International Journal of Latest Research in Science and Technology **ISSN (Online):2278-5299**
Volume 10, Issue 1: Page No.1-5 (2021)
https://www.mnkjournals.com/journal/ijlrst/index.php

MARKOV CHAINS MODEL APPLIED TO THE ANALYSIS OF AIR QUALITY

Brenda García Jarquín*, Mario Aguilar Fernández,
Imelda Calderón Vidales, Jezzica Jiovanna Zamudio

[1]UPIICSA, SEPI, Instituto Politécnico Nacional, jarquin_garcia@yahoo.com, México
[2]UPIICSA, SEPI, Instituto Politécnico Nacional, maguilarf@yahoo.com, México
[3]UPIICSA, SEPI, Instituto Politécnico Nacional, melaymelo@gmail.com, México
[4]UPIICSA, SEPI, Instituto Politécnico Nacional, jezz.zamu@gmail.com, México

Received: 28 December 2020; Accepted :10 January 2021; **Published** :20 January 2021

Abstract: ***Urban areas are characterized by high rates of population, high density of motor vehicles, industries, services, and commercial activities whose emissions alter air quality. In turn, the urban complexity presents challenges for the environmental problems so effectively manage environmental risks depends on the information available. Predict and prevent contaminant levels becomes necessary to reduce the risk to health and improve the environment through control strategies and emission reduction. The objective was to apply the analysis of Markov chains to analyze the behavior of pollutants in Mexico City. The study considered five criteria pollutants: CO2, NOx, SO2, O3, PM10. The values used to define the criteria to the analysis corresponded to the values defined by the metropolitan index of air quality in environmental law in Mexico from which five possible states from good to extremely bad are defined. The results obtained shows that: State changes have nonlinear behavior over time and even though the contaminant concentration increases exponentially, the probability of air quality worses is very low because it is kept in a "regular" state.***

Keywords- Air quality, Mexico City, Markov chains

I. INTRODUCTION

Urban complexity presents great challenges for the attention to environmental problems arising from qualities that are inherent in every city. Making decisions based on what's relevant and complete information possible by the authority, together with the participation of citizens in the definition of policies for environmental management should be part of normal urban life. However, the lack of access to technical information, the urgency in making decisions and the difficulty for integrate citizen participation in the construction of policies for environmental and social management make it even more difficult to effectively manage environmental risks [1].

A healthy environment is a prerequisite for the welfare and health. However, metropolitan areas such as Puebla, Puebla, Mexico, is characterized by high population indices and therefore have a high density of vehicles, industries, services and commercial activities, whose emissions alter air quality. Suspended particles are among the priority pollutants emitted into the air, which are being characterized by two conditions: the association of compounds of diverse nature and size, which depends on its environmental dynamics, i.e., retention and dispersion in the environment and their impact on health when inhaled [2],[3],[4]. The chemicals associated with particles can significantly affect the health of residents and alter the ecosystem and damaging architectural structures, leading to the need to implement strategies to control pollution levels and chemical exposure in order to reduce impacts on health and well-being of urban populations [5],[6].

Achieving these goals is difficult because of the complexity that characterizes the cities; however, it is necessary to introduce measures for managing environmental risks. One way to accomplish this is to assess the dynamics of pollutants in ambient air to establish preventive and contingency measures and thus reduce the impact of these in the urban area, so in this study, an exercise in multidisciplinary integration is done and then an intersectoral discussion approach to generate a program for managing exposure to pollutants in the city of Puebla, Pue., Mexico.

The experience generated by the working group in conjunction with the Government of the State of Puebla, for defining action plans for the prevention management of health effects and the preparation of an environmental plan about contingencies on environmental pollution is presented in the following sections. In the process, first it was developed a predictive analysis using Markov chains to evaluate the possibility of changes in air quality by increasing the concentration of particles of 10 microns and 2.5 microns in diameter (PM10 and PM2.5, respectively) at Ninfas monitoring station, and how the results of this analysis were used as input for discussion among representatives from sectoral institutions.

II. METHODS AND RESULTS

A. Description of Puebla, Pue. México

The metropolitan area of the Puebla City is one of the most important for their economic impact at the state and national level, due to the large number of population living in it. According to INEGI [7], has 3,192,434 inhabitants what makes to that region the fifth most populous after Mexico City, Veracruz and Jalisco; this region accounts for 71.47% of the state GDP, whose contribution to the national level is 3.4%, which puts Puebla in ninth place of the country.

