= observed values, P = average of predicted values, **O** = average of the observed values, S_p= Standard Deviation of Predicted valúes, S_o = Standard deviation of theobserved valúes.

Stochastic Method of Gaussian Mixtures

The grouping model most closely related to statistics is the distributions-based model. Groups can then easily be defined as the objects that most likely belong to the same distribution. A convenient property of this approach is that this closely resembles the way in which artificial data sets are generated: by random sampling of objects in a distribution.

One of the most prominent methods is known as the Gaussian mixing model (used in the expectation-maximization algorithm). Here, the data set is usually modeled with a fixed number (to avoid overfitting) of Gaussian distributions that is randomly initialized, and whose parameters are iteratively optimized to better classify the data set. This will converge to a local optimum, multiple runs can produce different results. To obtain a good grouping, the data are often assigned to the Gaussian distribution with a higher probability of belonging to such grouping.

The Gaussian mixture models are a probabilistic model to represent subpopulations normally distributed within a general population. Mixing models in general do not require knowing to which subpopulation a data point belongs, which allows the model to automatically learn the subpopulations, using Expectation-Maximization (EM).

GMMs are widely used for grouping and estimating. However, they have a wide range of applications in other fields, such as modeling meteorological observations in Geosciences (Zi, 2011), certain autoregressive models or the noise of some time series.

If you believe that your data comes from a set of different normal distributions, then the GMM is an appropriate analysis tool. The normal distribution is an underlying assumption, which means that, although it is assumed that the distributions are Gaussian, or it is possible that they are not. In some cases, you may not be able to count, but use logic or prior knowledge to assume that your data has a normal distribution. Therefore, models created from a GMM method carry a certain level of uncertainty.

A Gaussian mixing model means that each data point is put (randomly) from one of the data classes C, with probability p_i to be extracted from class i, and each class is distributed as Gaussian with mean standard deviation μ_i andσ_i. Given a set of data extracted from this distribution, we seek to estimate these unknown parameters.

The algorithm used here for the estimation is EM (Maximization of expectation). In short, if we knew the class of each of the N input data points, we could separate them, and use Maximum Probability to estimate the parameters of each class. This is the step that performs (soft) selections of (unknown) classes for each of the data points based on the previous round of parameter estimates for each class.

$$\phi_j := \frac{1}{m}\sum_{i=1}^{m} w_j^{(i)},$$

$$\mu_j := \frac{\sum_{i=1}^{m} w_j^{(i)} x^{(i)}}{\sum_{i=1}^{m} w_j^{(i)}},$$

$$\Sigma_j := \frac{\sum_{i=1}^{m} w_j^{(i)} (x^{(i)} - \mu_j)(x^{(i)} - \mu_j)^T}{\sum_{i=1}^{m} w_j^{(i)}}$$

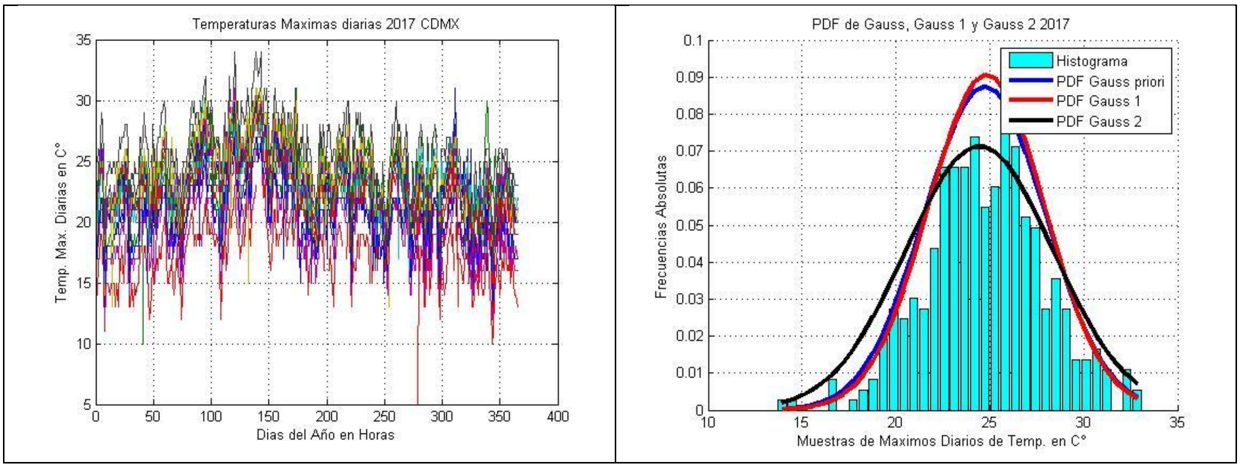
Figure3. Basic equations of EM Algorithms (Source: http://mccormickml.com/2014)

Results of the Obtaining New Gaussians the first 3 and last years

Normal PDF and its Means	Adjustment, error and approximation estimators	AverageConfidence Interval
2005 **Normal= 24.78** **Normal1= 24.78** **Normal2= 24.38**	**MSE= 3.6835e-04** **RMSE = 0.0199** **PA= 1.033** **R2=0.9632** **AI=0.9994**	**CI =19.23 26.48** **CI1 =21.16 28.40** **CI2 =20.76 28.00**
2006 **Normal= 21.55** **Normal1= 21.54** **Normal2= 21.31**	**MSE= 5.3664e-04** **RMSE= 0.0238** **PA=1.0355** **R2=0.9714** **AI=0.9992**	**CI =16.21 24.18** **CI1 =17.56 25.52** **CI2 =17.33 25.29**
2007 **Normal= 21.51** **Normal1= 21.38** **Normal2= 22.53**	**MSE=0.0031** **RMSE=0.0573** **PA=0.9704** **R2=0.9696** **AI=0.9956**	**CI =19.12 27.43** **CI1 =17.23 25.53** **CI2 =18.37 26.68**
2016 **Normal= 23.32** **Normal1= 23.43** **Normal2= 22.54**	**MSE=0.0010** **RMSE=0.0323** **PA=1.0397** **R2= 0.9853** **AI=0.9985**	**CI =18.81 22.61** **CI1 =21.53 25.33** **CI2 =20.64 24.44**
2017 **Normal= 24.73** **Normal1= 24.82** **Normal2= 24.43**	**MSE=2.3069e-04** **RMSE=0.0154** **PA=1.0158** **R2=0.9869** **AI=0.9997**	**CI =21.45 25.30** **CI1 =22.90 26.75** **CI2 =22.51 26.36**
2018 **Normal= 23.994** **Normal1= 24.13** **Normal2= 23.56**	**MSE=3.8493e-04** **RMSE=0.0198** **PA=0.9829** **R2=0.9885** **AI=0.9995**	**CI =23.67 24.30** **CI1 =23.82 24.45** **CI2 =23.25 23.88**

Figures 4. Temporal Series of Daily Maximum Concentrations of O3 and its PDF GEV

Table 5. 2017



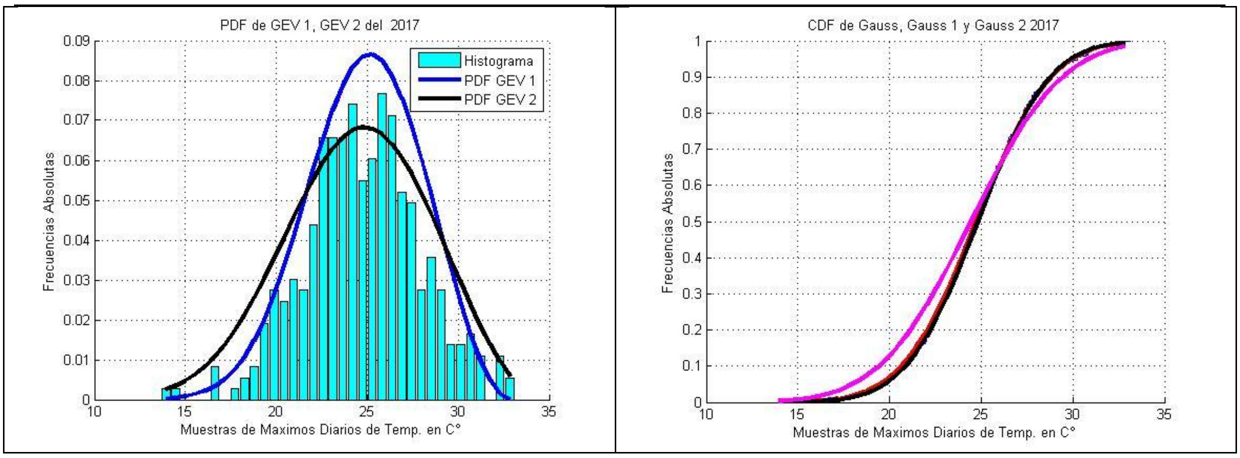

PDF de GEV 1, GEV 2 del 2017
Histograma
PDF GEV 1
PDF GEV 2
Frecuencias Absolutas
Muestras de Maximos Diarios de Temp. en C°
CDF de Gauss, Gauss 1 y Gauss 2 2017
Frecuencias Absolutas
Muestras de Maximos Diarios de Temp. en C°

