Clements, Adam, Herrera, Rodrigo, & Hurn, Stan
(2020)
Network analysis: A novel approach to identify PM2.5 hotspots and their spatio-temporal impact on air quality in Santiago de Chile.
Air Quality, Atmosphere and Health, *13*(9), pp. 1075-1082.

https://doi.org/10.1007/s11869-020-00862-2

Noname manuscript No.
(will be inserted by the editor)

Network analysis: a novel approach to identify $\mathrm{PM}_{2.5}$ hotspots and their spatio-temporal impact on air quality in Santiago de Chile

Adam Clements
School of Economics and Finance,
Queensland University of Technology.

Rodrigo Herrera
Facultad de Economía y Negocios,
Universidad de Talca, Talca, Chile.

Stan Hurn
School of Economics and Finance,
Queensland University of Technology.

Received: date / Accepted: date

Abstract Air pollution, particularly $\mathrm{PM}_{2.5}$ particulate matter, is a significant issue in Santiago, the capital of Chile. Santiago's pollution problem is exacerbated by its unique geographic location nestled against the Andes mountain range in the central valley of Chile. This paper uses network models that were developed primarily to analyze systemic risk in the financial system to identify those locations in the city that are most important for explaining $\mathrm{PM}_{2.5}$ levels. High average concentrations are associated with both systemically important locations and those that are most sensitive to pollution arriving from other areas. A detailed picture of the links across the city can help direct official efforts to combat pollution.

Keywords Particulate matter · $\mathrm{PM}_{2.5}$ · Santiago · Networks

Adam Clements
2 George St, Brisbane, Australia
Tel.: +61 7 31385293
E-mail: a.clements@qut.edu.au
orcid: 0000-0002-4232-0323

Rodrigo Herrera
2 Norte 685 Talca, Chile
Tel.: +56-71 2200200
E-mail: Rodrigo.Herrera@utalca.cl
orcid: 0000-0002-4604-0396

Stan Hurn
2 George St, Brisbane, Australia
Tel.: +61 7 31385293
E-mail: s.hurn@qut.edu.au

1 Introduction

Air pollution is a major environmental problem that has both health and economic implications. The component of pollution of most concern is particulate matter with a diameter less than 2.5 microns, known as $PM_{2.5}$. Given the negative and potentially dangerous consequences of $PM_{2.5}$ pollution, it is not surprising that the issue of modelling its concentration levels has attracted a great deal of attention.

Chile is an extreme case in point. It is estimated that at least 60% of Chile's inhabitants are exposed to $PM_{2.5}$ concentrations over recommended levels. The annual U.S. norm is $15\,\mu g/m^3$, while the WHO suggests an annual limit of only $10\,\mu g/m^3$ (Cifuentes 2010). It is estimated that there are 4,000 premature deaths due to chronic exposure to this component of pollution in Chile (MMA 2011) and the net economic benefit to Chile of effectively regulating $PM_{2.5}$ is estimated to be USD 7.1 billion (SINIA 2010). The Chilean capital, Santiago, where 41% of the country's total population resides, is of particular concern because of its location in a low mountainous valley near the Andes, and the simultaneous presence of the Pacific Anticyclone and the phenomenon of thermal inversion which inhibits the dispersion of pollution. Consequently exposure to $PM_{2.5}$ pollution in Santiago has significant impacts on mortality and morbidity (Pino-Cortés, Díaz-Robles, Campos, Vallejo, Gómez, Cereceda-Balic, Fu, and Figueroa 2020), particularly during the winter months when air pollution increases significantly relative to the summer months.[1]

There are a number of studies aimed at forecasting air pollution and particularly $PM_{2.5}$ levels in Chile and this body of research employs a number of different methods. These techniques include forecasting based on artificial neural networks (Perez and Reyes 2006; Díaz-Robles, Ortega, Fu, Reed, Chow, Watson, and Moncada 2008; Perez and Gramsch 2016); weather and chemical based models (Saide, Carmichael, Spak, Gallardo, Osses, Mena, and Pagowski 2011; Saide, Mena, Tolvett, Hernandez, and Carmichael 2016); and a systems approach based on a hierarchical set of linear regression equations where forecasts at a location are related to observations at other locations (Moisan, Herrera, and Clements 2018). While predictive models are clearly important, a growing area of research considers the interaction between concentration levels observed at different geographical locations within small-scale geographical areas. Spatio-temporal models (Sahu, Gelfand, and Holland, 2006) account for the full spatial distribution of monitoring stations. In the case of Santiago, Nicolis, Díaz, Sahu, and Marín (2019) develop a Bayesian scheme for forecasting from a spatio-temporal model at a 1-km high resolution grid, using $PM_{2.5}$ measurements from the coarser network of measuring stations. It is also possible to use the spatial correlations between monitoring stations to develop information-based indices to determine the quality of the monitoring network. This work has been undertaken in Santiago by Osses, Gallardo, and Faundez (2013) and Henriquez, Osses, Gallardo, and Diaz Resquin (2015). The core idea in this analysis is to determine the amount of information generated by $PM_{2.5}$ observations at each station, in relation to concentrations across the entire Santiago area.