Particularly the metropolitan area of Puebla record gross income in thousands of pesos to $ 3,819,627 due to 55% of the state's population is based in this city. Clearly, this is a region where exist complex relationships between production and reproduction activities in the context of social metabolism, from the use of space and the natural resources of the area, to the excretion of wastes and emissions. This situation was confirmed by two empirical factors, georeferential study of traffic density and the density of the urban plan. These two aspects showed that the downtown area is, indeed, the busiest in the city (Fig 1 and 2). Therefore, the analysis in this paper was made with data from the Ninfas monitoring station, located in the downtown area of the City.

Fig. 1 Geographical location of the four stations monitoring air
quality in the Valley of Puebla

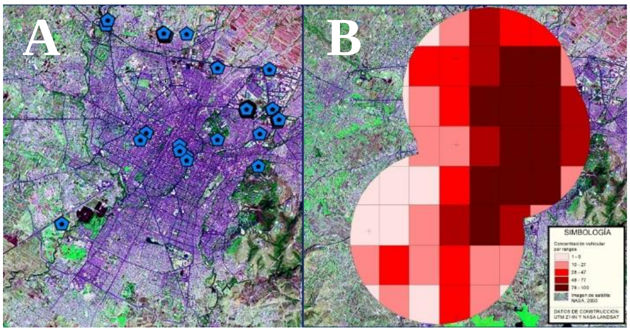

Fig. 2 Metropolitan Zone of Puebla: A. urban density maps, and location of point sources of emission. B. Layer georeferenced on the vehicular density [8].

B. Markov Chain analysis

The imprecision of climate forecasts, and natural phenomena such as earthquakes, storms, etc., is mainly due to the nature that describes an inherently deterministic behaviour, I.e., the occurrence of events are not determined by a historical memory, as If so, a simple statistical method could very accurately predict subsequent events. Because of the random nature of these events, some applications have been developed to model phenomena such as the recurrence of rains in drought-stressed regions [9], development models of epidemics such as process Galton-Watson [10], the pagerank of a search engine on the Internet [11], in genetics has been used to describe the change of gene frequencies as the diffusion model of Moto Kimura [12] as well as in other fields like music with various musical composition algorithms such as Csound or Max software or Boltzmann machines of neural networks [13].

As a mathematical model attempts to represent real-life phenomena, especially probabilistic models can be applied to achieve a better representation of reality. Because these consider a set of random variables (xt) instead of just one, the reliability of results increases. When these random variables are indexed using parameter t (time) they are known as a stochastic or random process. Then, the Markov chain model refers to the stochastic process for which the probability of an event occurring depends only on the immediately preceding event. This feature of memory is called Markov property and applies to events where the nature of occurrence is random and does not necessarily depend on a history of specific events.

Stochastic processes are able to specify probabilities for different possible values at any instant of time. This is possible because the subsequent states in the system are determined by the immediately preceding event as random element. In the natural sciences, a stochastic model will find emergent properties as a result of a phenomenon, then it is possible to identify new behaviours from joint interactions of system elements.

Mathematically a stochastic process where $X = \{X_n : n \geq 0\}$ is a Markov chain if it satisfies the following condition called Markov condition:

$$P(X_n = x_n | X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = X_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1})$$

$\forall n \geq 1$ y $\forall x0, x1, \dots, xn-1, xn \in S$

Intuitively, this equation is interpreted as: that given the "present" process, the "future" is independent of the "past".

Then, a Markov chain is defined as a stochastic process if the current state X_n and the previous state $X_1, \dots, X_{n-1}$ are known, then the probability of the future state is X_{n+1}.

That is, for n = 1,2, ... and for any succession of states $S_1, \dots, S_{n+1}$, it must be

$$P(X_{n+1} = S_{n+1} | X_1 = S_1, X_2 = S_2, \dots, X_n = S_n) == P(X_{n+1} = S_{n+1} | X_n = S_n)$$

Then, given a Markov chain X, its transition matrix is defined, P, as the matrix of transition probabilities of size |S|×|S|, that is:

$$P = (pij)\ i, j \in S \text{ where } pij = P(X1 = j | X0 = i)$$

$$
P = (p_{ij}) = \begin{bmatrix} Q & p & 0 & 0 & \dots & 0 & \dots \\ 0 & q & p & 0 & \dots & 0 & \dots \\ 0 & 0 & q & p & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & Q & p & \dots \\ 0 & 0 & 0 & 0 & 0 & q & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}
$$

To consider that the stochastic transition matrix corresponds to a Markov chain P, these must be met:

$$\text{pij} = \mathrm{P}(X_n + 1 = \mathrm{j} | X_n = \mathrm{i}) \geq 0; \forall \mathrm{i}, \mathrm{j} \in \mathrm{S}$$

Each row sum of P=1

$$\sum_{j \in S} pij = \sum_{j \in S} \mathrm{P}(\mathrm{X}_1 = \mathrm{j} | \mathrm{X}_0 = \mathrm{i}) = 1$$

So, the transition matrix of the n_{th} step, Pn = (pij (n)) is the matrix of transition probabilities in the nth step from the beginning, given by:

$$pij(n) = P(Xn = j | X0 = i)$$

That is, pij is the probability that based on i get j in n steps. Then the evolution of a Markov chain is described by its "transition probabilities", $pn\ ij = P(Xn + 1 = j | Xn = i)$ which initially depend on n, i and j.

