International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 04 - Issue 11 || November 2018 || PP. 15-35

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

M.Sc. José Roberto Zenteno Jiménez
Geophysical Engineering, National Polytechnic Institute,
Mexico City, ESIA-Unit Ticoman Delegation Gustavo A. Madero
jzenteno@ipn.mx

Abstract: The proposed methodology is oriented data with Gaussian behavior is to adjust the normal distribution function to the data time series maximum data ozone concentration of 2010 - 2017, to give a forecastof concentration, then we use the Bayesian inference for normal data. To validate the model used the following statistical estimators, measuring the root mean square error, mean square error, determination coefficient and prediction approach.

Using new means and variances of the new features of extreme distribution generate two new functions normal distribution a minimum and a maximum, so we obtain a new forecast data for further concentration to 150 ppb. And also having the three functions of normal probability distribution, the first set and two new functions normal distribution introduce stochastic method Gaussian mixture to give a new distribution function generally normal probability. We also use the sum of the three distribution functions to determine normal behavior or trend of maximum ozone concentrations for the city of Mexico probability. The database that was used is from page Mexico City http://www.aire.cdmx.gob.mx/

Keywords: Ozone Pollution, Random Distribution Functions and Variable Extreme, Bayesian Inference, Convolution.

Mexico City has had a very marked history with the evaluation of this air pollutant, ozone which for decades the government and government institutions and Scientific been given the task of trying to reduce the concentrations of the pollutant that both overwhelms the City and the Capitalinos through technological means established and Legislations well as research concerning the components of gasoline used by the automotive and transportation systems within the City. We know that the temperature directly affects the ozone concentration, and this year 2018 has been affected by such temperature rises so is expected for the coming months and represents a health risk,

Season high concentrations of this pollutant begins approximately half February and ends with the first rains June. Also, exposure of high ozone levels is associated with physiological and inflammatory effects in the lungs of healthy young adults who exercise outdoors. Therefore, it is recommended to reduce exposure time outdoors, especially the most vulnerable such as children, the elderly, pregnant women and people with respiratory and cardiovascular problems population, there is a direct relationship between chronic exposure to pollution and increased cases of morbidity and mortality.

According to Mexico Mexican Official Standard (NOM-020-SSA1-2014) recommended concentrations below 0095 ppm for 1 hour average, and less than 0070 ppm for the average of 8 hours (annual maximum). Therefore tropospheric ozone located at surface level in urban areas occurs when nitrogen oxides (NOX) and volatile organic compounds (VOCs) react in the atmosphere in the presence of sunlight. At high concentrations can endanger human health and vegetation.

First the data fit a Gaussian distribution which an analysis of these in a histogram is made and its parameters are obtained by the method of Maximum Likelihood The importance of this distribution model is that it allows many natural phenomena, while mechanisms that underlie much of this phenomenon are unknown, because of the enormous amount of uncontrollable variables involved in them, use the normal model can be justified by assuming that each observation is obtained as the sum of a few independent causes.

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

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 04 - Issue 11 || November 2018 || PP. 15-35

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

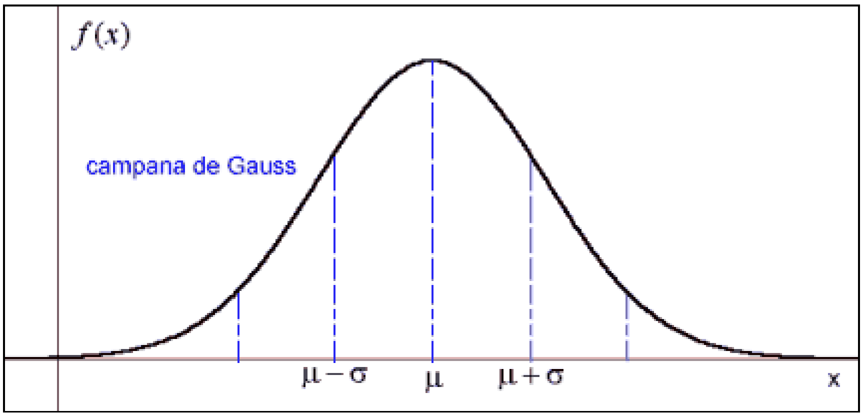


Figure 1. Bell curve or Gaussian density function (Source: Internet)

Now as we get the parameters of estimating a probability distribution function, we use the Maximum Likelihood technique to estimate parameters for adjustment. The maximum likelihood method is a method for obtaining apoint estimatorof a random variable.

Let (X1, ..., Xn) a random sample with a distribution function f (x | θ).

We define the likelihood function as:

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

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

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

So the maximum likelihood estimator is defined as:

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

Through the above description parameters of a normal distribution which are obtained by doing the following:

By the method of Maximum Likelihood, the likelihood function is:

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

Logarithms and differentiating with respect to parameters to be estimated have a system of equations as follows:

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

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

With the solution:

$$\mu = xy\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x-\mu)^2 \tag{7}$$

With the first set Normal we get

$$Normal(\mu,\sigma^2) \tag{8}$$

Now we are looking for extreme values are those that want to know the probability of occurrence so we use Bayesian inference to find this probability with a new distribution function will be part of the new features of normal distribution and function extremely variable or GeV, our new target will be the average.

Bayesian inference

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

$$P(Parametros|Datos) = \frac{P(Datos|Parametros) * P(Parametros)}{P(Datos)} \approx FVerosimilitud * PDFPriori \tag{9}$$

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

$$X|\theta \sim N(\theta,\sigma^2) \tag{10}$$

Where sigma is previously known and PDF to Priori is

$$\theta \sim N(\mu,\tau^2) \tag{11}$$

Here mu and tao are also known, we are looking for n samples of the data observed in the case of ozone peak values or above 150 ppb, the case of particulate PM10 above 120 microg / m3, the case PM2. 5 above 65 microg / m3 and in the case of maximum temperatures is the whole data sample and thus obtain the new Normal distribution function with the new required parameter:

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

Now these data contain noise, no nulls or zeros from the adjustment process and although it has good approximation is not fit quite right data so which produce inaccuracy to the distribution function Normal, therefore apply a random noise with a uniform distribution with the length of the terms of the time series data, we now apply a setting with the function extreme value distribution (GEV) to find the parameters that fit even better this data with the random noise.

$$GEV(\mu,\sigma,k) \tag{13}$$

Which a GEV fits with this **uniform random distribution**

$$GEVa(Xa,\mu,\sigma,k1) \tag{14}$$

And thereafter a GEVA is generated (with random parameters GEVa)

$$\boldsymbol{GEVA(\mu a, \sigma a, k1)} \tag{15}$$

We also adjusted now one GEV of the input data, here is where the theory of Extreme Value comes, and now seek a new distribution function, and is where the new equation is applied depending on the properties of the parameters that were previously obtained:

GEV 1

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

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

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

We now get a second equation for the new parameters

GEV 2

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

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

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

Basic theory of equations shown above

We study if the base is valid to this sum, we have for the case of two sums of the parameters of a GEV of the adjusted data, another random GEV and the parameters of a Normal Bayesian the probabilities for values above the target, now the Central Limit Theorem tells us the following:

The central limit theorem or central limit theorem states that, in very general terms, if Sn is the sum of n independent random variables and variance nonzero but finite, then the distribution function Sn “approaches well”a normal distribution

Teorema del límite central: Sea $X_1, X_2, ..., X_n$ un conjunto de variables aleatorias, independientes e idénticamente distribuidas con media μ y varianza $0 < \sigma^2 < \infty$. Sea

$$S_n = X_1 + \cdots + X_n$$

Entonces

$$\lim_{n \to \infty} \Pr\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right) = \Phi(z)\,.$$

Figure2. Central Limit Theorem

You want to calculate a new feature probability distribution with the sum of the three distribution functions prior probability, but we know that some other techniques would be used to do this, plus there are two functions of the same species and not so with the results of Theorem we can use the properties of the normal distribution function to find an equation for this.

