EREM 81/3
Journal of Environmental Research, Engineering and Management
Vol. 81 / No. 3 / 2025
pp. 124–133
10.5755/j01.erem.81.3.38701

Particulate Matter 2.5 Prediction in Osorno, Chile, Using a Discrete Markov Chain Model

Received 2024/09 | Accepted after revisions 2025/07

https://doi.org/10.5755/j01.erem.81.3.38701

Particulate Matter 2.5 Prediction in Osorno, Chile, Using a Discrete Markov Chain Model

Dafne Lagos*, Dabbo Campos
Industrial Processes Department, Faculty of engineering, Universidad Católica de Temuco, Chile

***Corresponding author:** dlagos@uct.cl

Air quality is of growing concern globally due to its impact on human health. One of the most important air pollutants is particulate matter, principally fine particulate matter (PM 2.5). This study was carried out in the city of Osorno, Chile, where high levels of PM 2.5 are recorded, specifically during autumn and winter. A large database was assembled of daily PM 2.5 concentrations, precipitations, ambient temperature and mean wind speed in the months of April to September from 2013 to 2023. Using a discrete Markov chain model, the evolution and projection of PM 2.5 concentrations associated with existing weather conditions (environmental variables) were analysed. The analysis showed that the combination of low PM 2.5 concentration, cold temperature, presence of precipitations and wind is the commonest and most stable state, while states with high PM 2.5 concentration are more probable when conditions of cold temperature without rain or wind occur. The transition matrix drawn up enabled us to identify patterns of change and recurrence, showing the importance of weather factors in the accumulation or dispersal of pollutants and the associated probabilities. For example, state 24 (Very high PM2.5 – Cold – No rain – No wind), considered an "emergency" level in Chile, has an average recurrence time of 18.875 days and a 14.29% probability of transitioning to state 3 (Low PM2.5 – Cold – With rain – With wind), which in turn has a 17.662% of self-transition probability. The results provide a basis from which to project air quality conditions and plan preventive measures.

Keywords: Markov chains, particulate matter 2.5, temperature, wind and precipitations.

Introduction

Air quality and its effects on public health constitute a very important global problem. From the air that we breathe in our cities to that which we share in our homes, air pollution presents critical challenges for human beings and is a major world health problem. The principal sources of pollution in ambient air are vehicle traffic, factories and domestic fuels (Zhou et al., 2024). This is moreover a problem that affects everyone, since approximately 99% of people breathe air containing significant pollutants at higher levels than the maximum recommended by the WHO (Azimi and Rahman, 2024).

Air pollution consists of a number of components. This article considers coarse and fine particulate matter which differ, among their other characteristics, in that coarse particles are eliminated relatively quickly from the atmosphere by the effects of gravity and other processes, while fine particles remain in the air for longer and can drift from one area to another (Munir, 2017). These minute particles, generally invisible to the naked eye, have a diameter of 2.5 μm or less; they are known as Particulate Matter 2.5 (PM 2.5), and constitute a key indicator of air quality. Because they are so fine, PM 2.5 may remain suspended in the air for hours, and can enter the bloodstream directly, where they have an impact on health (Ho and Lin, 2024). Various studies have confirmed that PM 2.5 not only damages the respiratory and cardiovascular systems, but also causes symptoms such as diabetes and cancer (Ho and Lin, 2024). Studies have shown a strong relation between PM 2.5 particle pollution and lung cancer (Fei et al., 2024). Strong evidence also exists of a relation between increased PM 2.5 exposure and higher risk of autism spectrum disorder (ASD) (Jin et al., 2024). In Chile, positive correlations have been observed between environmental particle levels, mortality, hospital admissions and acute respiratory infections due to cardiovascular and respiratory diseases (Nakamura et al., 2022). These are just some of the problems caused by particle pollution.

Previous investigations have shown that a variety of meteorological factors, such as temperature, relative humidity and wind speed (Nakamura et al., 2022), play a significant role in the concentrations of air pollutants. As weather factors may have a direct impact on the time that pollutants remain suspended, as well as on the direct causes of particle pollution, it should be noted that PM 2.5 may consist of many components, and that these vary depending on their origin. In southern Chile, for example, the principal component of fine particles is organic matter, accounting for more than 70% of PM 2.5 (Jorquera et al., 2021). There are also estimates suggesting that residential combustion may account for up to 40% of the environmental PM 2.5 in some low- and medium-income countries (Odo et al., 2023).

Vieira de Oliveira Salerno et al. note the high rates of exposure to PM 2.5 in the whole of South America (Vieira de Oliveira Salerno et al., 2023), while Carreño et al. indicate that several of the most polluted cities in South America are located in Chile (Carreño et al., 2022).

Firewood burning is the primary source of PM2.5 pollution in southern Chilean cities, responsible for 84.6% of emissions, followed by inorganic particles (4.8%) and coal combustion (4.4%) in cities like Temuco, which has similar conditions to Osorno (Villalobos et al., 2017). Air pollution also causes serious health impacts in the population. As a result, depending on the sector, restrictions are applied to emission sources when the mean emissions of PM 2.5 in a 24-hour period exceed 80 μg/m^3 (Perez and Gramsch, 2016). The study of Nakamura et al. (2022), which showed the characteristics of PM 2.5 pollution in Osorno, observed the times of day when the concentration of particulate matter was highest, detecting large peaks of PM 2.5 concentration between 19:00 and the 23:00 (local time), and two small peaks between 07:00 and 12:00. These peaks occurred during the period when many of the local population are engaged in domestic activities (Nakamura et al., 2022). The Atmospheric Decontamination Plan for Osorno (PDAO), active from April to September, and which depends on the Ministerio del Medio Ambiente (MMA), imposes restrictions during pre-emergency or emergency episodes, including limits on residential firewood heaters, industrial boilers, and use of multiple firewood devices (MMA, 2024)

The city of Osorno has the third highest PM 2.5 pollution levels in Chile, with an annual average of 30 μg/m^3 in 2023. The two highest were Coyhaique with 38 μg/m^3 and Padre Las Casas with 33 μg/m^3 (Fernández et al., 2024). Osorno also ranks as number 52 for air pollution in South America (IQAir, 2024).