Table6. 2018

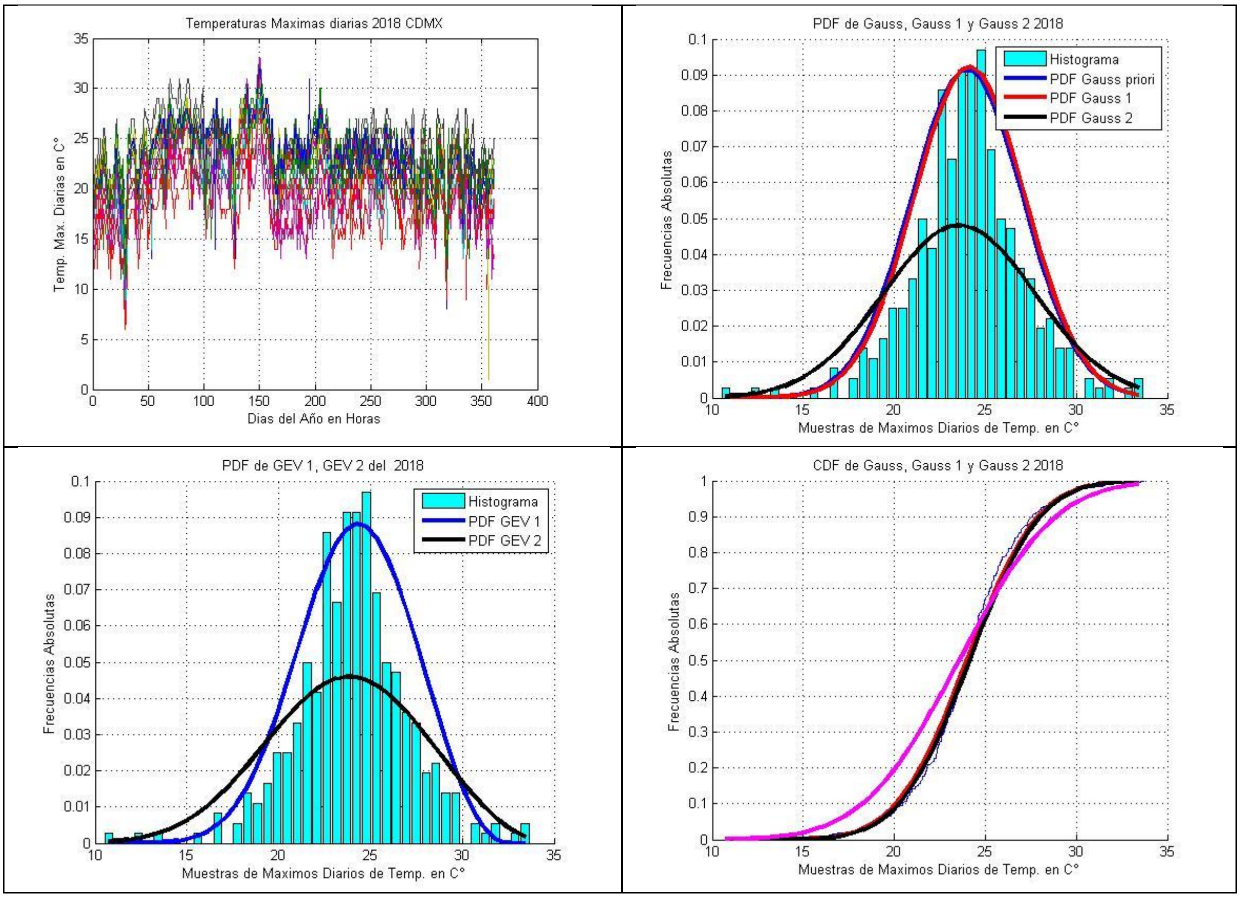
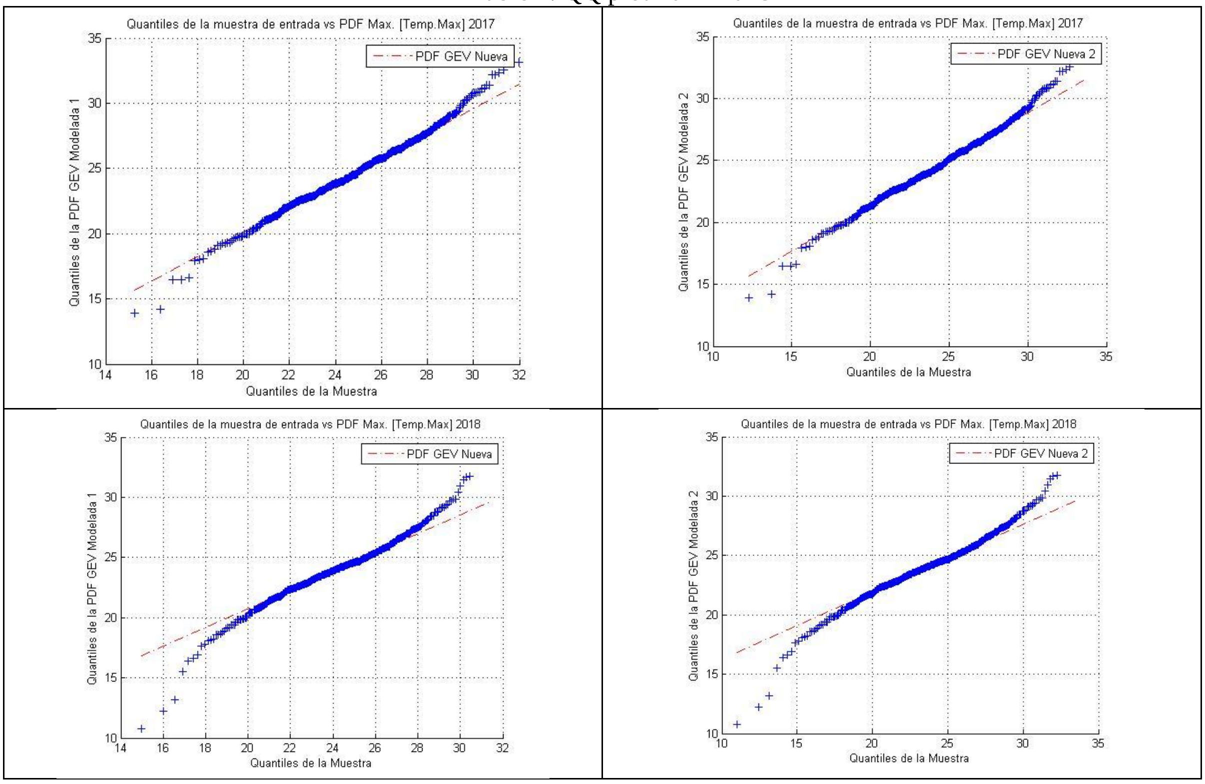
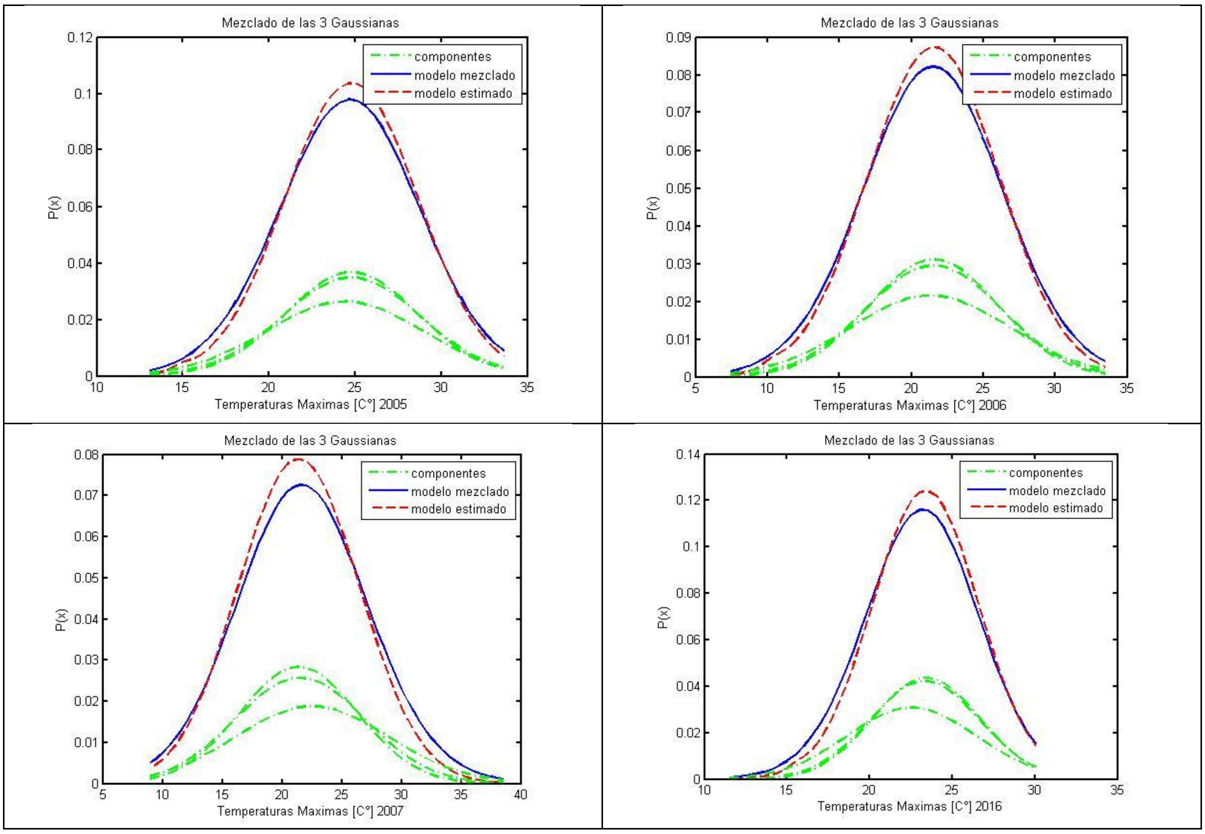
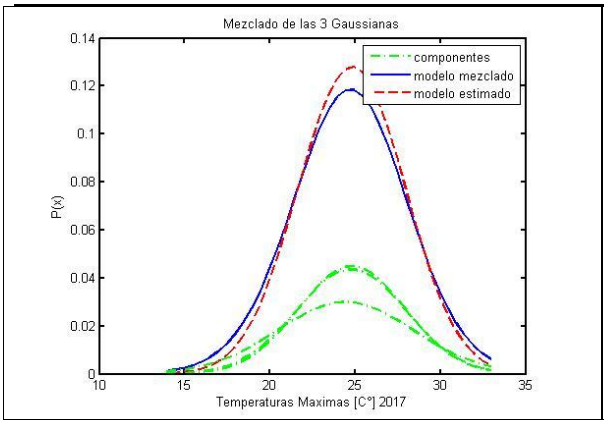