Thus we enunciate property: The sum is normally distributed with. Conversely, if two independent random variables whose sum normally distributed, should be normal. $U = X + Y \approx N(\mu x + \mu y, \sigma^2 x + \sigma^2 y)$

We are trying sums of n probability distribution functions, so we can make an equivalent with her stockings to start well, we can also take the theorem of convergence in r - th average:

It's a sequence of random variables Xn is said to converge to X in r - th media if the following is true:

$$\lim_{n\to\infty} E|Xn - X|^r = 0 \tag{22}$$

This ensures the convergence of random variables or data, then we can do this, **where** $E(X)_i = x_i$

$$En(X) = \sum_{i=1}^{n} \frac{x_i}{n} = \sum_{i=1}^{n} \frac{E(X)_i}{n} \tag{23}$$

Treat mean sums as the sum of the properties of normal data or normal behavior, but with the observation that this on n, in this case the number of averaging times have.

$$E(X) = \frac{E(x_1 + x_2 + x_3 + \cdots + x_n)}{n} = \frac{\mu_1 + \mu_2 + \mu_3 + \cdots + \mu_n}{n} = E(X)_i \tag{24}$$

We now proceed to also add the half-n as being a function of x, then we do the following:

$$E(X) = \sum_{i=1}^{n} \frac{x_i}{n} = \int_0^n \frac{x_i}{n} dx = \frac{x_i}{n} \int_0^n dx = \frac{x_i}{n} [x]_0^n = x_i = E(X)_i \tag{25}$$

Or

$$E(X) = E\left[\frac{1}{n} \sum_{i=1}^{n} x_i\right] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{n\mu_i}{n} = \mu_i$$

And so

$$E(X)_i = \sum_{i=1}^{n} \frac{\mu_i}{n} \tag{26}$$

Now the variance. The variance can be written as and integrate up to n

$$V(X)_i = \sum_{i=1}^{n} \frac{(n-i)x_i}{n(n-1)} = \frac{(n-i)x_i}{n(n-1)} \int_0^n dx = \frac{(n-i)x_i}{n(n-1)} [x]_0^n \tag{27}$$
$$= \sum_{i=1}^{n} \frac{(n-i)x_i}{(n-1)} = \sum_{i=1}^{n} \frac{nx_i - ix_i}{n-1}$$

Now notice that is divided by n-1 so started the sum of x between n-1 since n = 1 indetermina

$$\frac{1}{n-1}\sum_{i=1}^{n-1} nx_i - ix_i = \frac{1}{n-1}\sum_{i=1}^{n-1}\left(x_i - i\frac{x_i}{n}\right) = \frac{1}{n-1}\sum_{i=1}^{n-1}(x_i - i\mu_i) = \tag{28}$$
$$= \frac{1}{n-1}\sum_{i=1}^{n-1}\sigma_i = \sum_{i=1}^{n-1}\frac{\sigma_i{}^2}{(n-1)} = \sum_{i=1}^{n-1}\frac{\sigma_i{}^2}{n-1}$$

see also

$$conn = 2 \tag{29}$$
$$\frac{1}{n-1}\sum_{i=1}^{n-1}\sigma_i = \frac{1}{2-1}[\sigma_1]$$

$$conn = 3 \tag{30}$$
$$\frac{1}{n-1}\sum_{i=1}^{n-1}\sigma_i = \frac{1}{3-1}[\sigma_1 + \sigma_2]$$

$$conn = 4 \tag{31}$$
$$\frac{1}{n-1}\sum_{i=1}^{n-1}\sigma_i = \frac{1}{4-1}[\sigma_1 + \sigma_2 + \sigma_3]$$

So we are having these results we can say based on the Weak law of large numbers as follows:

Khintchine form (1929). Be$\{Xm\}$ a sequence of independent random variables and also distributed with mean and variance $E(Xm) = \mu Var(Xm) = \sigma^2$

Then we have:

$$\lim_{n\to\infty} P[|Xm - \mu| > \varepsilon] = 0 \tag{32}$$

Where Xm is given by

$$E(Xm)_i = \sum_{i=1}^{n}\frac{\mu_i}{n} = \frac{\mu_1 + \mu_2 + \mu_3 + \cdots + \mu_i}{n} = \frac{\mu_i}{n} = \mu \tag{33}$$

$$V(Xm)_i = \frac{1}{n-1}\sum_{i=1}^{n-1}\sigma_i = \frac{\sigma_1 + \sigma_2 + \sigma_3 + \cdots + \sigma_i}{n-1} = \frac{\sigma_i}{n-1} = \frac{\sigma_i}{n} \, or \, \frac{\sigma_i^2}{n-1} \tag{34}$$

A demonstration using the Chebyshev inequality which is

$$P(|Xm - \mu| > k\sigma) \le \frac{1}{k^2} \tag{35}$$

Using Chebychev Inequality as with n> 1

$$P\left[|Xm - E(Xm)| > k\sqrt{V(Xm)}\right] \le \frac{1}{k^2} \tag{36}$$

$$P\left[|Xm - \mu| > \frac{k\sigma}{\sqrt{n-1}}\right] \le \frac{1}{k^2} \tag{37}$$

Here now raising and clearing to k, so it turns out that $\varepsilon = \frac{k\sigma}{\sqrt{n-1}} sok^2 = \frac{(n-1)\varepsilon^2}{\sigma^2}$

$$\boldsymbol{P[|Xm - \mu| > \varepsilon] \leq \frac{\sigma^2}{(n-1)\varepsilon^2}} \quad \textbf{(38)}$$

$$\boldsymbol{P[|Xm - \mu| > \varepsilon] \leq \frac{V(Xm)}{\varepsilon^2}} \quad \textbf{(39)}$$

With V (Xm) = 0$V(Xm) \to n \to \infty$when $\frac{\sigma_i}{n-1}$ ór$\frac{\sigma_i}{n}$ and n $n \to \infty$the expretion is 0

Online references (9)

With this we prove that when n tends to infinity this division is zero which proves the law and that the expressions given if you give a single mean and standard deviation and since Xm is the data set maximum concentrations of daily ozone alone match three probability distribution functions, Chebyshev inequality is a result that offers a lower probability that the value of a random variable with finite variance is at a certain distance from their average expectation or dimension.

Another consequence of the theorem is that for each distribution with mean μ and standard deviation σ finite, at least half of their values are concentrated in the interval)$(\mu - \sqrt{2}\sigma, \mu + \sqrt{2}\sigma$and so we finally have the mean and standard deviation for sums of several distribution functions, now Moments of GEV$L(\lambda_{r)})$ the probability weighted moments sample must be the first moment is the mean and the second moment is the standard deviation or variance.

Now the parameters of the function extreme value or GEV we have:

Table 1. Parameters of the extreme variable function (GEV)

Location parameter	$fx0(x) = f(x - x0)$ It's called location parameter. Examples of location parameters include the mean, median and mode.
Scale parameter	The normal distribution has two parameters: a parameter location parameter mu and sigma scale. In practice, often the normal distribution is parameterized in terms of the square scale corresponding to the variance of the distribution.
Shape parameter	Many estimators measure the location or scale; However, there are also estimates for parameters of shape. More simply, they can be estimated in terms of the highest moments, using the method of moments, as in the asymmetry (third time) or the kurtosis (fourth time), if higher moments are defined and finite.