The spatial-temporal behaviour of pollution across the entire Santiago region is also the central focus of this paper. The primary objective is to identify the

[1] For the adverse impacts of $PM_{2.5}$ in in European cities see Maciejewska (2020).

locations, or hotspots, which have the greatest impact on $PM_{2.5}$ levels. To achieve this aim, the methodology developed by Diebold and Yilmaz (2014) for measuring systemic risk in the financial system is adapted to deal with the problem of modelling air pollution. The temporal dynamics of particulate matter is captured using a vector autoregressive (VAR) model. While there are clearly complex physical, chemical and meteorological processes underlying the dynamic behaviour of $PM_{2.5}$ measurements, VAR modeling is an efficient reduced form statistical approach that is particularly adept at capturing the interactions between observable data.[2] Publicly available weather data is easily incorporated in the methodology and can be safely assumed to be exogenous to the system.[3] In addition, one of the well-known strengths of VAR models is that the interactions between the observable data mitigate the effects of any omitted variables. For example, local emissions due to traffic flows are known to influence air pollution but are not included in the analysis because no data is available.

The network framework is perfectly suited to the task of modeling pollution as it provides estimates of directional links between each time series of $PM_{2.5}$ measured at individual monitoring stations throughout greater Santiago. Based on these directional links, two important summary measures of linkages can be constructed. *Fragility*, captures how sensitive concentrations at an individual station are in response to changes in concentrations at all other locations. *Centrality*, measures how strongly shocks to $PM_{2.5}$ levels at one station are propagated to other locations and hence how important that station is in terms of explaining the overall level of pollution across the city. The equivalent concept in the context of the banking system is what effect the loss made by one bank has on the performance of all other banks, which reflects its systemic importance.

Note that the analysis in this paper is quite different to that of Osses et al. (2013) and Henriquez et al. (2015). These papers use contemporaneous correlations between monitoring stations to determine the quality of the monitoring network in Santiago. By contrast, the analysis in this paper reveals the effect of shocks or perturbations to the pollution levels at each monitoring station on concentrations at other locations, taking the monitoring network as given. This work is similar in spirit to Liang, Zhou, Yang, Che, Wang, and Sun (2019) who consider large-scale spatial correlations in annual $PM_{2.5}$ levels across different parts of China, but the network approach used here is very different to the traditional spatial econometric models they employ. The biggest difference is that there is no need to impose a spatial weighting matrix, defined in terms of either adjacency or inverse distance. Instead, the VAR underlying the network captures the strength of the relationship between measurements at different locations directly from the observed data without the need for assigning arbitrary spatial weights.

The rest of the paper is structured as follows. Section 2 discusses the monitoring stations spread across Santiago along with some important properties of the observed $PM_{2.5}$ data. Section 3 outlines the network methodology used to estimate the systemic importance of locations relative to pollution across the entire city. Section 4 discusses the empirical results identifying the areas from which pollution has the biggest impact. Section 5 provides some concluding comments.

[2] For a general introduction to time-series modeling using VARs see Martin, Hurn, and Harris (2013), Chapters 13 and 14.

[3] See Acharya, Blackwell, and Sen (2016) for the potential to introduce bias by using potentially endogenous variables in this setting.

2 Data

The data used in this study are hourly historical observations of weather variables and environmental concentrations from 11 monitoring stations located in Santiago, Chile. The data were collected from the National Air Quality Information System (SINCA) for the period January 1, 2011 to August 31, 2015.

Fig. 1 shows the geographic distribution of the monitoring stations in the Santiago region. The values in parentheses report the annual average of $PM_{2.5}$ concentrations for the 2011 – 2015 period at each station. It is immediately apparent that the stations are not uniformly spaced. There are some stations such as Talagante and Las Condes that are separated by more than 50km and where there is likely to be a weak relationship between their concentration levels. In contrast, there is likely to be greater interaction between stations like Pudahuel and Cerro Navia, given their proximity.