Temperaturas Maximas diarias 2018 CDMX
Temp. Max. Diarias en C°
Dias del Año en Horas
PDF de Gauss, Gauss 1 y Gauss 2 2018
Histograma
PDF Gauss priori
PDF Gauss 1
PDF Gauss 2
Frecuencias Absolutas
Muestras de Maximos Diarios de Temp. en C°
PDF de GEV 1, GEV 2 del 2018
Histograma
PDF GEV 1
PDF GEV 2
Frecuencias Absolutas
Muestras de Maximos Diarios de Temp. en C°
CDF de Gauss, Gauss 1 y Gauss 2 2018
Frecuencias Absolutas
Muestras de Maximos Diarios de Temp. en C°

Table 7. QQ plot 2017-2018



Tables 8. Ofthe Gaussian Mixture





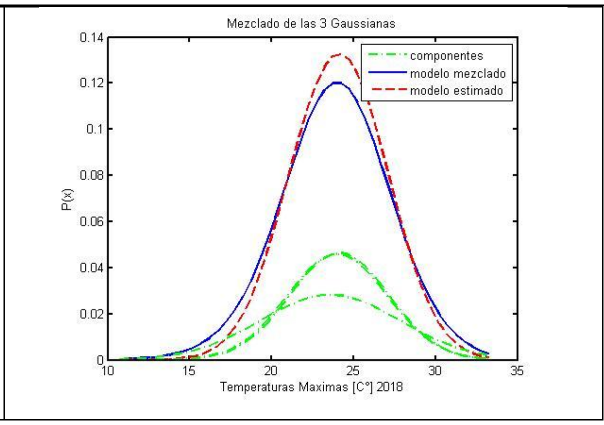


Table 9. Of Data Shown and Data Official

Years México City	**CONAGUA MaximumTemperaturesC°**	**Means Obtained from Gaussian Sums or Convolutions of Maximum Temperatures C °**	**Temperatureanomaly**	**Average Gaussian mixture temperature C°**
2005	23.7	24	+1.4 C°	24.77
2006	23.9	22.5	+0.5 C°	21.50
2007	23.9	23	+0.5 C°	21.41
2008	23.7	23.5	+1.65 C°	23.56
2009	24.4	21.5	+0.5 C°	20.37
2010	23.6	22	+1.6 C°	21.50
2011	24.2	23.6	+1.6 C°	23.95
2012	23.4	22	+1.5 C°	20.99
2013	24.6	23.5	+0.5 C°	23.11
2014	23.8	23	+1.0 C°	23.11
2015	23.9	24	+1.0 C°	24.76
2016	24.1	23	+1.7 C°	23.37
2017	24.6	24.7	+1.7 C°	23.37
2018	24.3	24.0	+0.7 C°	24.07

Data Sources:https://smn.cna.gob.mx/es/climatologia/temperaturas-y-lluvias/resumenes-mensuales-de-temperaturas-y-lluvias
http://www.aire.cdmx.gob.mx

Table 10. Temperature Averages Obtained with the new Normals and the Gaussian Mixture.

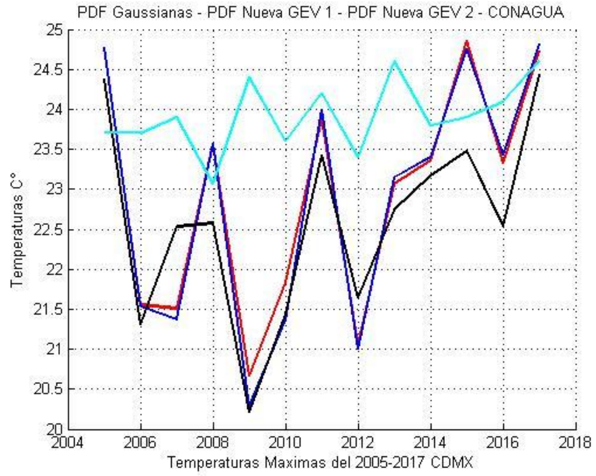
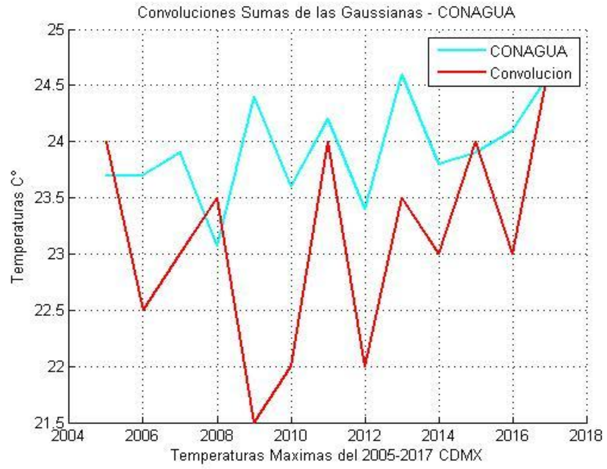


Figure 5. Temporal Series of Maximum Daily Temperatures and their found values

Gaussian adjustment in Red
Gaussian adjustment and GEV 1 in blue
Gaussian adjustment and GEV 2 in black
Light Blue Data CONAGUA CDMX
Informative references:
https://www.eleconomista.com.mx/politica/7-efectos-visibles-del-calentamiento-global-en-la-Ciudad-de-Mexico-20170603-0010.html



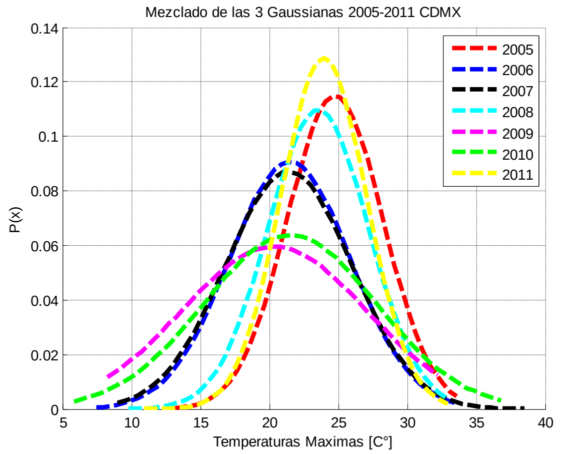
https://www.nytimes.com/es/2018/01/22/temperaturas-mundo-record-calor/