Thus function GEV end or can combine this value as follows which are the new parameters of the new features of extreme distribution:

$$\boldsymbol{GEV(\sum_{i=1}^{n} \frac{\mu_i}{n}, \frac{1}{n-1} \sum_{i=1}^{n-1} \sigma_i, k)} \quad \textbf{(40)}$$

With

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

According to the above definition the shape parameter we can deduce the method of higher moments is between two given that there are two parameters lower order composing this, according to the equations obtained, if we would not use a Newton- Raphson for obtain these parameters, so then we have:

$$\mu - \frac{\sigma}{k} = 0 yasik = \frac{\sigma}{\mu} \quad (41)$$

$$k = \frac{\sum_{i=1}^{n-1} \frac{\sigma_i}{n-1}}{\sum_{i=1}^{n} \frac{\mu_i}{n}} = \sum_{i=1}^{n} \frac{\frac{\sigma_i}{n-1}}{\frac{\mu_i}{n}} = \sum_{i=0}^{n} \frac{\sigma_i n}{\mu_i (n-1)} = \sum_{i=0}^{n} \frac{n k_i}{(n-1)} \quad (42)$$

Noting Series can reach the equivalent sum and thus we obtain an expression for the shape parameter

$$k = \sum_{i=0}^{n} \frac{k_i}{n} \quad (43)$$

Then we have the following

$$GEV\left(\sum_{i=1}^{n} \frac{\mu_i}{n}, \frac{1}{n-1} \sum_{i=1}^{n-1} \sigma_i, \sum_{i=0}^{n} \frac{k_i}{n}\right) \quad (44)$$

Now we have this range as we saw above, our shape parameter is greater or less than zero, also parameters mu and sigma are in ppb
The shape parameter is dimensionless.

$$\left(-\infty, \mu - \frac{\sigma}{k}, +\infty\right) \quad (45)$$

But our measurements are finite so

$$\left[\min(O_3), \mu - \frac{\sigma}{k}, \max(O_3)\right] \quad (46)$$

This should obey the following to calculate the new features GEV distribution

$$\left[\min(O_3) \le \mu - \frac{\sigma}{k} \le \max(O_3)\right] \quad (47)$$

Without exceeding these limits.

We have the following probability distribution functions, normal Bayesian data GEV GEV Randomly and therefore are 3 functions that have so the first two sums two probability distribution functions are:

Now new means and variances of the features found now only generated the new features extreme normal distribution

$$Normal1\left(E(GEV1), \sqrt{Var(GEV1)}\right) \quad (48)$$

$$Normal2\left(E(GEV2), \sqrt{Var(GEV2)}\right) \quad (49)$$

And so you have the three functions Gaussian probability distribution

$$Normal(\mu,\sigma^2) \tag{50}$$
$$Normal1(E(GEV1),\sqrt{Var(GEV1)}) \tag{51}$$
$$Normal2(E(GEV2),\sqrt{Var(GEV2)}) \tag{52}$$

Table 2. Probability distribution functions GEV and Normal

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

Adjustment indicators

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

Table 3. Indicators Set

Indicator	Equation
Root Mean Square Error	$RMSE=\sqrt{\left(\frac{1}{N-1}\right)\sum_{i=1}^{N}(Pi-Oi)^2}$
Mean Square Error	$RMSE=\left(\frac{1}{N}\right)\sum_{i=1}^{N}(Pi-Oi)^2$
Coefficient of Determination	$R^2=\left(\frac{\sum_{i=1}^{N}(Pi-P)(Oi-O)}{NS_pS_o}\right)^2$
Accuracy Prediction	$AP=\frac{\sum_{i=1}^{N}(Pi-Oi)^2}{\sum_{i=1}^{N}(Oi-Oi)^2}$

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

Normal operations between new found another comparison approach

Now we apply the sum of the three normal convolution to determine a trend and also an approximation.

two functions of normal distribution are as follows

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\left(-\frac{(x-\mu)^2}{2a^2}\right)} \tag{53}$$

$$g(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\left(-\frac{(x-\lambda)^2}{2b^2}\right)} \tag{54}$$

Now we know that convolution is:

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(\eta) g(t-\eta) d\eta \tag{55}$$

Be

$$f(x) = e^{-\frac{x^2}{2a^2}} \tag{56}$$

$$g(x) = e^{-\frac{x^2}{2b^2}} \tag{57}$$

$$(f * g)(t) = \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2a^2}} e^{\left(-\frac{(x-t)^2}{2b^2}\right)} dt \tag{58}$$

$$e^{-\frac{x^2}{2b^2}} \int_{-\infty}^{+\infty} e^{\left[-\frac{t^2}{2a^2}+\frac{xt}{b^2}-\frac{t^2}{2b^2}\right]} dt \tag{59}$$

$$e^{-\frac{x^2}{2b^2}} \int_{-\infty}^{+\infty} e^{\left[-\frac{t^2}{2}\left(\frac{a^2+b^2}{a^2b^2}\right)+\frac{xt}{b^2}\right]} dt \tag{60}$$

Making a change of variables

$$u = \frac{t}{\sqrt{2}} \frac{\sqrt{(a^2+b^2)}}{ab} \tag{61}$$

A) Yes

$$e^{-\frac{x^2}{2b^2}} \int_{-\infty}^{+\infty} e^{\left[-u^2+\frac{x}{b^2}\frac{\sqrt{2}abu}{\sqrt{(a^2+b^2)}}\right]} \frac{\sqrt{2}ab}{\sqrt{(a^2+b^2)}} du \tag{62}$$

$$\frac{\sqrt{2}ab}{\sqrt{(a^2+b^2)}} e^{-\frac{x^2}{2b^2}} \int_{-\infty}^{+\infty} e^{[-u^2+pu]} du \tag{63}$$

With

$$p = \frac{\sqrt{2}ax}{b\sqrt{(a^2+b^2)}} \tag{64}$$

Now this part is an integral Gaussian

$$I(p) = \int_{-\infty}^{+\infty} e^{[-u^2+pu]} du \tag{65}$$

Let's see we complete the square

$$I(p) = \int_{-\infty}^{+\infty} e^{\left[-\left(u-\frac{p}{2}\right)^2+\frac{p^2}{4}\right]} du = e^{\frac{p^2}{4}} \int_{-\infty}^{+\infty} e^{\left[-\left(u-\frac{p}{2}\right)^2\right]} du \tag{66}$$

And with

$$w = u - \frac{p}{2} \tag{67}$$

$$e^{\frac{p^2}{4}}\int_{-\infty}^{+\infty} e^{[-(w)^2]}dw \tag{68}$$

Now let's apply my alternative method to resolve this Gaussian Integral, you can resort to the conventional method for variable change, polar or otherwise

Now we have, Fubini theorem and Cartesian coordinates what we want is just the area under the curve is generated

$$\int_{-\infty}^{+\infty} e^{[-(x)^2]}dx = \left(\int_{a=0}^{b=\infty} e^{-x^2}dx\right)\left(\int_{a=0}^{b=\infty} e^{-x^2}dx\right) \tag{69}$$

$$I(x) = \int_{a=0}^{b=\infty} e^{-x^2}dx \tag{70}$$

Thus we have

$$I(x)^2 = \left(\int_{a=0}^{b=\infty} e^{-x^2}dx\right)^2 = \left(\int_{a=0}^{b=\infty} e^{-x^2}dx\right)^2 \tag{71}$$

Now equating the two results

$$I(x)^2 = \pi \tag{72}$$

Which has as

$$I(x) = \sqrt{\pi} = I(w) \tag{73}$$

$$e^{\frac{p^2}{4}}(\sqrt{\pi}) \tag{74}$$

$$e^{\left(\frac{\sqrt{2}ax}{b\sqrt{(a^2+b^2)}}\right)^2}\sqrt{\pi} = e^{\left(\frac{2a^2x^2}{4b^2(a^2+b^2)}\right)}\sqrt{\pi} \tag{75}$$

$$\frac{\sqrt{2}ab}{\sqrt{(a^2+b^2)}}e^{-\frac{x^2}{2b^2}} * e^{\left(\frac{2a^2x^2}{4b^2(a^2+b^2)}\right)}\sqrt{\pi} \tag{76}$$

$$\frac{\sqrt{2\pi}ab}{\sqrt{(a^2+b^2)}}e^{-\frac{x^2}{2(a^2+b^2)}} \tag{77}$$

Finally

$$(f*g)(x) = \frac{\sqrt{2\pi}ab}{\sqrt{(a^2+b^2)}}e^{-\frac{x^2}{2(a^2+b^2)}} = \frac{\sqrt{2\pi}ab}{\sqrt{(a^2+b^2)}}e^{-\frac{(x-\mu-\lambda)^2}{2(a^2+b^2)}} \tag{78}$$

