Selection of a Subset of Meteorological Variables for Ozone Analysis: Case Study of Pedregal Station in Mexico City

DAVID PUBLISHING

D

Sara Rodríguez1, Hortensia Reyes1, Paulino Pérez2 and Humberto Vaquera2

1. Department of Mathematics and Physics, Autonomous University of Puebla, Puebla 72570, Mexico

2. Post-Graduate Academy, Montecillo 56230, Mexico

Received: May 20, 2011 / Accepted: July 4, 2011 / Published: January 20, 2012.

**Abstract:** In this paper, we describe how we carried out a research using the methodology Extreme Value Theory to model a problem of environmental pollution. Through the likelihood ratio test, we selected the best subset of environmental variables. The procedure was applied to data of Pedregal Station, which belongs to the atmospheric monitoring system of Mexico City, managing the maximums of each three days in the period from 1990 to 2009. The importance of these studies on ozone concentrations is that exposures for a long period of time may cause health damage to alive/living beings, in short and long terms, mainly for people living in megacities. The environmental variables considered in this analysis are: wind, temperature and relative humidity. The chemical variables that we took are: carbon monoxide, nitrogen dioxide and sulfur dioxide. Although we did not examine all possible subsets, we took the most important ones by nested subsets, comparing them using the deviance statistic. Finally, we observed that the best subsets were those involving the time variable and, the one that contained all variables except sulfur dioxide, which lead us to think that explaining the ozone concentration evolutions through these variables, is still a complicated case, since more accurate methodologies are discovering new associations that are related with its creation and evolution. Environmental pollution is a very difficult problem that requires to carry out new researches, measurements and interactions, that help to put forth the acquired knowledge into new models.

**Key words:**Generalized extreme value distribution, ozone, selection of variables.

1. Introduction[[1]](#footnote-2)

The Extreme Value Theory is an area of statistics dedicated to develop models and techniques that estimate the behavior of rare events that in some cases have catastrophic results. These belong to the long-tailed of a distribution and are far from central statistics. These events exceed a threshold (maximum or minimum value) in a fixed time period and use the asymptotic distribution of the maximums properly normalized by sequences of random variables [1-3]. Extensions on this theory exist for the case of extreme modeling; one is in a continuous time [4] and the other uses Bayesian methods as it has been done by some researchers [1, 5, 6].

Ozone is the most harmful oxidant for life on earth that is produced by human activities. For this reason, the specialists, in conjunction with public institutions, have been carrying on investigations in areas related to ozone and health, to measure the resistance of individuals living in urban zones, in according with interesting variable, like age orthe exposure time, among others [7, 8]. Several statistical methodologies have been applied, for example multivariate models, non linear regressions, time series, non parametric models, Bayesian and neuronal networks, copula and spatial models [9-15].

Mexico city is a mega city since it has more than ten million people, who suffer of a deficiency of oxygen in their lungs of less than 23%, in comparison with people living in other places at sea level [16]. There are geographical conditions in this city (a place enclosed by mountains) that favor photochemical processes in the atmospheric stability, causing constant thermal inversions. Industry is located in the northern part of the city, and winds blow from north to south, leading to dispersion of pollutants (nitrous oxides, volatile organic compounds, among others), and also are dragged to the south, where complex chemical processes occur by sunlight, which gives raise to ozone creation [17].