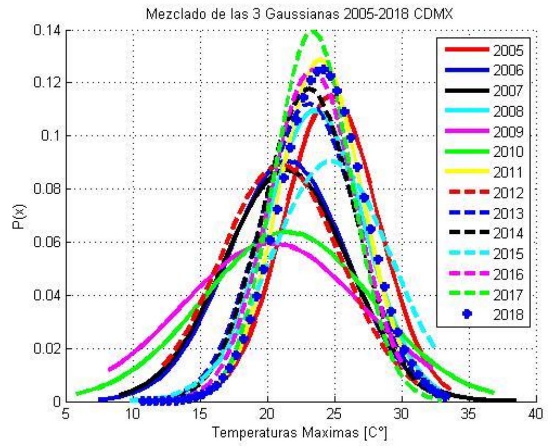
2015, 2016 and 2017 have been the hottest years since there are records

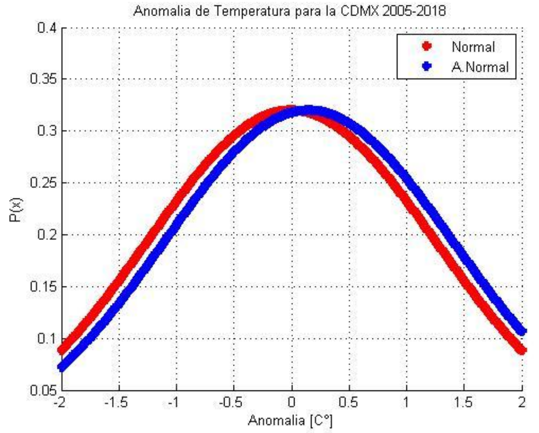
As you can see both the Gaussian sums as the new Gaussians and the new GEV, reference [14] are giving the trend with the means found and corroborate the given measurements, which we can express that temperatures are increasing in the CDMX.As you can see both the Gaussian sums as the new Gaussians and the new GEV, reference [14] are giving the trend with the means found and corroborate the given measurements, which we can express that temperatures are increasing in the CDMX.

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 05 - Issue 06 || June 2019 || PP. 36-54

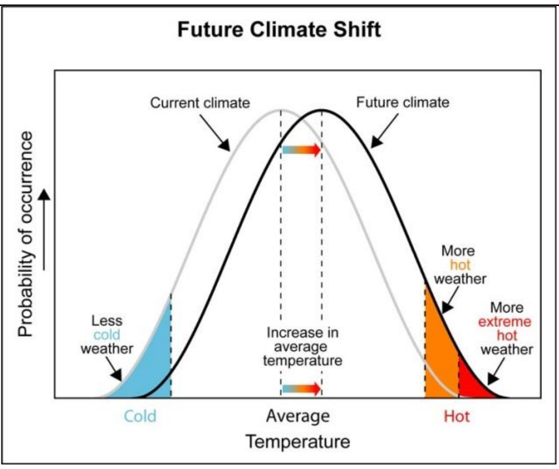
Figures 11. Maximum Temperatures in the CDMX 2005-2018







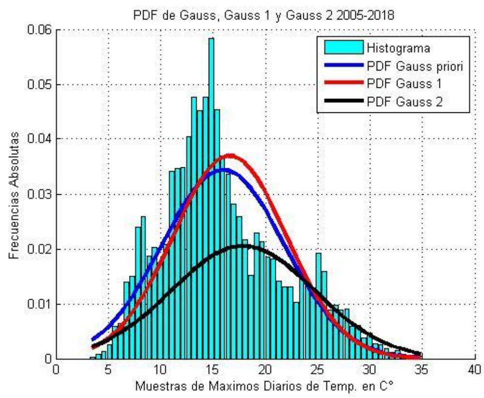
Informativereferences:
http://www.eluniversal.com.mx/nacion/sociedad/ultimos-tres-anos-han-sido-los-mas-calurosos-de-la-historia-asegura-la-unam

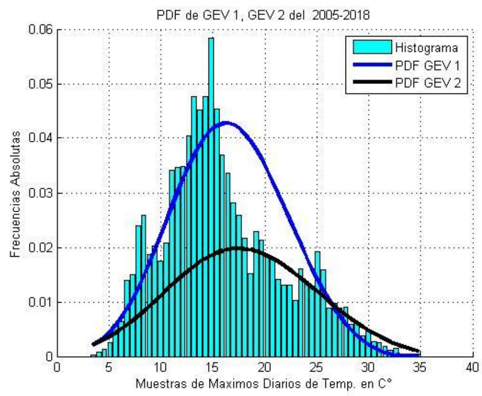


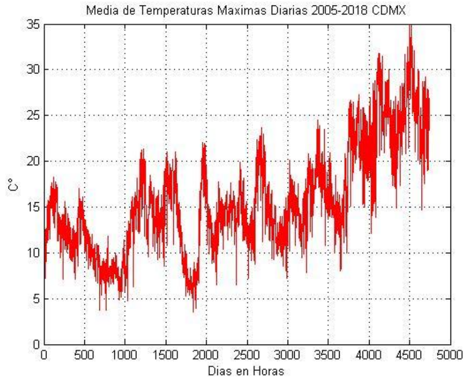
The Bell curve showing how an increase in average temperatures leads to an increase in warm and extreme weather. Note also that this does not mean that there will be no more cold weather: these cold events will get rarer but will not go away.(Source: US Climate Change Science Program / Southwest Climate Change Network)

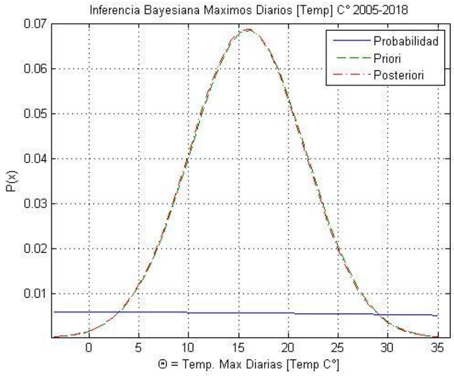
General Adjustment of the 2005-2018 Tendency