So the new Gaussian distribution function is:

$$\psi(x) = \frac{\sqrt{2\pi}ab}{\sqrt{(a^2+b^2)}}e^{-\frac{(x-\mu-\lambda)^2}{2(a^2+b^2)}} \tag{79}$$

With b Standard deviations for each distribution function of Entry and First Gaussian, now we have to get the Third Addition with this Convolution

$$\frac{\sqrt{2}ab}{\sqrt{(a^2 + b^2)}} e^{-\frac{x^2}{2b^2}} * e^{\left(\frac{2a^2x^2}{4b^2(a^2+b^2)}\right)} \sqrt{\pi} \tag{80}$$

$$\frac{\sqrt{2\pi}ab}{\sqrt{(a^2 + b^2)}} e^{-\frac{x^2}{2(a^2+b^2)}} \tag{81}$$

Finally

$$(f * g)(x) = \frac{\sqrt{2\pi}ab}{\sqrt{(a^2 + b^2)}} e^{-\frac{x^2}{2(a^2+b^2)}} = \frac{\sqrt{2\pi}ab}{\sqrt{(a^2 + b^2)}} e^{-\frac{(x-\mu-\lambda)^2}{2(a^2+b^2)}} \tag{82}$$

So the new Gaussian distribution function is:

$$\psi(x) = \frac{\sqrt{2\pi}ab}{\sqrt{(a^2 + b^2)}} e^{-\frac{(x-\mu-\lambda)^2}{2(a^2+b^2)}} \tag{83}$$

B with standard deviations of each distribution function the input and the first gaussian, now we have to get the third sum with this convolution

The third convolution is

$$\psi(x) = \frac{\sqrt{2\pi}}{\sqrt{\frac{(2c^2+a^2+b^2)}{2c^2(a^2+b^2)}}} e^{-\frac{\sqrt[4]{2}(x-\mu-\lambda-\gamma)^2}{8c^8\sqrt{\frac{(2c^2+a^2+b^2)}{2c^2(a^2+b^2)}}}} \tag{84}$$

Stochastic method of Gaussian Mixtures

The clustering model most closely related to statistical distributions is based on the model. Groups can then easily be defined as the most likely objects belonging to the same distribution. A convenient property of this approach is that this is very similar to the way in which the artificial data sets are generated: by random sampling of a distribution of objects.

One of the most prominent methods is known as Gaussian mixture model (used in the Expectation-Maximization algorithm). Here, the data set is usually modeled with a fixed number (to avoid overfitting) of Gaussian distributions is initialized randomly, and whose parameters are iteratively optimized to better classify the data set. This will converge to a local optimum, multiple runs may produce different results. For a grouping well, data is often assigned to the Gaussian distribution more likely to belong to such a grouping.

Grouping based on distributions produces complex models that can capture groups correlation and dependencies between attributes. Even so, these algorithms put an extra burden on the user: for many real data sets can be no definite mathematical model.

Examples of clustering using Expectation-Maximization (EM)

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

For example, the data modeling human height, the height is typically modeled as a normal for each gender with an average for men and women distribution. Since only the height data and no gender assignments for each data point, the distribution of all heights follow the sum of two (different average) normal distributions scale (different variance) and displaced. A model that makes this assumption is an example of a Gaussian mixture model (GMM), although either a GMM may have more than two components. The estimation of the parameters of individual components normal distribution is a canonical problem in data modeling with GMM.

The GMM is widely used to group and estimate the physical density. However, they have a wide range of applications in other fields, such as modeling meteorological observations geoscience (Zi, 2011), certain autoregressive models or some noise time series.

If you believe your data come from a different set of normal distributions, then the GMM is suitable analysis tool. The normal distribution is an underlying assumption, which means that although it is assumed that the distributions are Gaussian, or may not be. In some cases, you may not have, but use logic or prior knowledge to assume that your data are normally distributed. Therefore, the models created from a GMM method involve some level of uncertainty.

Gaussian mixture model means that each data point sets (randomly) from one of the data classes C, likely to be drawn from the class i, and each class is distributed as Gaussian with mean and standard deviation. Given a set of data extracted from said distribution estimate these unknown parameters seek.$p_i \mu_i \sigma_i$

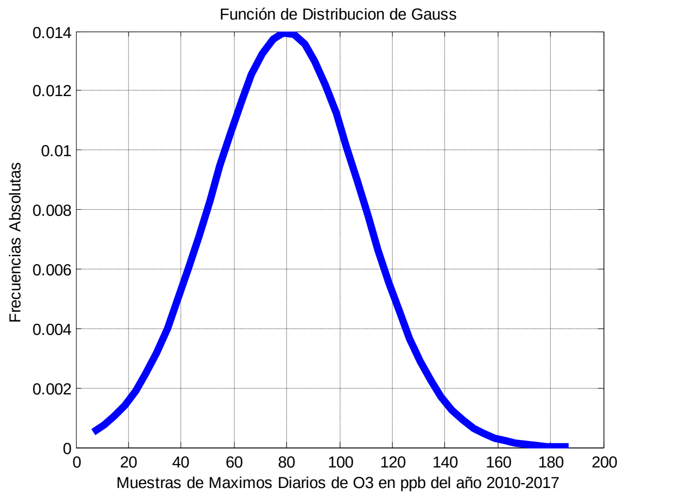
The algorithm used here is to estimate EM (Expectation Maximization). In short, if we knew the kind of each of the N input data points, we could separate them, and using Maximum Likelihood to estimate the parameters of each class. This is the step that makes selections (soft) class (unknown) for each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

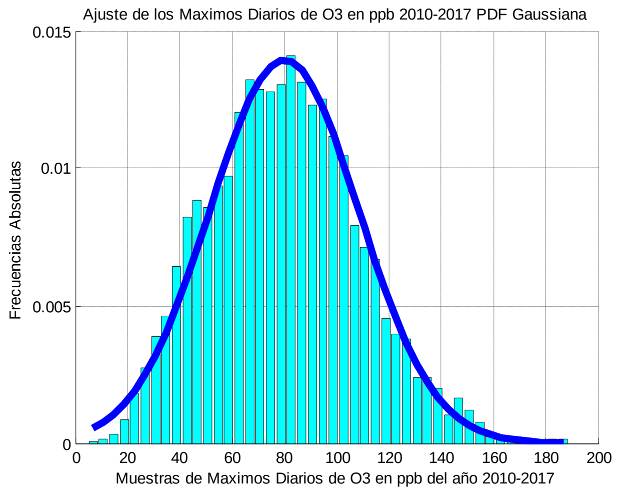
$$
\begin{aligned}
\phi_j &:= \frac{1}{m}\sum_{i=1}^{m} w_j^{(i)}, \\
\mu_j &:= \frac{\sum_{i=1}^{m} w_j^{(i)} x^{(i)}}{\sum_{i=1}^{m} w_j^{(i)}}, \\
\Sigma_j &:= \frac{\sum_{i=1}^{m} w_j^{(i)} (x^{(i)}-\mu_j)(x^{(i)}-\mu_j)^T}{\sum_{i=1}^{m} w_j^{(i)}}
\end{aligned}
$$

Figura4. basic equations of EM algorithms (Source:http://mccormickml.com/2014)

General Forecast of Maximum Several Ozone Exceedances for 2018

Table 4. Probability distribution functions GEV Maximum Daily Ozone [O3] of 2010-2017