Mexico City’s authorities issued environmental controls. One of them is the standard of air quality which can not exceed 0.11 parts per million of ozone, and that a person should not be exposed to this level for more than an hour [18]. The Atmospheric Monitoring System of Mexico City [19] as one of its policieshas to verify the validity of the information and constantly carries on analysis of its data. These information bases are the most used in Mexico due to their reliability and facility to obtain the data, which is available for the any user that in turn helps diverse pollution researchers, belonging to diverse fields on this matter like the statistics area, or others to the environmental physical chemistry and/or photochemical. Some of the works in the Bayesian statistics estimate the probability that a standard in air quality could be exceeded for a number of times within a time interval. Assigning a non informative initial distribution, researchers use a non homogenous Poisson distribution with multiple change-points and note that changes that occur in the data are reflections of environmental actions that have been taken by the inhabitants of Mexico City valley, and by the Mexican government in two different years in 2000 [20]. Another work on this Bayesian approach is the proposal for temporary space model in the interpolation and prediction of hourly ozone concentrations. Temperature was taken as covariate in some SIMAT stations, and was interpolated when data was missing. The model incorporates spatial covariance structure for observations and parameters that define the harmonic components, using a dynamic linear model. This model yields good results for short time prognostics, but it does not incorporate transport of pollutants, also does not respond to physical relations and, neither involves chemical reactions that occur at different atmospheric levels [6]. Subsequently, the authors propose a new approach assuming that the empirical distribution of ozone can be approximated by a Generalized Extreme Values type distribution, whose parameters define a time and space structure. The location parameter has a dynamic linear model and parameters of scale and form are depending on time. In the analysis of Mexico City data, an example is given in a flexible form to capture short and long term tendencies, also depending on time. The research that that was presented is a challenge to try variations in time lengths or the selection of a threshold. It was observed in the analysis that optimal selection of thresholds depended on neighboring stations [21]. Another article that uses the classic Extreme Generalized Values Distribution proposes a simple model to estimate trends in time series, that fits the ozone as the classic linear regressions models do, from the convergence of the quantile function to a standardized normal distribution, which allows to do hypothesis tests and confidence interval in the parameters of interest. The reported results confirm that the series of data presents two change points, one increasing before 1996 and another decreasing after this time [22].

In the field of environmental chemistry and physics, at the Atmosphere Science Center of the UNAM, researchers analyze the non lineal behavior among the precursors and ozone production in the city. They also inquire into the causes of the rising of high levels of ozone on a specific time. Their motivation was based on knowing why on a weekend, that are days off, pollution increased so much. They proposed hypothesis, and simulated real data bases, ozone creation through nitrous oxides, volatile organic compounds and climatic conditions on those days. This work got to demonstrate that strong winds carried pollution from Toluca metropolitan zone to the Valley [23, 24].

2. Material and Methods

Extreme Value Theory is based on Asymptotic Theory, derived from Fisher and Tippet principle [25], which establishes that there are only three types of distributions which are reached on the limit of maximums of stationary sequences. Fisher and Tippet found that distributions are unified in the Generalized Extreme Value Distribution, which has three parameters of location, scale and shaped, respectively.

(1)

Eq. (1) is generalized extreme value distribution with

Reiss and Thomaset al. [1, 3] established that the values of the shape parameter determine the distribution type. If then a Fréchet distribution type is obtained, if then a Weibull distribution type is obtained, in the case a Gumbel distribution is obtained. The results given in Ref. [22] allow us to use a linear regression model in order to analyze ozone concentrations through environmental and chemical variables. To accomplish this, Cox’s idea [26] was used, who employed the regression model in a Generalized Pareto. In our case, we took a covariates vector, the unknown parameter vector was expressed into . In this last expression, we introduced the maximums of the standardized covariates to avoid scale problems in the fixed time period. The likelihood function is given by:

(2)

The maximum likelihood estimators of the parameters are obtained by maximizing the likelihood given by Eq. (2), using numerical methods.

3. Procedure for Selecting the Best Model

Taking the information about observations and *k* covariates, hypothesis tests are done on the covariates coefficients to know if any of them influences the trend of ozone maximums concentrations. This was achieved by observing the estimator sign of the covariate coefficient, namely we need to contrast the hypothesis

Eq. (3) is test hypothesis for covariates.

where represents the parameter associated with time, it’s associated to the trend,..,k), if there will be an decreasing trend, if there will be an increasing trend and in the last case, there will be no tendency.

The test consists in rejecting if where is the significance level and is the *i-th* term of the diagonal matrix , for this, variances are obtained by using the observed Fisher information matrix.