Given that PM 2.5 depends on various weather factors, its concentration changes frequently. Because these changes tend to occur unexpectedly, investigators apply stochastic mathematical models to analyse them (Caraka et al., 2019). Markov chains are stochastic mathematical models which can model random phenomena that evolve over time, allowing the future state of a system to be predicted based on its current state and a transition probability matrix. Markov models are based on a few, non-restrictive hypotheses, and can be calibrated and applied to a database containing a set of data such as is normally available in practice, for example a set of particulate matter concentrations (Carpinone et al., 2015). The future condition of the system is completely determined by its current state and not by the sequence of events which led to that state. This property, known as "no memory" (or "memoryless"), allows simple results to be obtained by calculating probabilities and other indicators of interest (Batún et al., 2023). Some examples of research applying discrete-time Markov chains to the study of pollution include: a) the study of air and noise pollution in Oyo State (Nigeria) (Ogunlade et al., 2024),b) the analysis of air quality indices, such as Ontario's Air Quality Health Index (AQHI), which incorporates pollutants such as ozone, nitrogen dioxide, and fine particulate matter (Holmes and Hassini, 2021), c) predicting the Air Quality Index (AQI) and identifying major air pollutants such as O3, NO2, and PM10 in Taipei (Chen and Wu, 2020), d) forecasting the Air Pollution Index (API) in Miri, Sarawak (Zakaria et al., 2019), e) the study the persistence of polluted days with carbon monoxide(Rahimi et al., 2011), f) the analysis of air quality levels(Yousefi Kebriya and Nadi, 2024), among others. Based on the above, the object of this work was to determine the probability of the occurrence of high levels of PM 2.5 in the city of Osorno under different weather conditions – specifically wind speed, ambient temperature and quantity of precipitations – by means of a discrete Markov chain model. In this way we expect to be able to provide timely information to support the authorities in taking proper decisions, and thus planning interventions to improve the air quality in the city.

Methods

Procedure

Markov chains are a widely-used tool for analysing stochastic processes. In analysing environmental data, such as predictions of PM 2.5 concentrations, certain assumptions must be made on the distribution of the variables involved. A discrete time stochastic process is therefore considered to be a collection of random variables X_t, where the t index adopts values from a given T, and corresponds to a set of non-negative integers (t = 0,1,2...). Each random variable X_t represents a characteristic of interest that is quantifiable in time t. A stochastic process is a Markov chain if it is memoryless.

A X_t stochastic process is Markovian, or memoryless, if the conditional probability of any future event, given the current state and any past event, is independent of past events and only depends on the current state of the process. It is described as Markovian if for $t = 0,1,...,$ and the whole succession i, j, k_0, k_1,...,k_{t-1}. (Batún et al., 2023):

$$P\{X_{t+1} = j|X_0 = k_0, X_1 = k_1, \dots, X_{t-1} = \\ = k_{t-1}, X_t = i\} = P\{X_{t+1} = j|X_t = i\} \tag{1}$$

The conditional probabilities in a Markov chain are called one-step transition probabilities. If these probabilities do not change with time, they are said to be stationary. This implies that for each i and j: $P(X_{t+1} = j \mid X_t = i) = P(X_1 = j \mid X_0 = i)$, for any t = 1, 2, ... Then for probabilities of transition of n steps: $P(X_{t+n} = j \mid X_t = i) = P(X_n = j \mid X_0 = i)$, simplifiying the equation (Batún et al., 2023)

$$p_{ij}^{(n)} = P(X_{t+n} = j|X_t = i) \tag{2}$$

where (1) must meet the following properties:

$$p_{ij}^{(n)} \geq 0, \qquad \forall\, i, j\,;\, n = 0, 1, 2, \cdots \tag{3}$$

$$\textstyle\sum_{j=0}^{M} p_{ij}^{(n)} = 1, \qquad \forall\ i\,;\, n = 0, 1, 2, \cdots \tag{4}$$

The n-step transition matrix is then represented by (Batún et al., 2023):

$$p^{(n)} = \begin{array}{c} State \\ 0 \\ 1 \\ \vdots \\ M \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \dots & M \end{array} \\ \begin{bmatrix} p_{00}^{(n)} & p_{01}^{(n)} & \dots & p_{0M}^{(n)} \\ p_{10}^{(n)} & p_{11}^{(n)} & \dots & p_{1M}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M0}^{(n)} & p_{M1}^{(n)} & \dots & p_{MM}^{(n)} \end{bmatrix} \end{array} \tag{5}$$

Some of the characteristics of this matrix are the classes of states that it contains, such as recurrent states. A state in a Markov chain is considered recurrent if, once the process enters that state, it may return to it in the future. If the Markov chain contains only one class, i.e. if all the states intercommunicate, it is said to be irreducible. On the other hand, a transition matrix of a Markov chain is considered ergodic if all its states are aperiodic and recurrent. In simple terms, a Markov chain is ergodic when each state of the chain is aperiodic, meaning that it does not follow a regular pattern of visits over time. In other words, there is no fixed number of steps after which the chain inevitably returns to the same state.

To obtain a transition probability $p_{ij}^{(n)}$ in n steps for a database in which every combination of data is identified with a state, the frequency identified in each of the states is identified and the probability of moving from an initial state i to a destination state j is calculated (Guerry, 2013).

$$p_{ij}^{(n)} = P(X_{t+n} = j | X_t = i) \quad (6)$$

where z_{ij} (t, t + 1) represents the number of data in state i at moment t which transition to state j at moment t + 1 (for t = 0,...T − 1), and z_i (t) refers to the number of data in state i at time t.