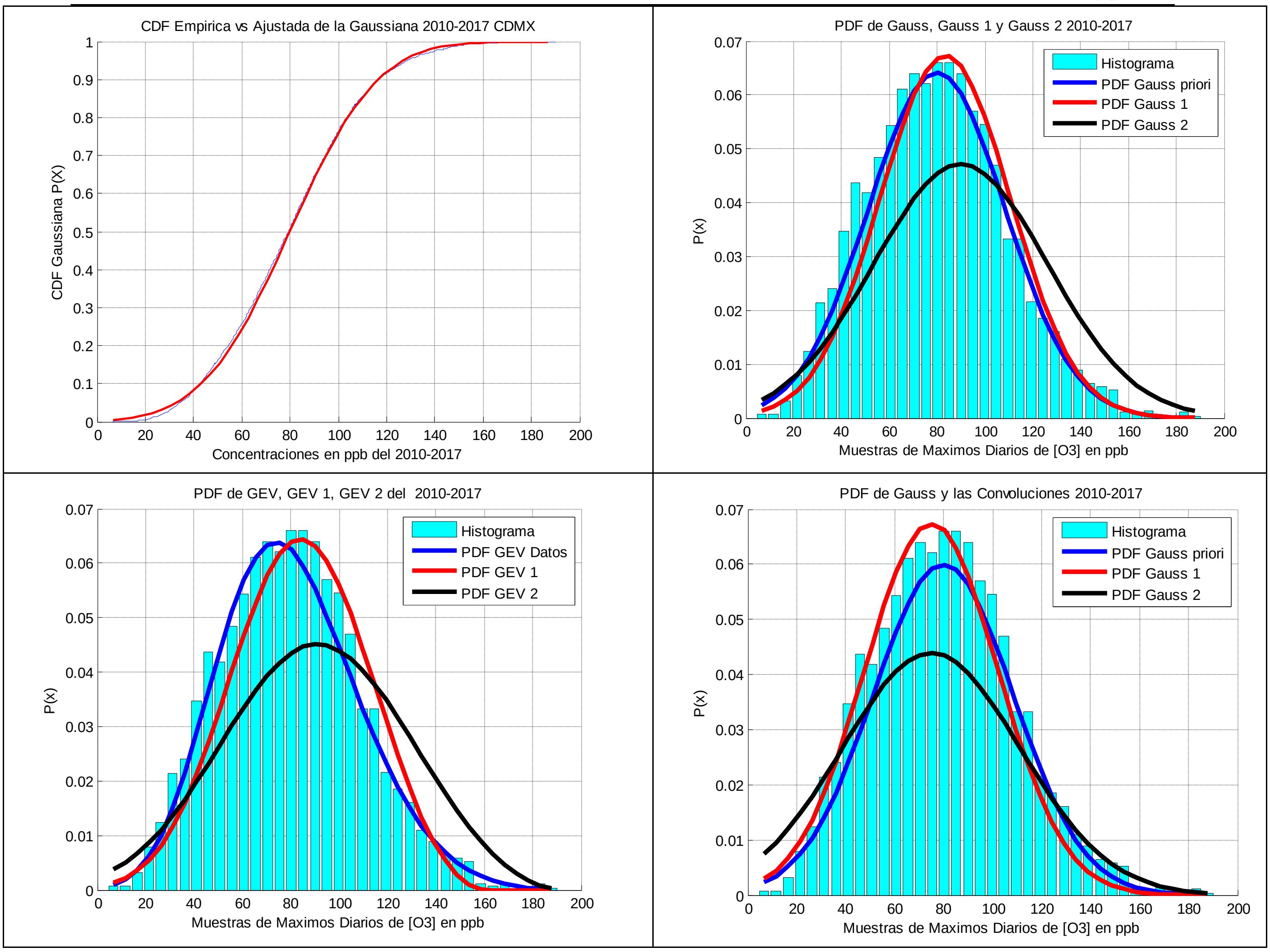


Table 5. QQ plots of GEV and Gaussian Mixture Maximum Daily Ozone [O3] of 2010-2017

Preliminary adjustment Normal

MSE = 0.00006095
RMSE = 0.0079
AP = 0.999
R2 = 0.9890

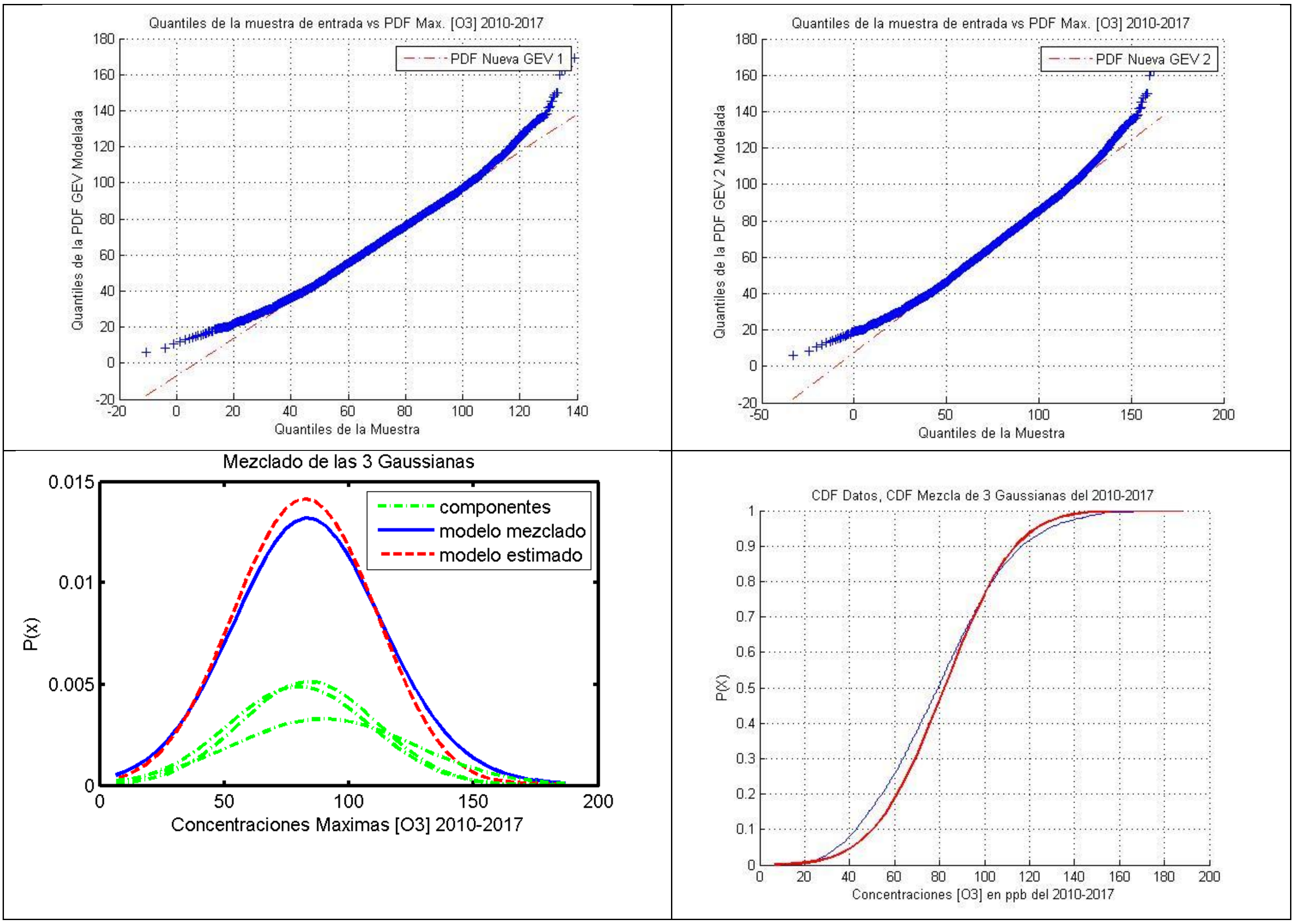


Table 6. Probability distribution functions GEV Maximum Daily Ozone [O3] of 2010-2017 Consider the table above 150 ppb of [O3]

PDF	quantile	From 2010-18 to 365 days	From 2010-18 to 3285 days Joined
Gaussian Normal	.9933	3	22
Gaussian Bayes	0.9998	1	1
Gaussian 1 (GEV 1)	0.9951	2	16
Gaussian 2 (GEV 2)	.9472	19	---
Gaussian convolution 1	.9966	2	11
Gaussian convolution 2	.9708	11	---
GEV	0.9866	5	---
GEV 1	0.9975	1	9
GEV 2	.9493	18	---
2 Gaussian Mixtures	0.9765	9	---
3 Gaussian Mixtures	.9972	1	10

Table with the final choice.