In the event that the null hypothesis is rejected with significance level α, it is evident that it exists a trend in the observations of Y. If all covariates estimators were zero, this would indicate that these are not influencing to explain the trend of Y.

Using Coles procedure of selecting nested models [1] that supposes a model with a parametric vector and a subset vector that is obtained by restricting components to a value, let us say they take zero value. This implies to partition vector where the sub-vector -dimensional fulfills that . The test involves to obtain the estimators ( and ( to prove the validity of the model using the relation of with to a significance level . From where, the hypothesis test is expressed as follows:

 (4)

Eq. (4) is test hypothesis for better model.

The decision rule is to reject the null hypothesis if when is the quantile () of the distribution , is the difference of model dimensions and . Large values of *D* indicate that model explains better data variations than model , but if *D* is small, it is not worth to incorporate more variables because the model has no improvements. The estimations with maximum likelihood using nested models lead to carry out pairwise testing [27, 28]. Akaike’s information criterion (AIC) [29]is given by Eq. (5):

 (5)

where is the number of parameters estimated in the model *M*, and is the value of the log-likelihood function for the set with estimated parameters. This criterion helps us to solve the problem of choosing between rival models that are non-nested in terms of their functional forms. For a given data set, several models can be fitted; according to the *AIC* criterion the model with smallest *AIC* is the best (more negative *AIC* value).

4. Applications

Mexico city was considered as the city with the hardest environmental problems in the worlda few years ago, since there were motor vehicles that circulated daily with no restrictions at all and there were not regulations on industries located in the city, whereby there was frequently maximum levels at 0.11 ppm with the most harmful variable, the ozone [30]. Currently, some situations have been controlled with the collaboration of everybody and nowadays, so it is no longer the most polluted in the world [31, 32].

The Sistema de MonitoreoAtmosférico (SIMAT) [33] of Mexico City began to operate in 1986. It is integrated by the Red Automática de Monitoreo Atmosférico (RAMA), the Red manual de Monitoreo Atmosférico (REDMA), the Red de Depósito Atmosférico (REDDA), and the Red de Parámetros Meteorológicos (REDMET), located in different places on the valley. Their geographic locations are: northwest (NW) which has 17 stations, southwest (SW) with 6 stations, northeast (NE) with 4 stations, in the southeast (SE) there are 5 stations and in the center (C) there are 3 stations. We chose Pedregal Station which is one of the SW stations, since it presents the highest levels of ozone in all the valley, its location is indicated in Fig. 1.

In all valleies, thermal inversions constantly occur, the station to be analyzed belongs to Alvaro Obregón delegation, with a population of 729,193 and 198,647 housing [34]. The reduction on very extreme pollution levels has been achieved through time [35], with efforts form the inhabitants as well as restrictions on daily movement of motor vehicles, accessible costs in public transportation and relocation of some polluting industries, like the refinery of Azcapozalco [36]. Unfortunately, the city grows by giant leaps that causes accumulation of atmospheric pollutants, this aggravates with more than 18 million people that inhabit in Mexico City and vehicle parking [37, 38]. There are frequent thermal inversions in more than 70 percent of the year, due to a natural phenomenon of air stagnation, this gets dispersed in a gradual way in the course of the day and breaks due to atmospheric warming (Table 1).

****

**N**

**W**

**S**

**E**

**Fig. 1 Geographic location of Pedregal station (PED) which is placed in the southwest of Mexico City that belongs to the** **Alvaro Obregón delegation [33].**

**Table 1 Summarized descriptive information about Pedregal Station (PED) of 1990-2009, n = 2073 [33]**.

|  |  |
| --- | --- |
| Summary of ozone | Value |
| Minimum | 0.029 |
| 1st quantile | 0.129 |
| Median | 0.169 |
| Mean | 0.174 |
| 3rd quantile | 0.215 |
| 4th quantile | 0.404 |
| Maximum | 2.073 |

We present a study to analyze the influence of 8 covariates (environmental and chemical) in ozone concentrations in Pedregal Station that belongs to the metropolitan zone of Mexico City. It shown the measurement unit and the variables code that wereused (Table 2).