One of the most important long-term properties of Markov chains is the stable state property called π_j, which presents the probability of terminating in state j in a large number of transitions. Only for an irreducible ergodic matrix, $\lim_{n\to\infty} p_{ij}^{(n)}$ exists (independent of i). The values of π_j are obtained by the following system of equations formed by the following expressions (Torres Delgado et al., 2023):

$$\pi_j = \sum_{i=0}^{M} \pi_j \cdot p_{ij}, \quad \forall\, j\{0, \dots, M\} \quad (7)$$

$$\sum_{j=0}^{M} \pi_j = 1 \quad (8)$$

The expected first passage time is a statistical measure which tells us the mean time taken to transition from one state to another in a Markov chain for the first time. Where μ_{ij} represents the expected first passage time from a state i to a state j, expressed below (Afzal et al., 2019):

$$p_{ij}^{(n)} = P(X_{t+n} = j | X_t = i) \quad (9)$$

The expected recurrence time refers to the mean time taken to return to a specific state after leaving it. If the initial state and the destination state are the same, the expected first passage time becomes the expected recurrence time for that state.

Data collection

To carry out the study and calculate the probabilities of the Markov chain, environmental data were compiled from the city of Osorno, Chile. Daily concentrations of particulate matter (μg/m^3) in Osorno were obtained. These data were provided by National Air Quality Information System (SIMCA, 2023), a platform that offers validated, standardised data on air quality in different regions of Chile. We also compiled the daily records of precipitation in millimetres (mm), mean daily temperatures in degrees Celsius (°C) and the daily wind speed (km/h), regardless of direction. These characteristics were selected based on their established relationship with pollution levels, as identified in the study by Nakamura et al. (2022). The data for precipitation, temperature and wind speed were obtained from Cañal Bajo aerodrome at Osorno (the only monitoring station), which maintains precise daily measurements (DMC, 2023).

Data for the years 2013 to 2023 (11 years) were included, taken during the autumn and winter (1 April to 15 September) since this is considered the critical period for PM 2.5 in Osorno (Fernández et al., 2024), mainly due to increased firewood use for heating. The data collected were reviewed prior to analysis to identify possible anomalies or missing data. Where missing or anomalous data were found, mean values for the same day in the other years were applied.

The PM 2.5 concentration was classified in four categories based on the value in microgrammes per cubic metre (μg/m^3). Values of 0 to 79 μg/m^3 were classified as “Low level of PM 2.5”, indicating relatively safe levels of air pollution. Values from 80 to 109 μg/m^3 were considered “Medium level of PM 2.5”, representing a moderate increase in pollution. In Chile this concentration is defined as the “Alert” level. Values of 110 to 169 μg/m^3 were classified as “High level of PM 2.5”, indicating a worrying state of air quality that could affect human health. In Chile this concentration is defined as the “Pre-emergency” level. Finally, any value of 170 μg/m^3 or more was classified as “Very high level of PM 2.5”, indicating severe pollution; this is defined as an “Emergency” level in Chile (MMA, 2011).

The precipitation was classified in two categories: “Rain” and “No Rain”, based on the precipitation measured in millimetres (mm) per day. A value greater than 1 mm was treated as “Rain”, indicating significant precipitation that could influence the dispersion of atmospheric pollutants. A value less than or equal to 1 mm was classified as “No Rain” (DGAC, 2021).

The daily temperature was classified in two categories: “Cold” and “Warm”. Temperatures lower than 10°C were classified as “Cold”. This condition may affect the dispersal of pollutants or the accumulation of particulate matter because of the methods used by the inhabitants to heat their homes (Odo et al., 2023), since firewood is the preferred fuel because of its low cost (Navarro-Espinosa and Thomas-Galán, 2023). Temperatures higher than 10°C were classified as “Warm”. This is also relevant since higher temperatures increase the concentration of particulate matter (Pateraki et al., 2012).

The approximate daily wind speed was classified in two categories: "No Wind" and "Wind". Wind speeds lower than 6 km/h were classified as "No Wind", since this condition can affect the dispersal of pollutants by allowing PM 2.5 to accumulate in the atmosphere. Wind speeds greater than or equal to 6 km/h were classified as "Wind", a condition more favourable to the dispersal of pollutants, reducing concentrations of particulate matter.

The categories of the variables used in the definition of states differ from those used in other studies.

Results and discussion

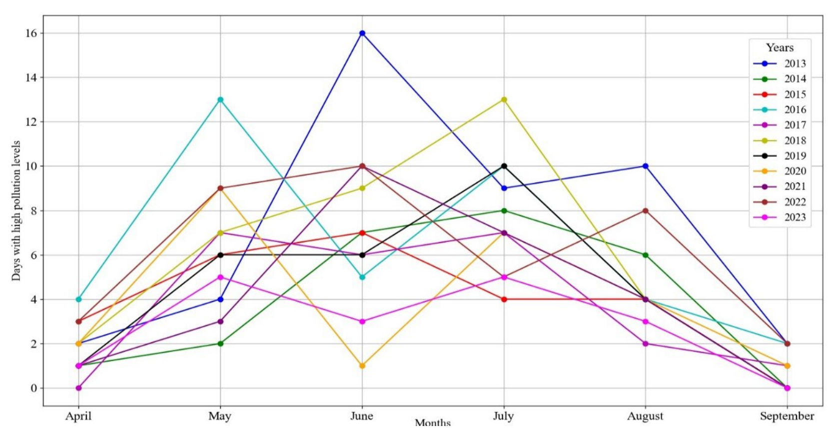
To implement the Markov chain model, states were defined as combinations of these four categories: PM 2.5 concentration, precipitation, temperature and wind speed. This generated a total of 32 states, as shown in *Table 1*.