PDF	From 2010-18 to 365 days	From 2010-18 to 3285 cumulative days
Gaussian 1 (GEV 1)	2	16
Gaussian convolution 1	2	11
GEV 1	1	9
3 Gaussian Mixtures	1	10

Table 7. Hosiery and CI is the confidence interval

PDF	Media in ppb	Standard deviation in ppb
Gaussian Normal	79.88	28.60 C1=78.84 80.92
Gaussian Bayes	79.81	28.59
Gaussian 1 (GEV 1)	83.37	27.25 CI= 82.33 84.41
Gaussian 2 (GEV 2)	89.87	36.39 CI= 88.83 90.91
Gaussian Adds 1	75	27.25
Gaussian Add 2	75	36.39
2 Gaussian Mixtures	86.41	32.38
3 Gaussian Mixtures	82.12	24.78

Table 8. Parameters obtained from GEVs

PDF	Media in ppb	Variance in ppb
GEV 1 k2 = -0.3164 SigmaSD = 27.7729 MupostMean = 74.1718	83.37	742.56
GEV 2 k2 = -0.3164 SigmaSD2 = 37.0936 MupostMean2 = 77.5822	89.87	13248

Possible Numbers Days Pre environmental contingencies to Mexico City for this 2018 are:
1 to 16 days over 150 ppb passing daily Maximum ozone [O3]

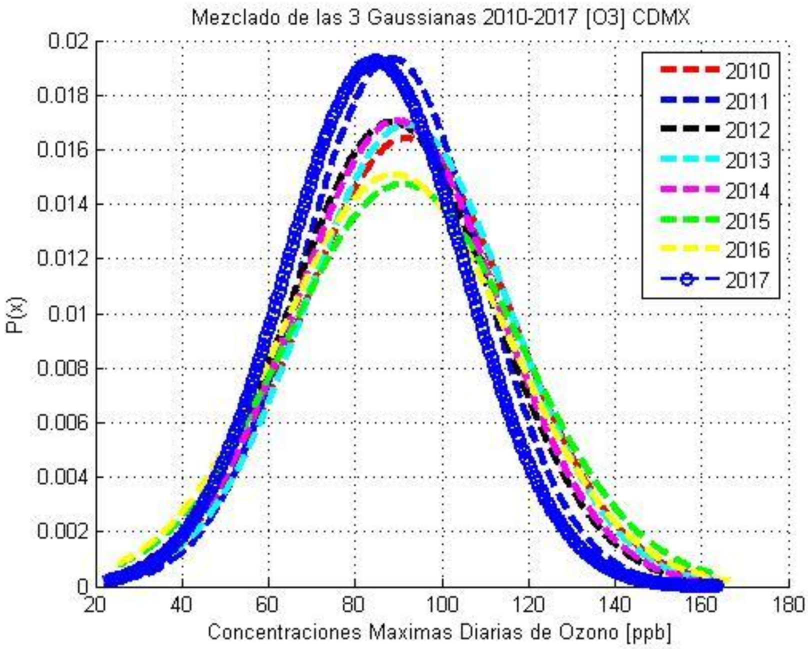
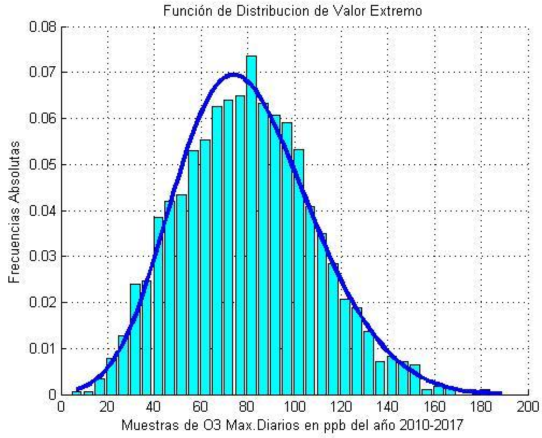
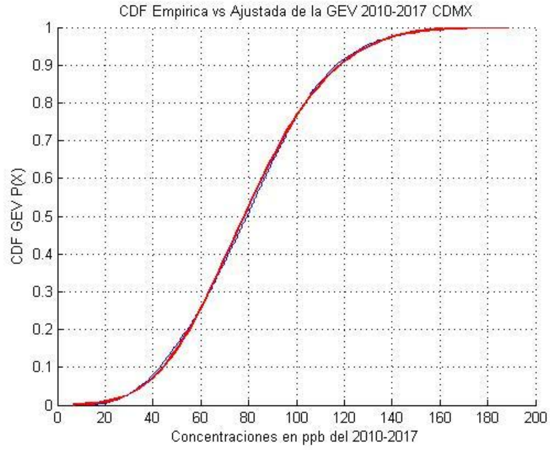
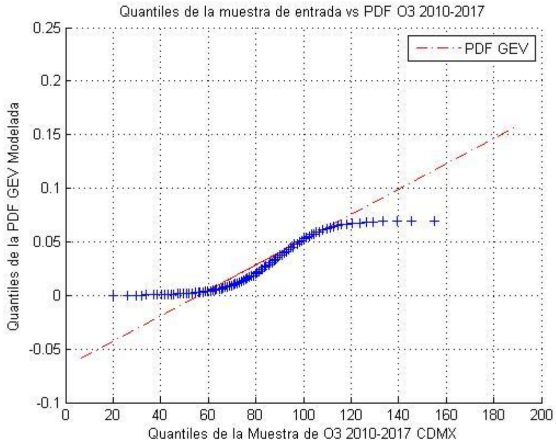


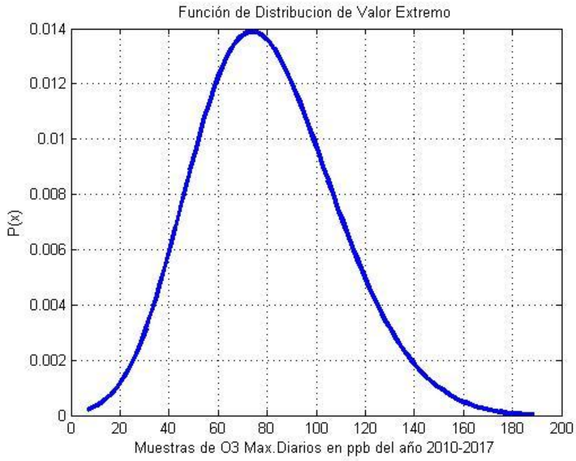
Figure 3. means obtained with the Gaussian Mixtures with Trend 3 Ozone downward (Author)

Table 9. Function variable properties and extreme GEV









GEV **K = -0.1831** **Sigma = 26.9508** **Mu = 68.5262** **Mean = 79.88 = 823.52 Variance**	**MSE = 0.0000943** **RMSE = 0.0083** **AP = 0.9991** **R2 = 0.9896**