Using a linear regression model in the scale parameter, for each group of 72 hour, we took the maximums of the 7 covariates.

Then, we present graphics of the 7 covariates measurements that have different measuring units and that were standardized to avoid scale problems (Fig. 2).

The maximum likelihood estimators for high and fixed *p* were obtained using a program written in Ref. [39]. Some nested models were estimated for various parameters in

 (6)

Eq. (6) is generalized extreme value distribution with covariates.

Based on this information hypothesis testing can be done about ozone concentrations trend, in order to know the way some covariates get involved in the model.

Fig. 3 shows probability plot, quantile plot, density plot for the ozone concentrations from 1990 to 2009. Theses graphics can be used to check the validity of distributional assumptions.

Table 3 shows the maximum likelihood estimator and standard deviation (SV) for the parameter associated to time. In this table, we show that the estimator of time parameter is negative. That is to say, ozone decreases over time.

**Table 2 Names of** **covariates with its units of measurement and their respective codes, in Pedregal Station.**

|  |  |  |
| --- | --- | --- |
| Variable code | Measurementunit | Renamed |
| Ozone (O3) | ppm |  |
| Time (t) | hours | X1 |
| Temperature (tmp) | DegreeCelcius (°C) | X2 |
| Windspeed (wsp) | m/s | X4 |
| Relative humidity (hr) | % | X3 |
| Carbonmonoxide (CO) | ppm | X5 |
| Nitrogendioxide (NO2) | ppm | X6 |
| Sulfurdioxide (SO2) | ppm | X7 |

****

****



**2.0**

**1.5**

**1.0**

**0.5**

**-0.5**

**0.0**

**Fig. 2** **Graphics of the 7 covariates that were related to the ozone variable at Pedregal station during the period from 1990 to 2009. These are** **O3 (dependent variable), tmp, hr, wsp, CO, SO2 and NO2 (independent variables). Due to the limit of space in the text, we only present some graphics.**

Another example with other model is as shown in Table 4, where we have a negative sign in the coefficients of time, wind speed and nitrous dioxide.

The remaining variables have opposite signs, indicating positive contribution for ozone creation.

Selection of the Best Subset of Variables: The number of models to be selected can be big if the number of variables is high, in our case, there are 26 = 64 subsets to analyze. The models that will be analyzed are those ones nested in model, and that the biggest model contains the 6 covariates, whereby we present 6 models of interest (Table 5).

Table 6 shows the log-likelihood, deviance (with model as reference) and degrees of freedom for the tested models. The last column of this table shows the result of testing hypothesisthat a given model fits the data better than for different significance levels, and .

Whereby, it can be observed that the best models





**Fig. 3 Decriptive analysis of distributional assumptions of GEV at Pedregal Station.**

**Table 3 Model that uses ozone and time where MLE and SV are the maximum likelihood estimators and their standard deviation, repectively.**

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter |  | MLE | Standard Deviation (SV) |
|  |  | -2.24470 | 1.699 × 10-2 |
|  |  | -0.00057 | 1.99 × 10-6 |

**Table 4 Model where MLE are maximum likelihood estimators, in brackets is the standard deviation with .**

|  |  |  |
| --- | --- | --- |
| Parameter |  | MLE (Standard Deviation) |
|  |  | -1.9470 (1.83 × 10-2) |
|  |  | -0.0008 (1.99 × 10-6) |
|  |  | -0.1006 (6.59 × 10-2) |
|  |  | 0.0045 (7.86 × 10-4) |
|  |  | -0.0514 (1.97 × 10-2) |
|  |  | 2.6421 (9.36 × 10-1) |
|  |  | -1.2019 (1.26 × 10-1) |
|  |  | 1.4989 (3.31 × 10-1) |