A preliminary review of high-concentration data during selected months is shown in *Fig. 1*.

Table 1. Definition of states in relation to PM2.5 and weather conditions

State	PM2.5/Temperature/Precipitation/Wind	State	PM2.5/Temperature/Precipitation/Wind
0	Low – Cold – No rain – No wind	16	High – Cold – No rain – No wind
1	Low – Cold – No rain – Wind	17	High – Cold – No rain – Wind
2	Low – Cold – Rain – No wind	18	High – Cold – Rain – No wind
3	Low – Cold – Rain – Wind	19	High – Cold – Rain – Wind
4	Low – Warm – No rain – No wind	20	High – Warm – No rain – No wind
5	Low – Warm – No rain – Wind	21	High – Warm – No rain – Wind
6	Low – Warm – Rain – No wind	22	High – Warm – Rain – No wind
7	Low – Warm – Rain – Wind	23	High – Warm – Rain – Wind
8	Medium – Cold – No rain – No wind	24	Very high – Cold – No rain – No wind
9	Medium – Cold – No rain – Wind	25	Very high – Cold – No rain – Wind
10	Medium – Cold – Rain – No wind	26	Very high – Cold – Rain – No wind
11	Medium – Cold – Rain – Wind	27	Very high – Cold – Rain – Wind
12	Medium – Warm – No rain – No wind	28	Very high – Warm – No rain – No wind
13	Medium – Warm – No rain – Wind	29	Very high – Warm – No rain – Wind
14	Medium – Warm – Rain – No wind	30	Very high – Warm – Rain – No wind
15	Medium – Warm – Rain – Wind	31	Very high – Warm – Rain – Wind

Fig. 1. Frequency of days with high pollution in different months of each year



From *Fig. 1*, it is evident that the frequency of high PM2.5 concentrations tends to decline as the summer months approach (from January to April and from September to December). This pattern aligns with the findings of Molina et al. (2017), therefore, these months were excluded from the analysis.

The historical data of the variables were used to calculate the frequencies of transitions between the different states, thus obtaining the transition probabilities. The transition matrix was constructed with the transition probabilities obtained in this way. The Python code used for the calculations, based on the equations presented in the Methods / *Procedures* section, is available at the following link: https://github.com/ParticulateMatterOsorno/Markov-chain/blob/main/Python%20code.PY

Fig. 2 presents a bubble chart showing the occurrence of each state, with the X-axis representing the current state and the Y-axis the state on the following day.

There are no data for states 14, 29, 30 and 31 in the study case, meaning that these states did not occur during the observation period, or are extremely rare. This may be due to the nature of the data set collected or to the geographical area, in which certain specific weather conditions do not occur. These states were therefore eliminated from the matrix.

Fig. 2. *Frequency of transitions in different states*

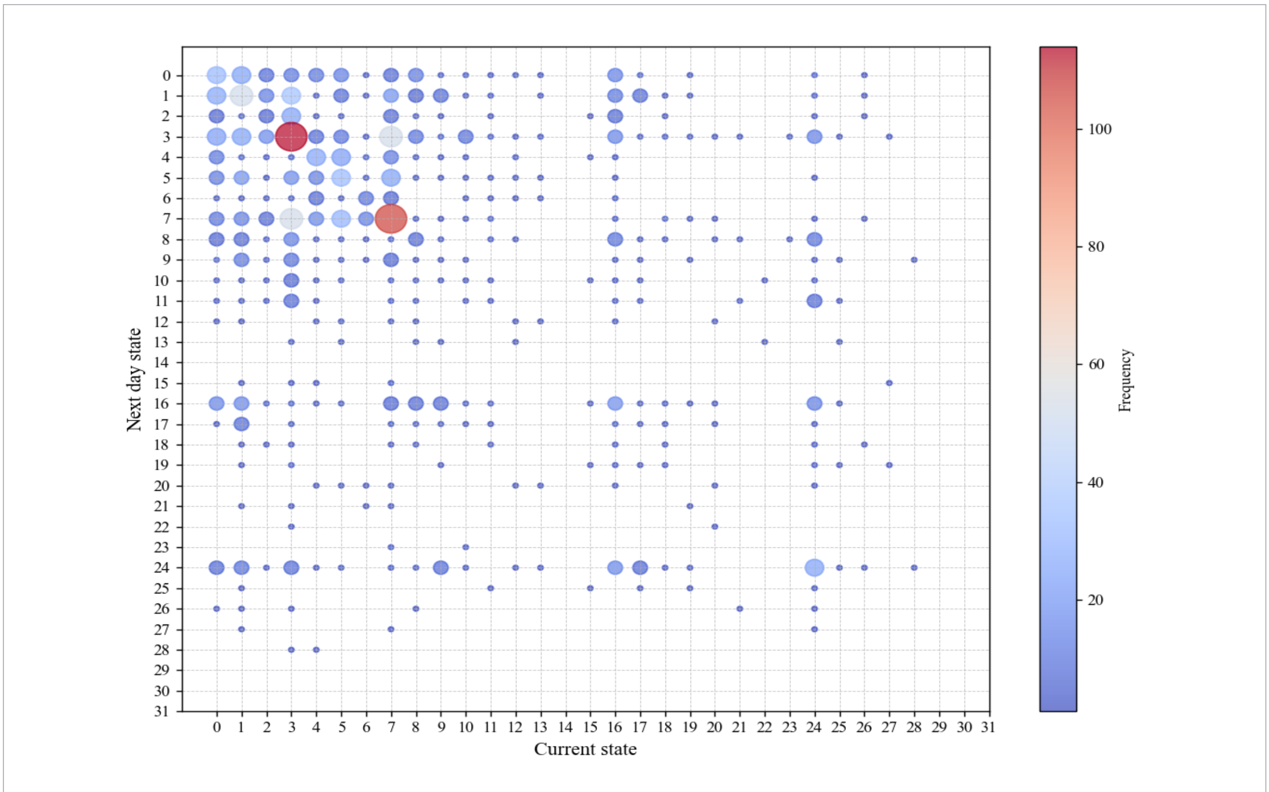


Fig. 3 shows the resulting matrix, with probabilities calculated to three decimal places.