Trend of Ozone

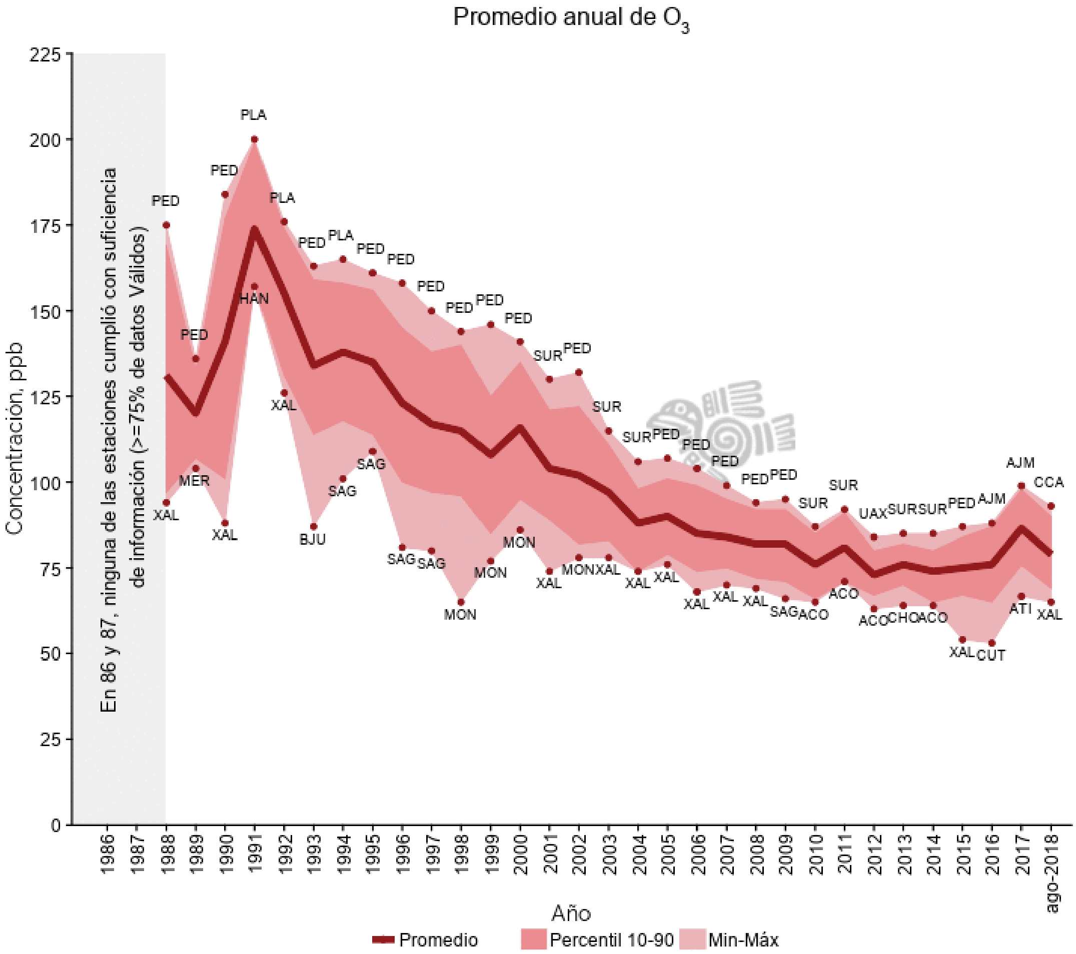


Figure 4. Average Ozone Trends (Source:http://www.aire.cdmx.gob.mx/)

Table 10. Average approximated by the sum of the Standard clearly see the downward trend of ozone concentration

YEARS	Averages Ozone ppb	Gaussian sums ppb Mean
2010	81	87
2011	83	82.50
2012	76	90
2013	78	94.5
2014	76	90
2015	78	87
2016	78	82.50
2017	79	75

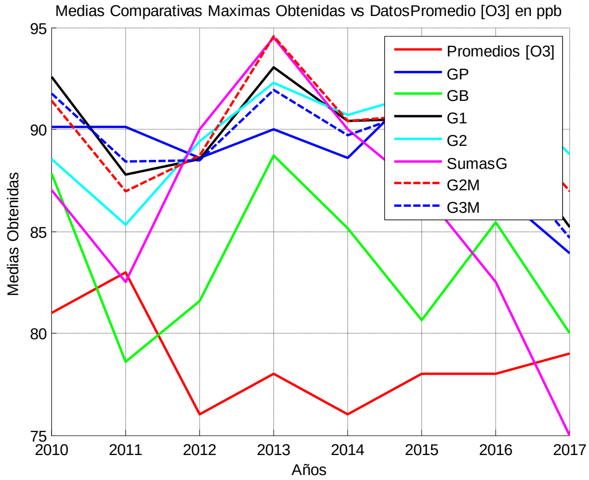


Figure 5. Means obtained from different PDF Gaussian found against the official average ozone trend data

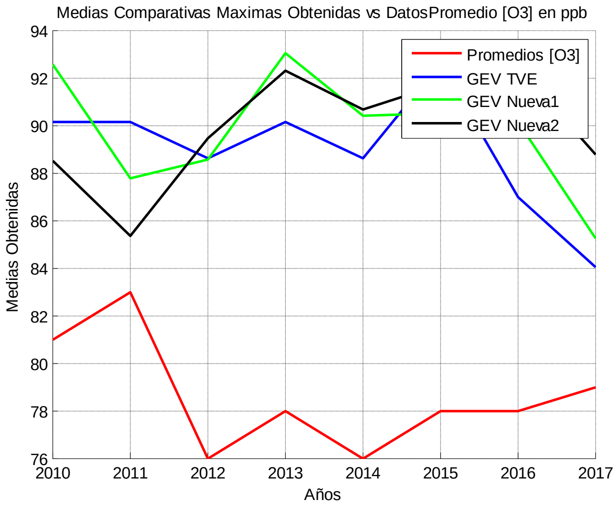
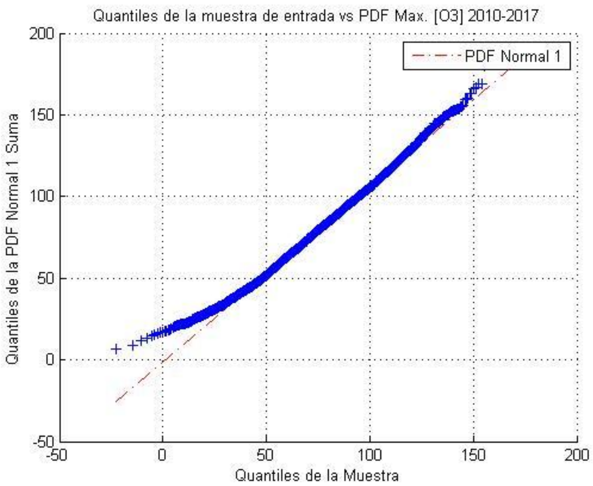
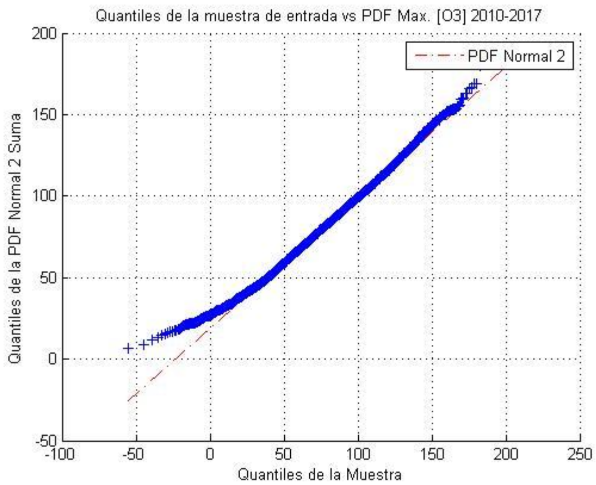
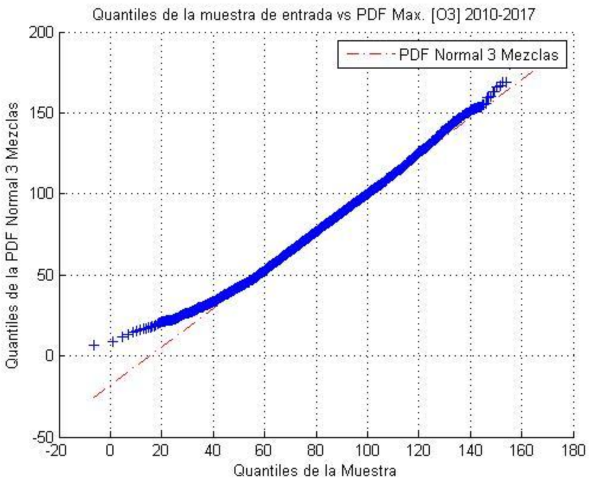