**Table 5 Nested models of interest with the covariates.**

|  |  |  |
| --- | --- | --- |
| Models |  | Covariates |
|  |  | *t* |
|  |  | *t and tmp* |
|  |  | *t, tmp* and *NO2* |
|  |  | *t, tmp, NO2 and wsp* |
|  |  | *t, tmp, NO2, wsp and SO2* |
|  |  | *t, tmp, NO2, wsp, SO2* and *CO* |

and are significant to ; in the first case, the model is in terms of time and in the second one all variables are present, except sulfur dioxide. From Table 7 it can be concluded the models that best fits the data are and , which agrees with the results obtained using the deviance criterion.

Fig. 4 shows the graphics of residuals of the model (residual probability plot). The residuals present a good behavior in general, since the residuals of the model follow a straight line tendency, but not all graphics are good, for example the residual quantile plot, we can see that the last real values were highly deviated with respect from the straight line. The central data of model shows that is a little deviated from the estimated line (left figure), and in Model the recent data is outside of the estimated line (right figure).

**Table 6 Deviance function results (D), log-likelihood, degrees of freedom (*k*), for Pedregal Station.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Models | log-lik |   | k |   |   |
|   | 2757.335 |   |   |   |   |
|   | 3055.176 | -595.682 | 7 | < |  is better than  |
|   | 2095.814 | 1323.042 | 6 | > | Rejected that is better than  |
|   | 2351.725 | 811.22 | 5 | > | Rejected that is better than  |
|   | 2463.099 | 588.472 | 4 | > | Rejected that is better than  |
|   | 2806.472 | -98.274 | 3 | < |  is better than  |
|   | 2167.648 | 1775.056 | 1 | > | Rejected that is better than  |

**Table 7 Results Akaike’s criterion, for Pedregal Station**.

|  |  |  |  |
| --- | --- | --- | --- |
| Models | log-lik | k |  |
|   | 2757.335 | 11 | -5492.67 |
|   | 3055.176 | 4 | -6102.352 |
|   | 2095.814 | 5 | -4181.628 |
|   | 2351.725 | 6 | -4691.45 |
|   | 2463.099 | 7 | -4912.198 |
|   | 2806.472 | 8 | -5596.944 |
|   | 2167.648 | 10 | -4315.296 |



**Fig. 4 Residuals probability plot for and models respectively**.

5. Conclusions

The station with the higher records of pollution is Pedregal. Its geographical location is ideal for air flows to carry pollutants to the southeast zone and since air circulation is poor, due to that the valley is surrounded by mountains, this produces higher levels of ozone concentration, much above those health standards allowed by Mexico City authorities. The proposed methodology analyses the behavior of covariates included in the model. This supposes observations on Generalized Extreme Value distribution. It is important to observe the value of signs in the coefficients related with the increasing or decreasing tendency of the model. In the case of comparison of the proposed nested sub-models using chemical and environmental variables, it turns out that the models that explain better ozone concentration are two: one of them contains time variable and the other one contains all variables, except sulfur dioxide.