Some transitions are seen to have a high probability, meaning that certain combinations of PM 2.5 concentration and weather conditions are more common than others. This shows how the air quality and weather conditions may evolve in specific scenarios.

We observe that some states have a high probability of persisting for a period of days. For example, State 3 (Low level of PM 2.5 - Cold - Rain - Wind) shows a high probability (34.97%) of remaining unchanged. This high probability implies that these conditions of air quality and weather are relatively stable and tend to persist once they occur. Similarly, State 7 (Low level of

Fig. 3. *Transition matrix for the defined states*

$P_{ij}=$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19	20	21	22	23	24	25	26	27	28
0	0.199	0.161	0.037	0.143	0.068	0.075	0.012	0.062	0.037	0.031	0.012	0.006	0.006	0	0	0.093	0.006	0	0	0	0	0	0	0.043	0	0.006	0	0
1	0.116	0.251	0.024	0.116	0.024	0.092	0.01	0.063	0.034	0.053	0.01	0.01	0.005	0	0.005	0.077	0.039	0.005	0.005	0	0.005	0	0	0.043	0.005	0.005	0.005	0
2	0.114	0.186	0.086	0.243	0.029	0.014	0.029	0.1	0.043	0.043	0.014	0.014	0	0	0	0.057	0	0.014	0	0	0	0	0	0.014	0	0	0	0
3	0.034	0.11	0.074	0.350	0.012	0.052	0.012	0.163	0.04	0.028	0.018	0.021	0	0.006	0.003	0.015	0.009	0.003	0.012	0	0.003	0.003	0	0.025	0	0.003	0	0.003
4	0.112	0.031	0.02	0.082	0.255	0.133	0.071	0.163	0.02	0.01	0.01	0.01	0.02	0	0.01	0.02	0	0	0	0.01	0	0	0	0.01	0	0	0	0.01
5	0.094	0.058	0.014	0.072	0.167	0.239	0.014	0.217	0.014	0.014	0.022	0	0.022	0.007	0	0.022	0	0	0	0.007	0	0	0	0.014	0	0	0	0
6	0.048	0.048	0	0.048	0.119	0.095	0.238	0.31	0.024	0.024	0	0	0	0	0	0	0	0	0	0.024	0.024	0	0	0	0	0	0	0
7	0.026	0.07	0.026	0.196	0.044	0.089	0.022	0.391	0.015	0.022	0.011	0.007	0.007	0	0.011	0.022	0.011	0.004	0	0.004	0.004	0	0.007	0.007	0	0	0.004	0
8	0.16	0.08	0.053	0.133	0.04	0.04	0	0.04	0.093	0.013	0.04	0.027	0.013	0.027	0	0.093	0.053	0.04	0	0	0	0	0	0.04	0	0.013	0	0
9	0.083	0.167	0.021	0.104	0.021	0.042	0	0.021	0.063	0.021	0.021	0	0	0.021	0	0.146	0.083	0	0.021	0	0	0	0	0.167	0	0	0	0
10	0.129	0.097	0	0.29	0.032	0.065	0.032	0.065	0	0.032	0.032	0.032	0	0	0	0.097	0.032	0	0	0	0	0	0.032	0.032	0	0	0	0
11	0.063	0.094	0.031	0.125	0	0.063	0.031	0.156	0.094	0	0.031	0.031	0	0	0	0.125	0.031	0.094	0	0	0	0	0	0	0.031	0	0	0
12	0.188	0	0	0.063	0.188	0.188	0.063	0	0.063	0	0	0	0.063	0.063	0	0	0	0	0	0.063	0	0	0	0.063	0	0	0	0
13	0.111	0.111	0	0.222	0	0.111	0.111	0	0	0	0	0	0.111	0	0	0	0	0	0	0.111	0	0	0	0.111	0	0	0	0
15	0	0	0.143	0	0.143	0	0	0	0	0	0.143	0	0	0	0	0.286	0	0	0.143	0	0	0	0	0	0.143	0	0	0
16	0.124	0.097	0.062	0.133	0.009	0.009	0.027	0.027	0.08	0.009	0.009	0.018	0.018	0	0	0.159	0.035	0.018	0.018	0.009	0	0	0	0.142	0	0	0	0
17	0.143	0.229	0	0.086	0	0	0	0	0.029	0.029	0.029	0.057	0	0	0	0.086	0.057	0	0.029	0	0	0	0	0.2	0.029	0	0	0
18	0	0.059	0.059	0.176	0	0	0	0.294	0.059	0	0	0	0	0	0	0.118	0.059	0.059	0.059	0	0	0	0	0.059	0	0	0	0
19	0.125	0.125	0	0.188	0	0	0	0.125	0	0.063	0	0	0	0	0	0.063	0	0	0	0	0.063	0	0	0.188	0.063	0	0	0
20	0	0	0	0.111	0	0	0	0.111	0.111	0	0	0	0.222	0	0	0.111	0.111	0	0	0.111	0	0.111	0	0	0	0	0	0
21	0	0	0	0.2	0	0	0	0	0.2	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0
22	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0.667	0	0	0	0	0.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0.051	0.041	0.02	0.143	0	0.01	0	0.01	0.092	0.02	0.031	0.071	0	0	0	0.133	0.02	0.031	0.031	0.01	0	0	0	0.245	0.02	0.01	0.01	0
25	0	0	0	0.143	0	0	0	0	0	0.143	0	0.143	0	0.143	0	0.143	0	0	0.143	0	0	0	0	0.143	0	0	0	0
26	0.167	0.167	0.167	0	0	0	0	0.167	0	0	0	0	0	0	0	0	0	0.167	0	0	0	0	0	0.167	0	0	0	0
27	0	0	0	0.333	0	0	0	0	0	0	0	0	0	0	0.333	0	0	0	0.333	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0