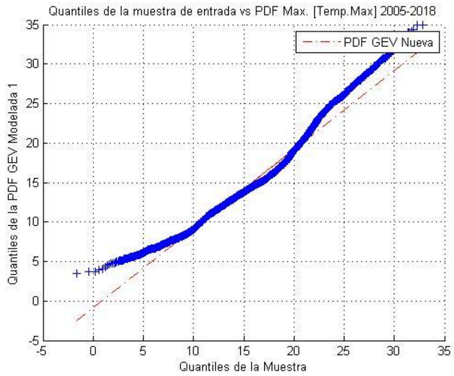
Figures 6

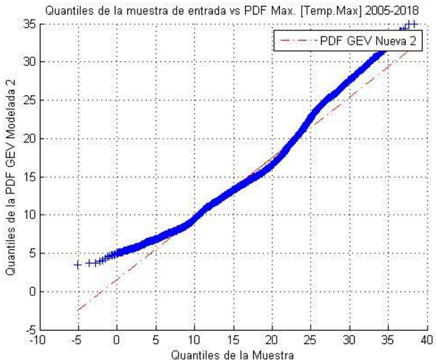


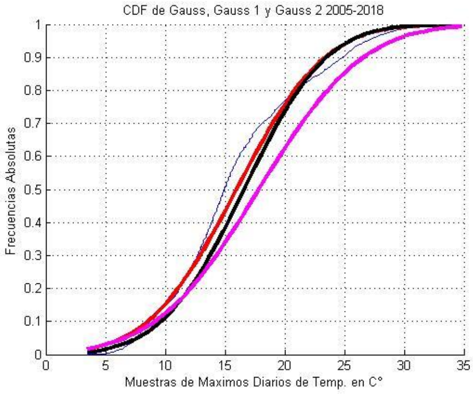












Normal Mean: 15.96 Sdt= 9.31

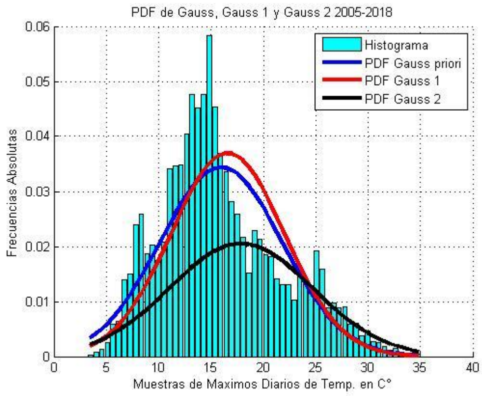
CI = 15.8002 16.1305

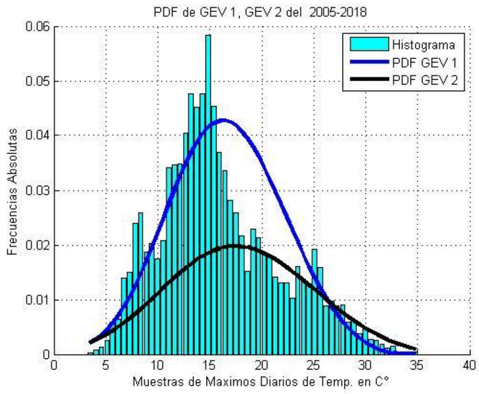
Normal 1 Mean: 16.59 Sdt= 5.39

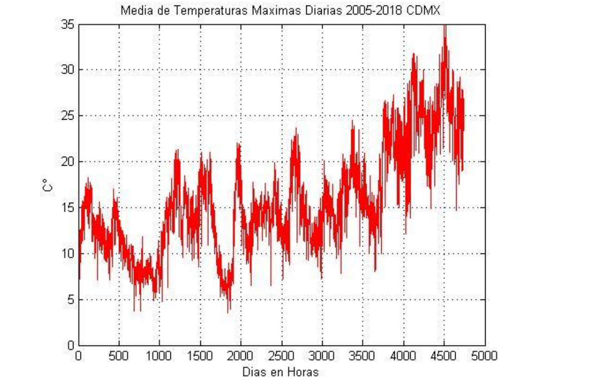
CI1 = 16.4285 16.7587

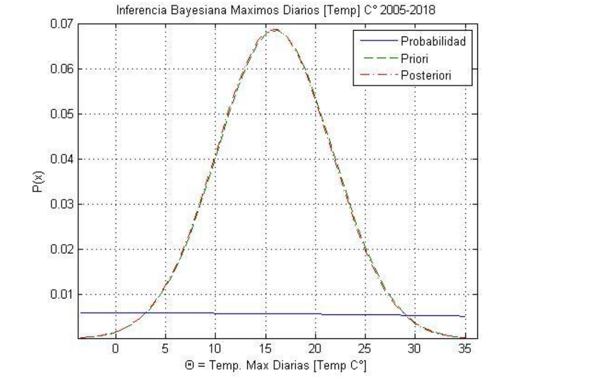
Normal 2 Mean 17.86 Sdt= 6.80

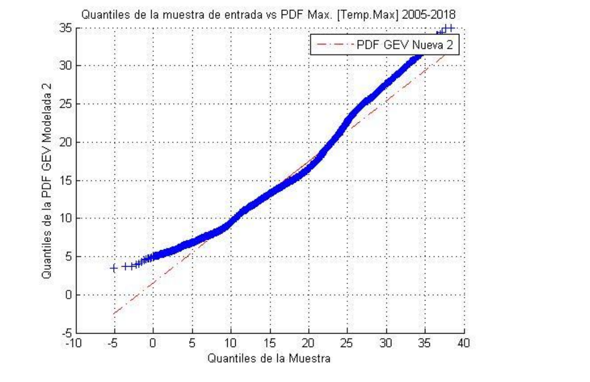
CI2 = 17.6971 18.0273

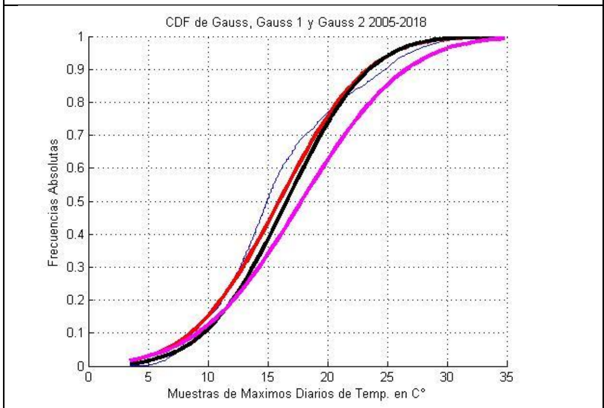












Normal Mean: 15.96 Sdt= 9.31

CI = 15.8002 16.1305

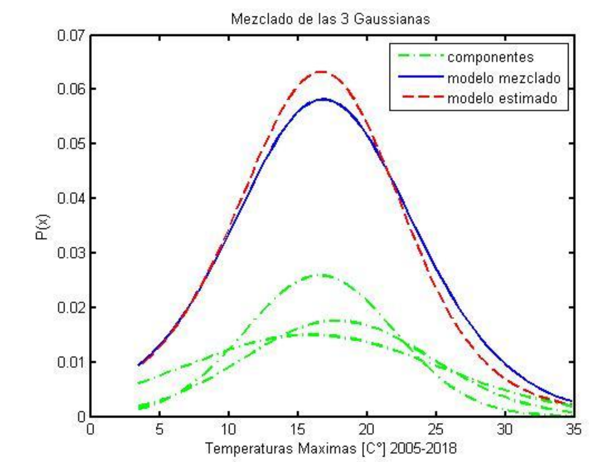
Normal 1 Mean: 16.59 Sdt= 5.39

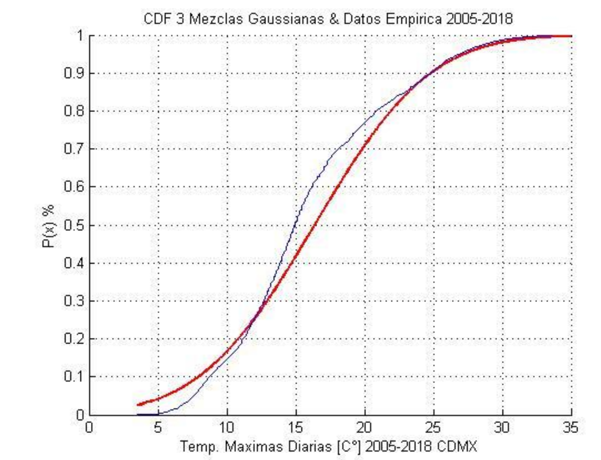
CI1 = 16.4285 16.7587

Normal 2 Mean: 17.86 Sdt= 6.80

CI2 = 17.6971 18.0273

Figures 7. From Gaussian Mixing





Hastings Metropolis Algorithm

Monte Carlo Markov Chain Methods (MCMC) are simulation methods to generate samples of distributions to Posteriori and estimate quantities of interest a posteriori. MCMC methods simulate values successively of a proposed density, whichdoesnotnecessarilyhaveto be akinto a posteriori density.

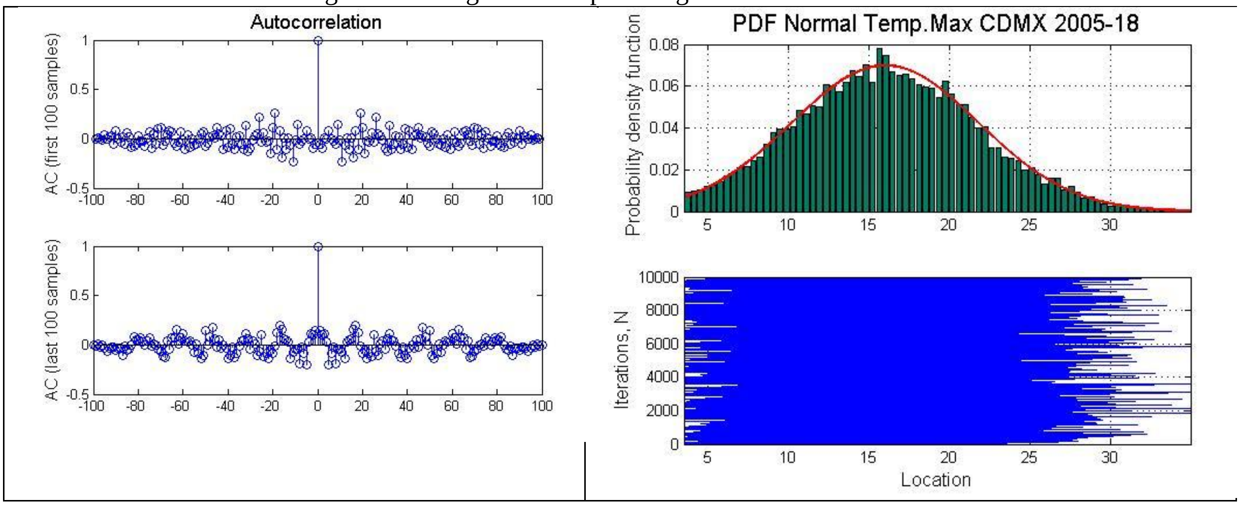
They are notuniquetoBayesianinference, but can be usedtosimúlatevalues from a distributionfromwhichitisnoteasy to generatesamples.