References

1. Coles, A., ed. 2004. *An Introduction to Statistical Modeling of Extreme Values*. Springer Verlag.
2. Galambos, J., ed. 1978. *The Asymptotic Theory of Extreme Order Statistics*. Roberte: Temple University Philadelphia Pennsylvania, Krieger Publishing Company.
3. Reiss, R., and Thomas, M., eds. 2001. *Statistical Analysis of Extreme Values*. Germany: BirkhauserVerlag. 2001.
4. Albeverio, S., Jentsch, V., and Kantz, H. 2005. *Extreme Events in Nature and Society*. Springer.
5. Coles, A., and Powell, E. 1996. “Bayesian Methods in Extreme Value Modeling: A Review and New Developments.” *International Statistical Review* 64: 119.
6. Huerta, G., Sansó, B., and Stroud, J.R. 2004. “A Spatiotemporal Model for Mexico City Ozone Levels.” *Royal Statistical Society* 53: 231-248.
7. Hernández, C., Téllez, R., Sanín, A., and Lacasaña, N. 2000, “Relationship between Emergency Room Visits Forrespiratory Illness and Air Pollution in Ciudad Juarez, Chihuahua.” *Public Health in Mexico* 42: 288-297. (in Spanish)
8. PRB (Population Reference Bureau). 2002. Nexos. <http://www.prenatal.tv/lecturas/efectos%20del%20medio%20ambiente%20en%20la%20salud%20infantil.pdf>. (in Spanish)
9. Guyón, X. 2010. “Modelación Para la Estadística Espacial.” *Revista de Investigación Operacional* 31 (1): 1-33. (in Spanish)
10. Davison, A. 2009. Spatial Extremes and Applications: Discussion Outcomes. <http://extremes.epfl.ch/files/content/sites/extremes/files/users/111184/public/WorkshopJuly09/ExtremesDiscussionsOutcomes.pdf>.
11. Porter, S., Rao, T., Zurbenko, G., Dunker, M., and Wolff, T.G. 2001. “Ozone Air Quality Over North America: Part II­An Analysis of Trend Detection and Attribution Techniques.” *Air and Waste Management Association* 51: 283-306.
12. Quintela del Río, A., and Fernández, M.F. 2010. “Extreme Values Analysis of Ozone Concentrations Using Nonparametric Methods.” The Science of the Total Environment 409: 1123-1133.
13. Sang, H., and Gelfand, A. 2009. “Continuous Spatial Process Models for Extreme Values.” [*Journal of Agricultural, Biological and Environmental Statistics*](http://www.google.com.mx/url?sa=t&rct=j&q=jabes&source=web&cd=21&ved=0CCUQFjAAOBQ&url=http%3A%2F%2Fwww.amstat.org%2Fpublications%2Fjabes.cfm&ei=7pCnTqSPEYeDgAfZgYku&usg=AFQjCNGOTI_tskmnpCn03s1BkBLST-lV6w) 15: 49-65.
14. Thompson, M.L., Joel, R., Lawrencw, H.C., Peter, G., and Paul, D.S. 2001. “A Review of Statistical Methods for the Meteorological Adjustment of Tropospheric Ozone.” *Atmospheric Environment* 35: 617-630.
15. Wolff, T., Dunker, M., Rao, T., Porter, S., and Zurbenko, G.I. 2001. “Ozone Air Quality over North America: Part I-A Review of Reported Trends.” *Air and Waste Management Association* 51: 273-282.
16. Molina, J., Molina, L. 2004. “The Impacts of Magacities on Air Pollution.” Presented at the Environmetal Aspects of Urbanization Conference, Goteborg, Sweden.
17. Young, A., and Betterton, E., Saldivar de Rueda, L. 1997, “Atmósfera, México.” *UNAM* 10: 161-178.
18. NOM. 2002. “Modification Mexican Official Standard NOM-020-SSA1-1993.” *Official Journal of the Federation-Mexico*.
19. SIMAT (Sistema de Monitoreo Atmosférico de la Ciudad de México). 2010. “Agricultura No. 21, Primer Piso, Col. Escandón, Del. Miguel Hidalgo, C.P. 11800.” http: //www:sma.df.gob.mx./simat2/.
20. Achcar, J., Rodríguez, E., and Tzintzun, G. 2009. “Using Non-homogenous Poisson Models with Multiple Change-points to Estimate the Number of Ozone Exceedances in Mexico City.” *Environmetrics* 22 (1): 1-12.