PM 2.5 - Warm - Rain - Wind) also shows a high probability (39.11%) of remaining unchanged. This indicates considerable stability in conditions of this kind, reflecting weather patterns encountered in Osorno.

Another important result is the existence of certain transitions with a probability that is very low or close to zero. These unlikely transitions are found in certain combinations of PM 2.5 levels with weather conditions that are unusual or improbable for the system modelled. For example, states with high or very high levels of PM 2.5 with rain and wind are unlikely to occur. There are also states with a low probability of persisting. For example, there are several states along the diagonal of the matrix with 0% possibility of persisting, suggesting great instability and a significant tendency to transit rapidly to another state.

Among the states with Very high levels of PM 2.5, State 24 (Very high level of PM 2.5 - Cold - No Rain - No Wind) presents 24.49% probability of remaining unchanged, indicating high persistence of this combination of pollution and weather conditions. This situation is completely different to the other states with Very high levels of PM 2.5, since they present 0% probability of remaining unchanged. State 24 is also interesting in that it presents several probabilities of transition to another state. The most representative case is the possibility of transiting from State 24 to State 3 (Low level of PM 2.5 - Cold - Rain - Wind), with a probability of 14.29%; this is indicative of the strong impact of wind and rain on particulate matter levels. An important feature of the PM 2.5 concentration and weather conditions defined in State 24 is the fact that it is the most frequent combination that gives rise to the definition of the air quality situation classed as an “Emergency” under Chilean law. The transition matrix indicates the probability of this condition persisting or changing.

According to the order of the matrix results shown in *Fig. 3*, it can be seen that the highest concentration of probabilities for transition, other than zero, occur in the states that include a low level of PM 2.5, followed, to a lesser degree, by states with medium PM 2.5 levels.

The figure also shows which states have a tendency to transition to a high concentration of PM 2.5; for example, changes in this direction are more frequent in States 16 and 24 than in other states with equivalent levels of PM 2.5. We likewise see an increase in PM 2.5 with cold mean temperature and no rain or wind; while in states with similar conditions but a warmer temperature, the probabilities of transition are considerably lower. This is because, although higher temperatures increase the concentration of pollutants in the air, the intensive use of polluting domestic heating devices in cold weather has a greater impact on pollution levels.

The probability information determined in the matrix of *Fig. 3* allows us to establish the probability of a stationary state for each state defined, as well as the recurrence time. This information is shown in *Table 2*.

Table 2. *Steady state probabilities and recurrence times of the defined states*

States	0	1	2	3	4	5	6
π_j (%)	8.71	11.211	3.791	17.662	5.235	7.467	2.269
μ_{ii} (days)	11.482	8.92	26.378	5.662	19.102	13.392	44.066
States	7	8	9	10	11	12	13
π_j (%)	14.742	4.061	2.6	1.679	1.733	0.865	0.487
μ_{ii} (days)	6.784	24.627	38.468	59.574	57.711	115.614	205.24
States	15	16	17	18	19	20	21
π_j (%)	0.379	6.118	1.896	0.921	0.867	0.486	0.271
μ_{ii} (days)	263.726	16.345	52.742	108.603	115.389	205.554	369.097
States	22	23	24	25	26	27	28
π_j (%)	0.108	0.163	5.306	0.379	0.325	0.163	0.108
μ_{ii} (days)	923.94	613.717	18.847	263.785	307.783	614.632	929.403

According to *Table 2*, State 3 (Low level of PM 2.5 - Cold - Rain - Wind) is the most likely to occur (probability 17.66%), indicating that this is the commonest (most probable) state in the long term. It is followed by State 7 (Low level of PM 2.5 - Warm - Rain - Wind) with a probability of 14.74%. States with a lower probability of occurrence include State 22 (High level of PM 2.5 - Warm - Rain - Wind) and 28 (Very high level of PM 2.5 - Warm - Rain - Wind) with only 0.11% each. These states are therefore extremely rare.

In parallel, the expected recurrence times show that, on average, State 3 recurs after only 5.66 days. On the other hand, states with a low probability, such as State 22 and State 28, present very high recurrence times, of 923.94 and 929.40 days respectively.

The transition matrix was developed using data from the city of Osorno, and is therefore valid in this context. However, the defined states, variables used, and the processing mechanism are replicable for other geographical areas in the country, potentially generating valuable information in those contexts.

Using the transition matrix and the stationary states, both short- and long-term predictions can be made of the air quality and weather conditions. Thus future scenarios can be anticipated and suitable responses prepared, improving air quality management. Detailed analysis of the transition matrix enables local authorities to develop better informed policies and take preventive or mitigation measures based on the probability of the occurrence of states of high pollution. This is necessary in order to implement effective strategies to protect public health and reduce pollution levels. The model also provides a basis for further studies to adjust and improve its accuracy. This investigation work can contribute to the development of more robust and accurate models, further improving air quality prediction and management in Osorno.