The stochastic simulation and a verycommonproblemaistosimulate a random variable with a givendistribution, over a space of states, althoughitis true thatthere are severalmethodsto do this, they are insufficient.

So Nicolas Metropolis (1953) proposedanalgorithmwhichisbasedon the construction of a Markovchain, ergodicwhich has the distributionfunction as a stationarydistribution.

Therefore, itisenoughtogenerate the Markovchain and from a certainmoment the values generatedwillhave a distributionvery similar to the givendistributionfunction.

Figures 8. Using the Metropolis Algorithm - Markov Chain



Bayesian Point Change Detection

The Algorithm used to observe the biggest changes in the time series of maximum temperatures in the city of mexico, The points of change are abrupt variations in the generative parameters of a sequence of data the average obtained with the one used for the Normal setting. Although frequentist methods have produced online filtering and prediction techniques, a probability distribution of the length of the current "execution", or the time since the last change point, is calculated. The implementation is highly modular so that the algorithm can be applied to a variety of data types.

Results with the means of each year from 2005 to 2018

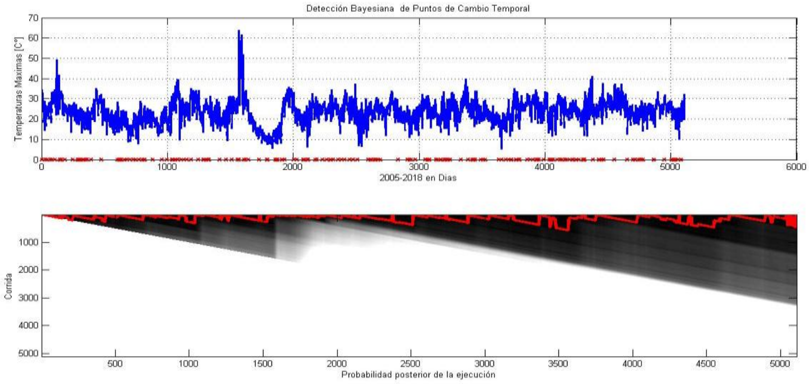


Figure 9. Detection of Temporal Bayesian Point Change with the darkest areas of greatest change and highest probability shown, coinciding from 2010 onwards.

Using the WRF to observe the annual temperature trends in the CDMX

According to the long-term forecast established by climate analysis models, it is expected that Mexico City will achieve several degrees of increase in temperature for the coming years. It is very important to consider your annual analysis and establish growth patterns with the presence of new phenomena that can alter the conditions of the environment, resulting in an increase in temperature. An analysis of the temperature obtained in the last seven years and the present is proposed in order to observe the average annual temperature trends for the CDMX by using the WRF model (version 4.0).

For the analysis, data from meteorological monitoring stations from government sources (SEDEMA, GIRPC) are considered, from which the maximum average annual values of temperatures for each of the stations were obtained. The WRF model is used to make annual projections based on the interpolation algorithms considering the atmospheric and land use characteristics within its arbitrary distribution in the behavior of the temperature and the study area. Obtaining as result the following maps of annual distribution of temperature.