21. Huerta, G. and Sansó, B. 2005. “Time-varying Models for Extreme Values.” *Environmental and Ecological Statistics* 14: 285-299.
22. Reyes, H., Vaquera, H., and Vilaseñor, J. 2010. “Estimation of Trends in High Urban Ozone Levels using the Quantiles of (GEV).” *Environmetrics* 21: 470-481.
23. García, A., Jazcilevich, A., Ruíz, L., Torres, R., Suárez, L., and Reséndiz, N. 2009. “Ozone Weekend Effect Analysis in México city.” *Atmósfera* 22 (3): 281-297.
24. Muñoz-Cruz, R., Granados-Gutiérrez, M., and del C.Jaimes Palomera, M. 2007. “Análisis Del Comportamiento Semanal del Ozono en la Zona Metropolitana Del Valle de México en el periodo 1990-2006.” <http://www.sma.df.gob.mx/simat/informes.tec.htm>.
25. Pickands, J. 1975. “Statistical Inference using Extreme Order Statistics.” *The Annals of Statistics* 3 (1): 119-131.
26. Cox, W., and Chu, S. 1996, Assessment of Interannual Ozone Variation in Urban Areas from a Climatological Perspective. *Atmospheric Environment* 30: 2615-2625.
27. Casella, G., and Berger, R., eds. 2002. *Statistical Inference*. Thomson Learning. Duxbury.
28. Hogg, V., McKean, W., and Craig, T.A., eds. 2007. *Introduction to Mathematical Statistics*. Prentice Hall. Jersey.
29. Akaike, H. 1974. “A New Look at the Statistical Model Identification.” *IEEE Transactions on Automatic Control* 19 (6): 716-723.
30. Mexican Official Regulation, NOM-020-SSA1-1993 [Online], http://www.sma.df.gob.mx/sma/links/download/archivos/nom-020-ssa1-1993.pdf.
31. Blacksmith Instituto, 2010. Los lugares más contaminados del mundo. Blacksmith Instituto Web site. <http://blog.nuestroclima.com/?p=183/2010/12/31/los-lugares-mascontaminados-del-mundo/>. (in Spanish)
32. Bravo, H. 1992. “La Contaminación Atmosférica por Ozono en la Zona Metropolitana de la Ciudad de México: Evolución Histórica y Perspectivas.” Presented at IX Comisión Nacional de los Derechos Humanos. (in Spanish)
33. Norma OficialMexicana NOM-020-SSA1-1993. 2004. SIMAT. [http://www.sma.df.gob.mx/simat/proteccion/(ante)nom-020-ssa1-1993.pdf](http://www.sma.df.gob.mx/simat/proteccion/%28ante%29nom-020-ssa1-1993.pdf). (in Spanish)
34. INEGI (National Institute of Statistics and Geography) Website. http://www.inegi.org.mx/default.aspx?.
35. Rodríguez, S., Reyes, H., Vaquera, H., and Linares, G. 2009. “Modelación Estadística en Contaminación Del Aire Para dos Estaciones Vecinas del Estado de México, Aportaciones y Aplicaciones de la Probabilidad y la Estadística.” *Benemérita Universidad Autónoma de Puebla Dirección General de Fomento Editorial* 3: 43-53. (in Spanish)
36. Montiel, R.A., López, O.A.M., Lichtinger, V., Frenk, M.J. “Program to Improve Air Quality in the Metropolitan Area Mexico 2002-2010.” World Resources Institute. http://projects.wri.org/sd-pams-database/mexico/program-improve-air-quality-mexico-city-metropolitan-area-proaire
37. Jazcilevich, A.D., Agustin, R.G., and Gerardo, R.S. 2003. A Study of Air Flow Patterns Affecting Pollutant Concentrations in the Central Region of Mexico.” *Atmospheric Environment* 37: 183-193.
38. Censo. 2010. “Instituto Nacional de Estadística y Geografia Website.” [www.inegi.org.mx/Sistemas/temasV2/Default.aspx?s=est&c=17484](http://www.inegi.org.mx/Sistemas/temasV2/Default.aspx?s=est&c=17484).
39. Stephenson y Gillel, Ismev Package: Extreme Values in R. 2004. R Foundation for Statistical Computing. http://www.r-project.org.
1. **Corresponding author:** Sara Rodríguez, master, main research field: statistics. E-mail: rguez.sara@gmail.com. [↑](#footnote-ref-2)