Conclusions

The development of a transition matrix based on a discrete Markov chain allows us to model the concentration of fine particulate matter (PM 2.5) in the city of Osorno. This analysis reveals the patterns of change in the combinations of particulate matter 2.5 concentration with daily weather characteristics. Common transitions between states were identified, for example transition from State 19 to State 24 with a probability of 18.8%. The analysis also highlighted states with a low probability of remaining unchanged, reflecting the high instability of certain combinations of PM 2.5 and weather variables. This knowledge can add value to current tools used for applying measures under the Osorno Air Decontamination Plan (PDAO). These findings are important for identifying when and how large variations in the PM 2.5 concentration may occur, thus providing basic information for the implementation of more targeted and proactive mitigation strategies. Furthermore, identifying states with higher particulate matter 2.5 levels, and the probabilities of these changing or persisting, provides information on which to base the design of preventive actions that could reduce pollution without eliminating the community's primary heating method.

By defining states based on combinations of the PM 2.5 concentration with specific weather factors, it is possible to analyse air quality at a given moment and show the probabilities that air pollution will get better or worse, associated with predicted changes in the weather.

In the social context, specifically for people living in Osorno, this model offers an opportunity to develop better-informed policies and apply preventive measures based on the probability that states of high pollution will occur. This is essential for protecting public health and reducing pollution levels in the city

Acknowledgements

The authors thank the Research Department of Universidad Católica de Temuco for financial support for the translation of this document. They also thank the Centre for Industrial Management, Operations and Logistics of the Engineering Faculty of Universidad Católica de Temuco for suggesting the research topic.

References

Afzal H., Mufti M. R., Awan I., Yousaf M. (2019) Performance analysis of radio spectrum for cognitive radio wireless networks using discrete time Markov chain. Journal of Systems and Software 151: 1–7. Available at: https://doi.org/10.1016/j.jss.2019.01.053

Azimi M. N., Rahman M. M. (2024) Unveiling the health consequences of air pollution in the world's most polluted nations. Scientific Reports 14(1). Available at: https://doi.org/10.1038/s41598-024-60786-0

Batún J. L., Cool R., Pantí H. (2023) Analysis of academic trajectories of higher education students by means of an absorbing Markov chain. Revista de La Academia Colombiana de Ciencias Exactas, Fisicas y Naturales 47(185): 1008–1023. Available at: https://doi.org/10.18257/raccefyn.1973

Caraka R. E., Chen R. C., Toharudin T., Pardamean B., Yasin H., Wu S. H. (2019) Prediction of status particulate matter 2.5 using state Markov chain stochastic process and hybrid VAR-NN-PSO. IEEE Access 7: 161654–161665. Available at: https://doi.org/10.1109/ACCESS.2019.2950439

Carpinone A., Giorgio M., Langella R., Testa A. (2015) Markov chain modeling for very-short-term wind power forecasting. Electric Power Systems Research 122: 152–158. Available at: https://doi.org/10.1016/j.epsr.2014.12.025

Carreño G., López-Cortés X. A., Marchant C. (2022) Machine Learning Models to Predict Critical Episodes of Environmental Pollution for PM2.5 and PM10 in Talca, Chile. Mathematics 10(3): Available at: https://doi.org/10.3390/math10030373

Chen J. C., Wu Y. J. (2020) Discrete-time Markov chain for prediction of air quality index. Journal of Ambient Intelligence and Humanized Computing: Available at: https://doi.org/10.1007/s12652-020-02036-5

DGAC (Dirección General de Aeronáutica Civil) (2021) Reporte anual de la evolución del clima en Chile [Annual report on the evolution of the climate in Chile]. Available at: https://www.meteochile.gob.cl (*in Spanish*).

DMC (Dirección Meteorológica de Chile), Servicios Climáticos, Estación Cañal Bajo, Osorno Ad. (400009) [Climate Services]. Available at: https://climatologia.meteochile.gob.cl (accessed on 14 July 2024) (*in Spanish*).

Fei G., Li H., Yang S., Wang H., Ge Y., Wang Z., Zhang X., Wei P., Li L. (2024) Burden of lung cancer attributed to particulate matter pollution in China: an epidemiological study from 1990 to 2019. Public Health 227: 141–147. Available at: https://doi.org/10.1016/j.puhe.2023.12.005

Fernández R., Peña R., Bravo-Alvarado J., Maisey K. R., Reyes E. P., Ruiz-Plaza De Los Reyes D., Márquez-Reyes R. (2024) Prevalence Distribution of Chronic Obstructive Pulmonary Disease (COPD) in the City of Osorno (Chile) in 2018, and Its Association with Fine Particulate Matter PM2.5 Air Pollution. Atmosphere 15(4): 482. Available at: https://doi.org/10.3390/atmos15040482

Guerry M. A. (2013) On the embedding problem for discrete-time Markov chains. Journal of Applied Probability 50(4): 918–930. Available at: https://doi.org/10.1239/jap/1389370090

Ho Y. H., Lin Y. C. (2024) Searching for the fine particulate matter (PM 2.5) pollutant emission source using a drone. Measurement: Journal of the International Measurement Confederation 232: Available at: https://doi.org/10.1016/j.measurement.2024.114726

Holmes J., Hassini S. (2021) Discrete-Time Markov Chain Modelling of the Ontario Air Quality Health Index. Water, Air, and Soil Pollution 232(4): 158. Available at: https://doi.org/10.1007/s11270-021-05096-1

IQAir (2024) World's most polluted cities, Most polluted city ranking based on annual average PM2.5 concentration (μg/m^3) South America: Available at: https://www.iqair.com/world-most-polluted-cities

Jin T., Pang Q., Huang W., Xing D., He Z., Cao Z., Zhang T. (2024) Particulate matter 2.5 causally increased genetic risk of autism spectrum disorder. BMC Psychiatry 24(1): Available at: https://doi.org/10.1186/s12888-024-05564-y

Jorquera H., Villalobos A. M., Schauer J. J. (2021) Wood burning pollution in Chile: A tale of two mid-size cities. Atmospheric Pollution Research 12(4): 50–59. Available at: https://doi.org/10.1016/j.apr.2021.02.011

MMA (Ministerio del Medio Ambiente) (2024) Chile. Calidad del aire. Plan de descontaminación del aire Osorno [Air quality. Osorno air pollution control plan.]: Available at: https://pdao.mma.gob.cl/calidad-del-aire (accessed on 17 May 2025) (*in Spanish*).