Figure9. Annual Maximum Temperature of the CDMX 2012

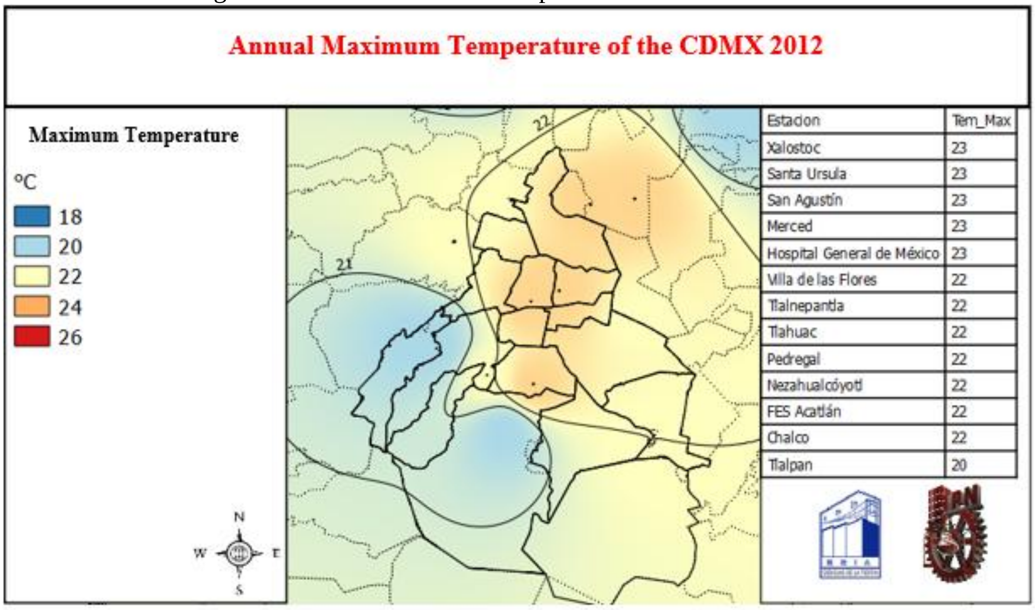


Figure10. Annual Maximum Temperature of theCDMX 2013

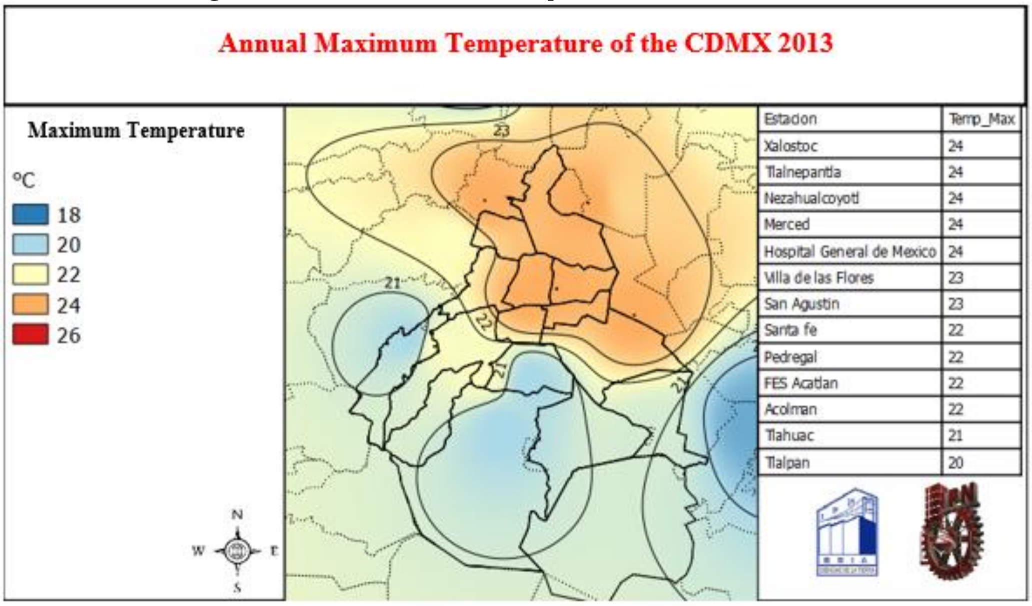


Figure11. Annual Maximum Temperature of the CDMX 2014

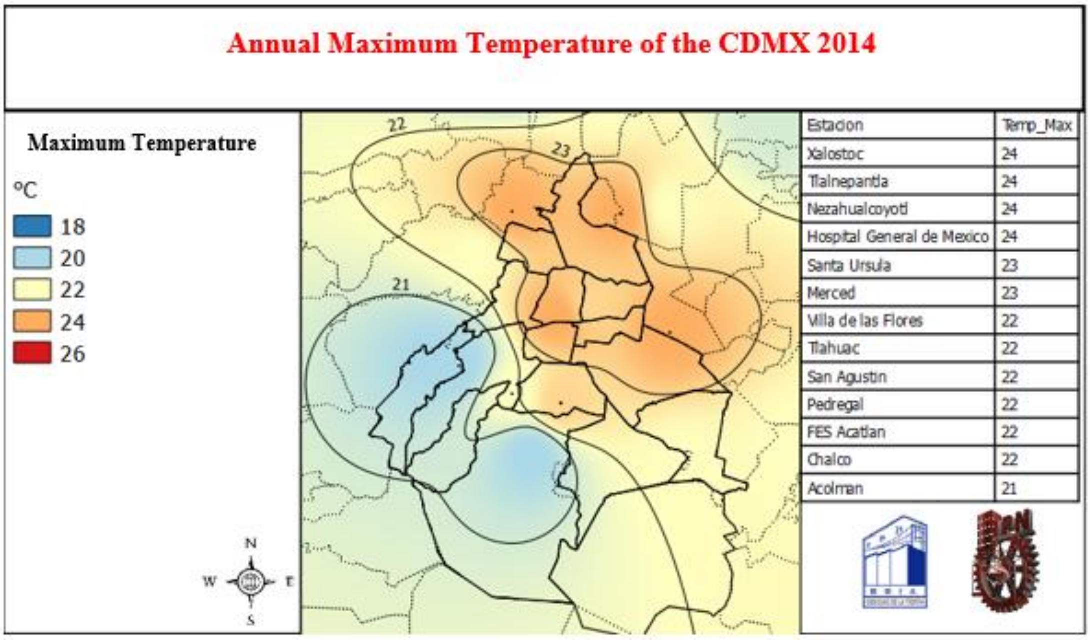


Figure12. Annual Maximum Temperature of the CDMX 2015

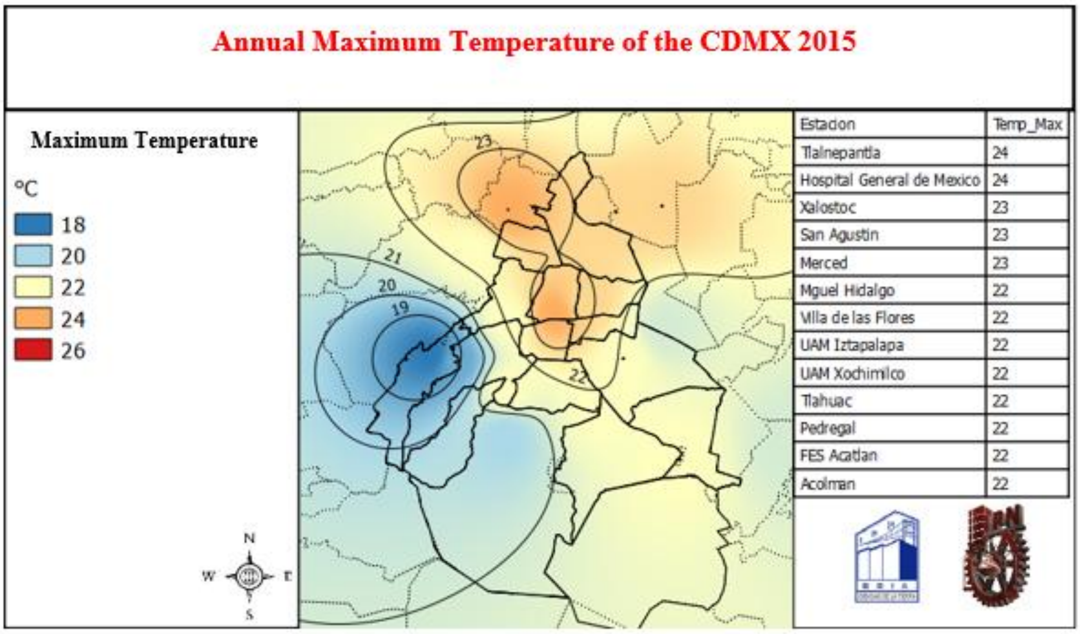


Figure13. Annual Maximum Temperature of the CDMX 2016

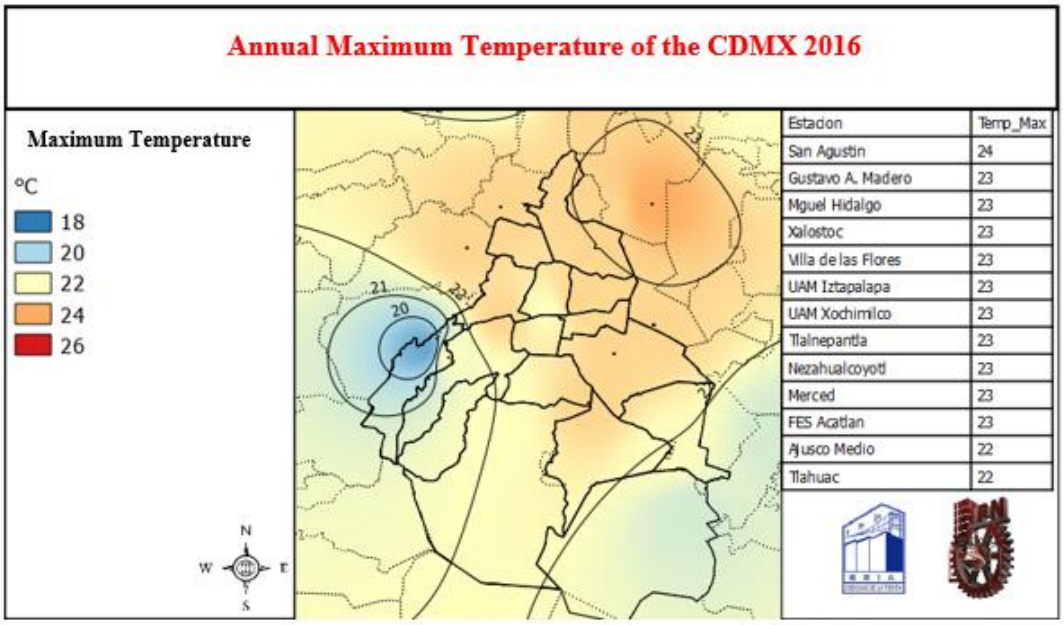


Figure14. Annual Maximum Temperature of the CDMX 2017

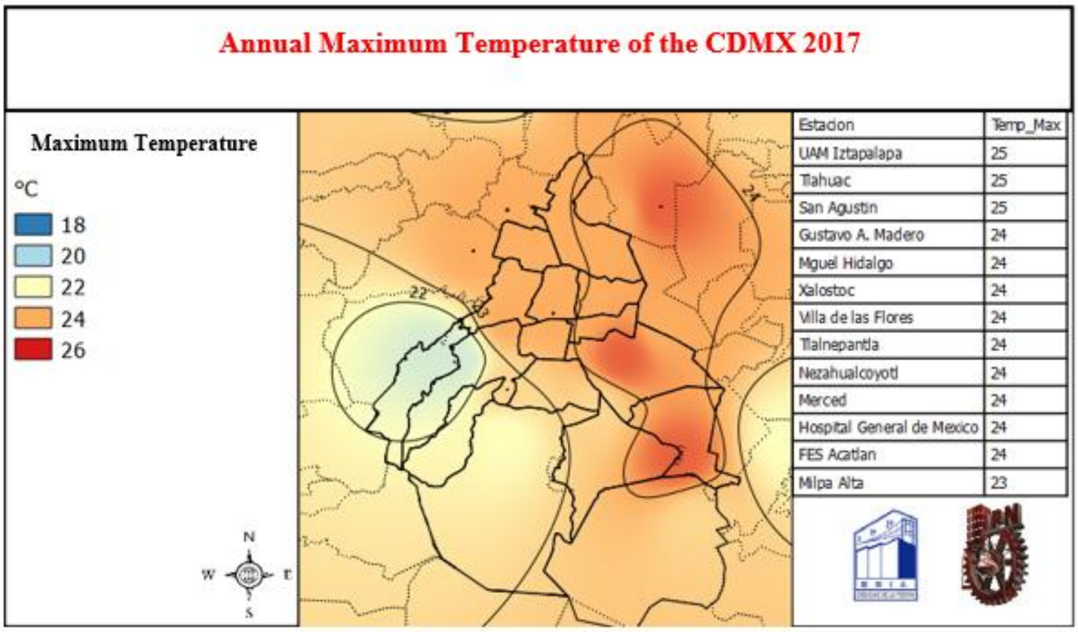


Figure15. Annual Maximum Temperature of the CDMX 2018

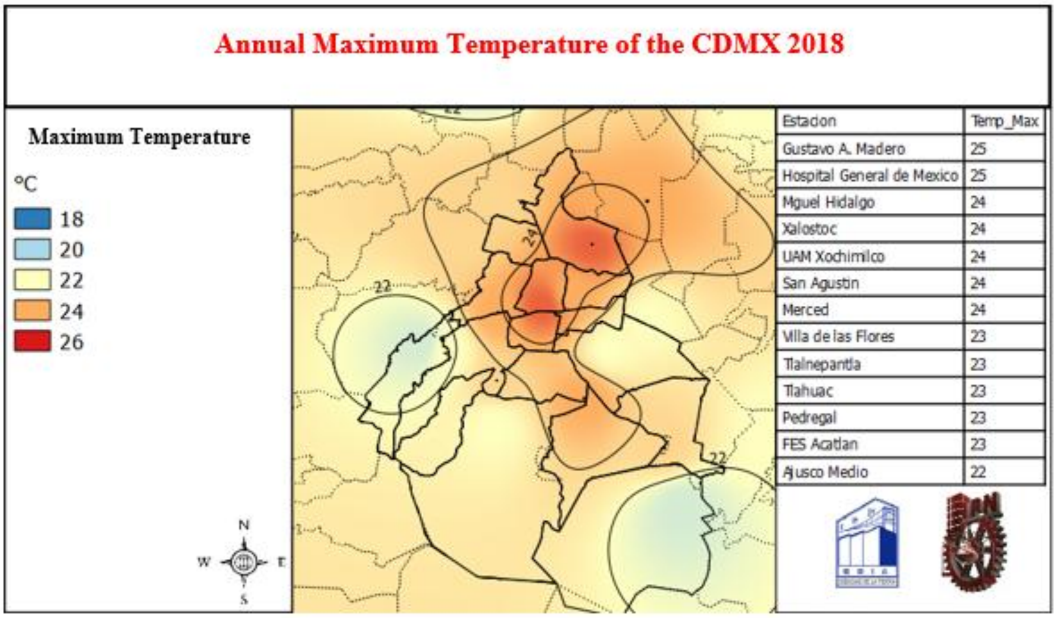
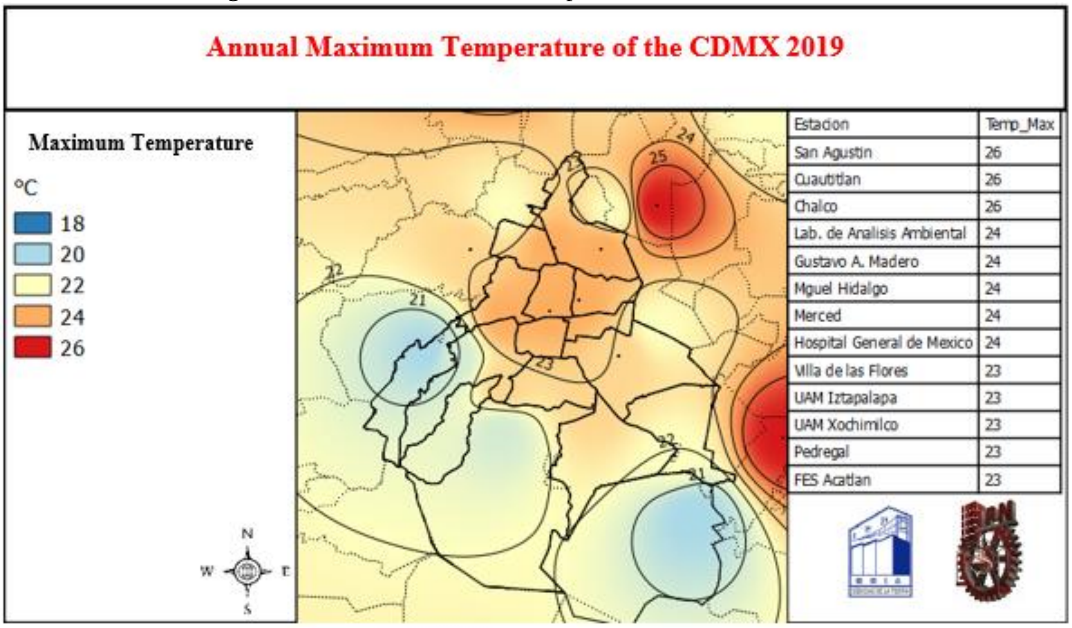


Figure16. Annual Maximum Temperature of the CDMX 2019



Results

The products obtained by the WRF model show an increase in temperature in recent years. Relating the results with the table 9 described above we can see that the behavior of the values of temperature from 2012 to the present shows great similarity, showing with it the reliability that the WRF model can have when making forecast taking data from weather stations that are within the application area.

Figure 14 referring to the year 2017 shows an atypical increase in temperature, which marks the beginning of the increase of the atypical temperature for the next years obtaining a percentage of 1 annual grade which is critical for the CDMX. Although the studies carried out with the model show trends of atypical increase, it gives a guideline to think about external factors that are altering the temperature values. Therefore, it is very important to continue conducting studies that allow us to give an adequate forecast of temperature behavior without forgetting to include in its variability scenarios external factors that cause anomalies of increase in our measurements, such is the case of the heat islands phenomenon [17] that currently has become a big problem for large cities and reflect that CDMX is not exempt from this phenomenon.

Conclusions

According to the method demonstrated and the database of the CDMX official page, the new normal and extreme value distribution functions show the increase of the temperature in the city, we find new functions and when mixing the normal pdfs it gave us even more with greater veracity the result, so also with the WRF program gives us based on the methodology shown the increase of temperature in the CDMX.

References

[1]. A.J. Jakeman, J.A. Taylor, R.W. Simpson, Modeling distributions of air pollutant concentrations - II. Estimation of one and two parameters statistical distributions, Atmos. Environ., 20 (1986) 2435-2447.