This briefly explains the process of formation of the chains because events are random, so this study will be evaluated by a stochastic process to assess the likelihood of change in the concentration of two priority pollutants (PM10 and PM2.5) in the air of the Valley of Puebla, Puebla, Mexico. In this case the time is considered as a discrete parameter, which, despite being able to be infinite (because a contaminant may constantly prevail) is discretized for better handling. In that way the state space (x_t = i) of stochastic process comprises the set of possible values of parameters monitored at stations considered in this study.

For the application of Markov Chain, five possible states were defined taken in account the ranges of air quality for PM10 and PM2.5 established in the NADP-009-AIR-2006 standard (Table 1).

TABLE I

PAPERSCLASSIFICATION OF STATES FOR POLLUTANTS PM10 AND PM 2.5, FROM THE RANGES OF AIR QUALITY UNDER THE NADP-009-AIR-2006

State	State definition	PM10 Range	PM2.5 Range
1	Good	0 - 60	0 - 15.4
2	Acceptable	61 - 120	15.5 - 40.4
3	Tolerable	121 - 220	40.5 - 65.4
4	Bad	221 - 320	65.5 - 150.4
5	Extremely bad	>320	>150.4

Given the conditions of Table 1, the transition matrices were formed for each pollutant. From these, the results showed that for PM10 nine iterations (P^9) were required, unlike for PM2.5 were required 10 iterations (P^{10}). The transition matrices for each pollutant P are shown in Fig 3 and 4, respectively.

	1	2	3	4	5
1	0,988212181	0,011787819	0	0	0
2	0,631578947	0,368421053	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0

Fig. 3 Transition matrix P of PM10 daily means of Ninfas monitoring environmental air station in Puebla, Puebla, Mexico

	1	2	3	4	5
1	0,686567164	0,310447761	0,002985075	0	0
2	0,157232704	0,781446541	0,056603774	0,004716981	0
3	0,089285714	0,625	0,285714286	0	0
4	0	0,333333333	0,666666667	0	0
5	0	0	0	0	0

Fig. 4 Transition matrix P of PM2.5 daily means of Ninfas monitoring environmental air station in Puebla, Puebla, Mexico

According to the results obtained after iterations, it was observed that for PM10, the phase states resulting from matrix stabilized were 1 and 2, corresponding to qualities of air between Good and Acceptable (Table 2).

TABLE 2

STABILIZED MATRIX FOR PM10 DAILY MEANS MONITORED AT NINFAS, PUEBLA, PUE. MÉXICO.

States/ Probabilities	1	2
1	98%	2%

Table 2 shows that the concentration of PM10 particles in the air in the downtown area of Puebla, has a 98% probability of remaining in state 1, implying that tend to remain in the state of air quality "good" and likely to change the "good" state to "acceptable" state is 2%, which means that the data dispersion of PM10 due to a distribution that does not increase linearly in relation to time, although fluctuations in the concentration of the contaminant over time would tend to increase, it is unlikely, in the emission and dispersion conditions currently observed in the region, the air quality may deteriorate into a state of "acceptable" or "regular".

TABLE 3

STABILIZED MATRIX FOR PM2.5 DAILY MEANS MONITORED IN NINFAS, PUEBLA, PUE. MÉXICO

States/Probabilities	1	2	3
1	33%	62%	5%

C. Intersectoral Colaboration

Intersectoral collaboration was given through the coordination of the area of environmental quality of the Secretariat of Environment of the State of Puebla, who joined in the specialized committees covering different areas of environmental management. Given the need to involve different actors in society and promote responsibility in the exercise of environmental management, the participation of institutions under its social aim and ability to contribute information and experience on the issues raised was considered.

For the committee on the management of health risks, two university educational institutions were integrated, the Autonomous University of Puebla, through representatives of the area of medicine and the Autonomous University of Tlaxcala, through the working group of Chemistry and Environmental Mutagenesis, group actually bring newly generated information of research projects supported by the National Science and Technology Council, which had helped members of various university bodies and even other universities in the state of Puebla.