MMA (Ministerio del Medio Ambiente), Chile (2011) Decreto 12. Establece norma primaria de calidad ambiental para material particulado fino respirable MP 2,5 [Decree 12. Establishes primary environmental quality standard for fine respirable particulate matter MP 2.5]: Available at: https://www.bcn.cl/leychile/navegar?idNorma=1025202 (accessed on 14 July 2024) (*in Spanish*).

Molina C., ToroA R., MoralesS R. G. E., Manzano C., Leiva-Guzmán M. A. (2017) Particulate matter in urban areas of south-central Chile exceeds air quality standards. Air Quality, Atmosphere and Health 10(5): 653–667. Available at: https://doi.org/10.1007/s11869-017-0459-y

Munir S. (2017) Analysing temporal trends in the ratios of PM2.5/PM10 in the UK. Aerosol and Air Quality Research 17(1): 34–48. Available at: https://doi.org/10.4209/aaqr.2016.02.0081

Nakamura A., Nakatani N., Maruyama F., Fujiyoshi S., Márquez-Reyes R., Fernández R., Noda J. (2022) Characteristics of PM2.5 Pollution in Osorno, Chile: Ion Chromatography and Meteorological Data Analyses. Atmosphere 13(2): 168. Available at: https://doi.org/10.3390/atmos13020168

Navarro-Espinosa A., Thomas-Galán M. (2023) Firewood electrification in Chile: effects on household expenditure and energy poverty. Energy Policy 173: Available at: https://doi.org/10.1016/j.enpol.2022.113337

Odo D. B., Yang I. A., Dey S., Hammer M. S., van Donkelaar A., Martin R. V., Dong G. H., Yang B. Y., Hystad P., Knibbs L. D. (2023) A cross-sectional analysis of ambient fine particulate matter (PM2.5) exposure and haemoglobin levels in children aged under 5 years living in 36 countries. Environmental Research 227: Available at: https://doi.org/10.1016/j.envres.2023.115734

Ogunlade T. O., Bakare K. A., Ogunyebi S. N., Fadugba S. E., Ogunmiloro O. M., Ogunrinde R. B., Adetunji S. V. (2024) Application of a Stochastic Model to Environmental Pollution. 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG): 1–7. Available at: https://doi.org/10.1109/SEB4SDG60871.2024.10630334

Pateraki S., Asimakopoulos D. N., Flocas H. A., Maggos T., Vasilakos C. (2012) The role of meteorology on different sized aerosol fractions (PM 10, PM 2.5, PM 2.5-10). Science of the Total Environment 419: 124–135. Available at: https://doi.org/10.1016/j.scitotenv.2011.12.064

Perez P., Gramsch E. (2016) Forecasting hourly PM2.5 in Santiago de Chile with emphasis on night episodes. Atmospheric Environment 124: 22–27. Available at: https://doi.org/10.1016/j.atmosenv.2015.11.016

SINCA (Sistema de Información Nacional de Calidad del Aire), Ministerio del Medio Ambiente [National Air Quality Information System]: Available at: https://sinca.mma.gob.cl/index.php/estacion/index/key/A01 (accessed on 14 July 2024) (*in Spanish*).

Torres Delgado M. X., Queiroz J., Átila O. (2023) Previsão de intervalos de preço no mercado de ações brasileiro usando cadeias de Markov de tempo discreto [Forecasting Price Intervals in the Brazilian Stock Market Using Discrete-Time Markov Chains]. Revista Brasileira de Computação Aplicada 15(1): 34–47. Available at: https://doi.org/10.5335/rbca.v15i1.13400 (*in Portuguese*).

Vieira de Oliveira Salerno P. R., Briones-Valdivieso C., Motairek I., Palma Dallan L. A., Rajagopalan S., Deo S. V., Petermann-Rocha F., Al-Kindi S. (2023) The cardiovascular disease burden attributable to particulate matter pollution in South America: analysis of the 1990–2019 global burden of disease. Public Health 224: 169–177. Available at: https://doi.org/10.1016/j.puhe.2023.07.035

Villalobos A. M., Barraza F., Jorquera H., Schauer J. J. (2017) Wood burning pollution in southern Chile: PM2.5 source apportionment using CMB and molecular markers. Environmental Pollution 225: 514–523. Available at: https://doi.org/10.1016/j.envpol.2017.02.069

Yousefi Kebriya A., Nadi M. (2024) Examining Air Pollution Continuity in Tehran Province using Markov Chain Model. Iranica Journal of Energy and Environment 15(2): 201–210. Available at: https://doi.org/10.5829/IJEE.2024.15.02.10

Zakaria N. N., Othman M., Sokkalingam R., Daud H., Abdullah L., Abdul Kadir E. (2019) Markov Chain Model Development for Forecasting Air Pollution Index of Miri, Sarawak. Sustainability 11(19): 5190. Available at: https://doi.org/10.3390/su11195190

Zhou R. X., Liao H. J., Hu J. J., Xiong H., Cai X. Y., Ye D. W. (2024) Global Burden of Lung Cancer Attributable to Household Fine Particulate Matter Pollution in 204 Countries and Territories, 1990 to 2019. Journal of Thoracic Oncology. Available at: https://doi.org/10.1016/J.JTHO.2024.01.014