[2]. Bayesian Online Changepoint Detection Ryan Prescott Adams, David J.C. MacKay https://arxiv.org/abs/0710.3742
[3]. Forecasting and Estimating Multiple Change-point Models with an Unknown Number of Change-points Gary Koopyand, Simon M. Potterz2006
[4]. Casella, G and Robert, C. Introducing Monte Carlo Methods with R (Use R)
[5]. Compact approximations to Bayesian predictive distributions Edward Snelson, ZoubinGhahramaniICML2005
[6]. Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press, New York, p. 164.
[7]. Hoel, P; Port, S and Stone, C. Introduction to Stochastic Processes
[8]. P.G. Georgopoulous, J.H. Seinfeld, Statistical distributions of air pollutant concentrations, Environ. Sci. Technol., 16 (1982) 401A-416A.
[9]. Roberts, E.M.,1979. Review of statistics of extreme values with applications to air quality data, part II. Applications. Journalof Air Pollution Control Association 29, 733–740.
[10]. Trabajo presentado en el Congreso de la Unión Geofísica Mexicana 2017 Pronostico de Concentraciones de Ozono por Distribuciones de Probabilidad para la CDMX https://www.raugm.org.mx/2017/pdf/constancia.php?clave=809
[11]. Prescott, P., and A. T. Walden, Maximum-likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples, J. Stat. Comput. Simul., 6, 241–250, 1983.
[12]. Otten, A., and M. A. J. Van Montfort, Maximum-likelihood estimation of the general extreme-value distribution parameters, J. Hydrol., 47, 187–192, 1980.
[13]. Zenteno Jiménez José Roberto, Prediction of Concentrations of Ozone Levels in México City using Probability Distribution Functions, International Journal of Latest Research in Engineering and Technology (IJLRET) || Volume 04 - Issue 07 || July 2018 || PP. 35-45
[14]. Zenteno Jiménez José Roberto. A Methodology for Obtaining news Probability Distributions Functions Normal and Extreme Value for Bayesian Inference and Stochastic Mixed Gaussian Case One: For Daily Concentration Data Maximum Ozone. International Journal of Latest Research in Engineering and Technology (IJLRET)|| Volume 04 - Issue 11 || November 2018 || PP. 15-35
[15]. Statistical Modeling and Computation, Dirk P. Kroese – Joshua C.C. Chan, Springer Ed. 2014, Bayesian Inference Chapter 8, 236 page.
[16]. Zenteno Jiménez José Roberto Prediction of Concentrations of Suspended Particle Levels of 2.5 micrometers (PM2.5) in Mexico City with Probability Distribution Functions and its Trend. http://www.ijlret.com/Papers/Vol-05-issue-04/1.B2019025.pdf
[17]. Heat Islands in México City: Perspective from Remote Sensing Satellite Images, Author: Fernando Mireles Arellano, Amanda Oralia Gómez González, Carlos Hernández López; International Journal of Latest Research in Engineering & Technology (IJLRET) || Volume 04 - Issue 10 || October 2018 || PP. 1-12