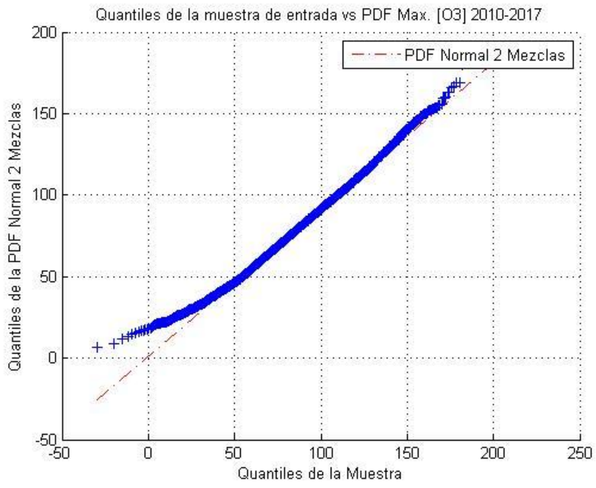
Figure 6. averages obtained from different PDF GEV found against the official average ozone trend data

Table 11. Graphics QQ plot of Normal Normal addition and 3 and 2 Mixtures









Let us now limits according to the theory proposed

Table 12. Measurements Testing Parameters derived GEV

Let us now limits for GeVs	**GEV of 2010-2017 Trend**
GEV $[\mathbf{21ppb} \leq \mathbf{215.7}\boldsymbol{ppb} \leq \mathbf{215.7ppb}]$ GEVA $[\mathbf{21ppb} \leq \mathbf{208}\boldsymbol{ppb} \leq \mathbf{215.7ppb}]$ GEV2 $[\mathbf{21ppb} \leq \mathbf{194.81}\boldsymbol{ppb} \leq \mathbf{215.7ppb}]$ GEV1 $[\mathbf{21ppb} \leq \mathbf{161.9}\boldsymbol{ppb} \leq \mathbf{215.7ppb}]$	**K = -0.1831** **Sigma = 26.9508** **Mu = 68.5262** **With 215.7 ppb**
GEV 1 **k2 = -0.3164** **SigmaSD = 27.7729** **MupostMean = 74.1718** **With 161.9 ppb**	**GEV 2** **k2 = -0.3164** **SigmaSD2 = 37.0936** **MupostMean2 = 77.5822** **With 194.81 ppb**

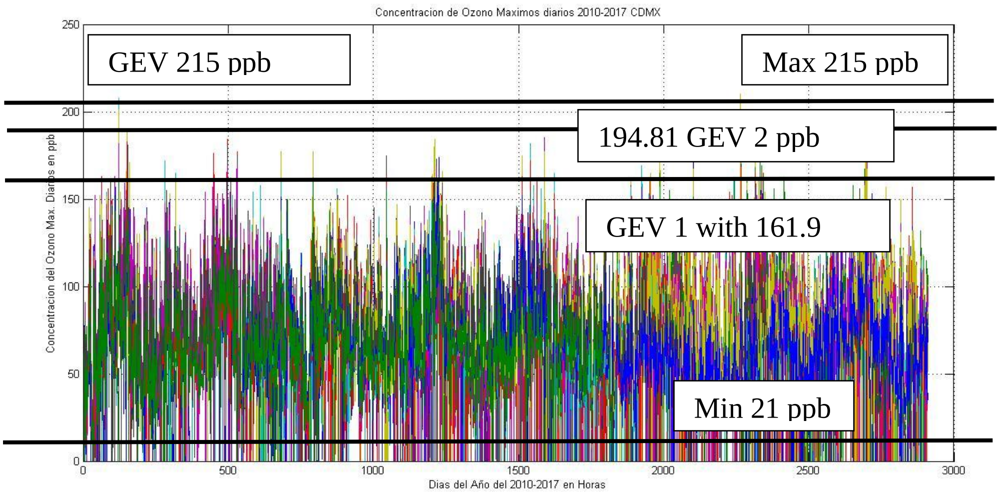


Figure 7. Series Temporary Maximos Ozone 2010-2017

Conclusions

With this methodology we have a choice to find day forecast, we can also see that the equations proposed for the parameters for the new functions extremely variable is satisfactory and comply with the given theory and demonstrated also that when applied to the data actual give their forecasts each and with the functions of normal distribution new grant pretty good day forecast, this method can see that not only a single function fitted to the data distribution gives favorable results even being a perfect fit as well the amounts of these normal operations also give their approach these results were made with every year but here only shows the trend of the years 2010-2017,the trend shown also coincides with the official chart trend monitoring page Mexico City, although stockings have been given higher or lower than shown.

The software used was Matlab 2015, with some functions and subroutines as Stochastic Gaussian mixture.

References

[1]. Berger, JO, Statistical Decision Theory and Bayesian Analysis, Springer Ser. Stat., 2nd ed., Springer-Verlag, New York, 1985.
[2]. Christopeit, N., Estimating parameters of an extreme value distribution by the method of moments, J. Stat. Plann. Inference, 41, 173-186, 1994.
[3]. Davison, AC, and RL Smith, Models for exceedances over high thresholds, JR Stat. Soc., Ser. B, 52, 393-442, 1990.
[4]. Hosking, JRM, L-moments: Analysis and estimation of distributions using linear combinations of order statistics, JR Stat. Soc., Ser. B, 52, 105-124, 1990.
[5]. Landwehr, JM, NC Matalas, and JR Wallis, Probability weighted moments Compared With some traditional techniques in estimating Gumbel parameters and quantiles, Water Resour. Res., 15 (5), 1055-1064, 1979.
[6]. Li, Z. Applications of Gaussian Mixture Model to Weather Observations. IEEE Geoscience and Remote Sensing Letters (Volume: 8, Issue: 6, Nov. 2011)
a. McLachlan, G. & Peel, D. Finite Mixture Models.
[7]. Prescott, P., and AT Walden, Maximum-likelihood estimation of the
a. parameters of the three-parameter generalized extreme-value distribution
b. from censored samples, J. Stat. Comput. Simul., 6, 241-250, 1983.
[8]. Robert, CP, The Bayesian Choice: A Decision-Theoretic Motivation, Springer Ser Stat, Springer-Verlag, New York, 1994
[9]. Ryan Prescott Adams and David J.C. MacKay, Bayesian Online Changepoint Detection, University of Cambridge, https://arxiv.org/abs/0710.3742
[10]. Otten, A., and MAJ Van Montfort, Maximum-likelihood estimation of overall the extreme-value distribution parameters, J. Hydrol., 47, 187-192, 1980.

Online references

[1]. https://es.wikipedia.org/wiki/Desigualdad_de_Chebyshov
[2]. https://en.wikipedia.org/wiki/Location_parameter
[3]. https://en.wikipedia.org/wiki/Scale_parameter
[4]. https://en.wikipedia.org/wiki/Shape_parameter
[5]. https://brilliant.org/wiki/gaussian-mixture-model/
[6]. https://ieeexplore.ieee.org/document/5936097/?reload=true
[7]. http://mccormickml.com/2014/08/04/gaussian-mixture-models-tutorial-and-matlab-code/
[8]. https://es.wikipedia.org/wiki/An%C3%A1lisis_de_grupos
[9]. http://academic.uprm.edu/eacuna/