Along with them, participated in this health committee, representatives of the following state and federal agencies: Ministry of Education, Ministry of Health, Secretariat of Communications and Transportation. In addition, also participated an environmental communication consultant of the Federal Secretariat of Environment and Natural Resources.

The objectives set for the work of this commission was established from the assumption of the existence of imminent risk to health from the presence of environmental contaminants as well, we sought to update environmental legislation, have strategies for managing risks health and a contingency plan. These objectives would be able to generate a proposal to amend state law, for the integration of environmental communication actions and management of health risks. However, based on monitoring data and that presented on the dynamics of atmospheric particles and the incidence of respiratory diseases, the urgency to also ensure the management of health risks for other non-anthropogenic pollution factors was established. Instead of a contingency plan, the workgroup decided to extend the range of communication actions and management of health risks.

The synergy observed upon interaction of the various institutions and input on information from each of them, but above all, the possibility of integrating the formal discussion of the research results on the dynamics of pollutants allowed not only targets of the entity he summoned, but progress in identifying other areas of opportunity for environmental management and quality of life as exposure geographic areas, population groups of interest, productive activities associated to environmental and health exposure risks plus effective risk communication strategies about the quality air and other environmental non anthropogenic factors as UV radiation.

III.CONCLUSIONS

The face state of concentration levels expected for suspended particle material (PM10 and PM2.5) showed high stability, meaning that no changes in concentration or exposure to contamination increase are expected, not in the immediate time. In addition, the behaviour observed in dry and rainy throughout the year is similar to the hourly data throughout the day, which implies self-similarity in micro and macro scales, aspect that could be explored in order to define the micro dynamic behaviour of air suspended matter. These observations guided the intersectoral working group to develop strategies for managing risks: Exposure geographic areas and population groups defined by productive activities and lifestyles plus risk communication strategies.

The synergy observed upon interaction of the various institutions and input on information from each of them was possible because of the efficient coordination, the respectful environment in which the work was made and the conjunction of particular interest of each one. On the other hand, the possibility of integrating the formal discussion of the research results on the dynamics of pollutants, allowed not only get the targets summoned, but in identifying other areas of opportunity for environmental and social research in order to improve the ecological resilience of the city.

REFERENCES

[1] Tang, J., Wang, L., & Yao, Z. (2008). Analyses of urban landscape dynamics using multi-temporal satellite images: A comparison of two petroleum-oriented cities. Landscape and Urban Planning, 87(4), 269–278.

[2] Plaia, A., Di Salvo, F., Ruggieri, M., & Agró, G. (2013). A Multisite-Multipollutant Air Quality Index. Atmospheric Environment, 70, 387–391.

[3] Plaia, A., & Ruggieri, M. (2010). Air quality indices: a review. Reviews in Environmental Science and Bio/Technology, 10, 165–179.

[4] Srogi, K. (2007). Monitoring of environmental exposure to polycyclic aromatic hydrocarbons: a review. Environmental Chemistry Letters, 5(4), 169–195.

[5] Andrews, G. J., Chen, S., & Myers, S. (2014). The “taking place” of health and wellbeing: towards non-representational theory. Social science & medicine (1982), 108, 210–22.

[6] Eakin, H., & Luers, A. L. (2006). Assessing the vulnerability of Social-Environmental systems. Annual Review of Environment and Resources, 31(1), 365–394. doi:10.1146/annurev.energy.30.050504.144352

[7] Instituto Nacional de Estadística y Geografía (INEGI), web page: www.inegi.org.mx. 25/07/2010.

[8] Flores Carro, Diana K. (2011). Determinación de zonas de riesgo a la salud por la incidencia de infecciones respiratorias agudas y la presencia de contaminantes ambientales en el Valle de Puebla, a través de la Evaluación Multicriterio. Durante el periodo 2004-2008. Tesis de Maestría. Universidad Autónoma de Tlaxcala, CIISDER. Tlaxcala, México. 148 p.

[9] Gabriel & Neumann (2006): A Markov chain model for daily rainfall occurrence at Tel Aviv.

[10] Heyde, C.C.; Seneta, Eugene (1977). I.J. Bienaymé : statistical theory anticipated, Studies in the history of mathematics and physical sciences, 3, Berlín: Springer Verlag, ISBN 3540902619.

[11] Langville, Amy N.; Meyer, Carl D. (2006). Google's PageRank and Beyond: The Science of Search Engine Rankings. Princeton University Press. ISBN 0-691-12202-4.

[12] Kimura, 1968. Evolutionary rate at the molecular level. En: Nature 217, 1968. pp. 624-626
[13] Cycling 74, 2003. Extraído de https://cycling74.com/.