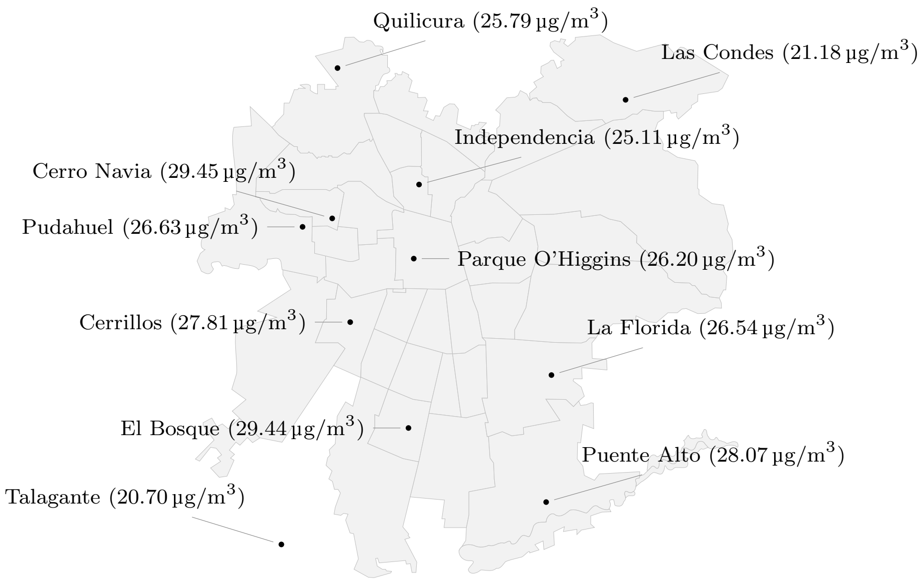

Fig. 1 Map of 11 monitoring stations in Santiago, Chile. Values in parentheses indicate hourly average $PM_{2.5}$ levels across period (2011 – 2014) for each commune corresponding to each monitoring station.

Table 1 reports average hourly $PM_{2.5}$ concentrations, together with firewood consumption and a number of poverty and population statistics for each commune surrounding each monitoring station. A number of interesting observations may be made by comparing the characteristics of each commune to its average $PM_{2.5}$ concentration. The communes with the greatest average $PM_{2.5}$ levels, namely El Bosque and Cerro Navia, are not the areas with the largest population. The two communes with the largest populations, Puente Alto and La Florida, experience relatively high pollution levels but not the highest. The most polluted communes, El Bosque and Cerro Navia, have both the highest population densities and the worst poverty. Although it is reasonable to expect that greater concentrations of

Station	Hourly Mean $PM_{2.5}$($\mu g/m^3$)	Consumption of firewood (ton/year)	Poverty due to income (%)	Multidimensional Poverty Index (%)	Population	Density (hab/km2)
Cerrillos	27.81	169	8.1	19.7	71.906	3424
Independencia	25.11	227	9.8	21.3	65.479	8849
Las Condes	21.18	6044	0.6	4.8	249.893	2514
El Bosque	29.44	1712	14.5	27.0	175.594	12453
Parque O'Higgins	26.20	1996	5.9	11.6	200.792	8964
Talagante	20.70	11892	12.0	29.9	59.805	477
Quilicura	25.79	962	7.8	18.5	126.518	2200
Pudahuel	26.63	2172	7.8	20.5	195.653	991
Cerro Navia	29.45	1568	12.1	35.6	148.312	13361
La Florida	26.54	3006	3.1	17.0	365.674	5165
Puente Alto	28.07	6848	8.0	27.1	492.915	5589

Table 1 Average $PM_{2.5}$ concentrations for each monitoring station. Firewood consumption, poverty and population statistics for each commune based around each station are also reported, see (Gajardo 2016), (Gramsch 2014) and (INE 2015) respectively.

relatively poor households will be associated with increased burning of firewood, which is one of the most important contributory factors leading to higher $PM_{2.5}$ levels (Molina, Toro, Morales, Manzano, and Leiva-Guzman 2017), this is in fact not the case. Apart from Puente Alto, where both firewood consumption and $PM_{2.5}$ levels are high, there is little association between firewood consumption and $PM_{2.5}$ levels.

In summary, there appear to be no obvious patterns between the characteristics of the local commune areas and the pollution levels they experience. The network analysis undertaken here will provide a formal analysis of the links between concentration levels at all 11 stations. The results will reveal whether locations with the highest average concentrations have the biggest impact on pollution levels across the entire city, or whether these areas merely suffer from pollution generated in other areas.

3 Measuring network connectedness

The methodology used to estimate the connectedness between $PM_{2.5}$ concentrations at the 11 stations is based on Diebold and Yilmaz (2014), who demonstrate how a traditional VAR model and associated variance decomposition is useful for measuring network connectedness. This framework provides estimates of the total directional connectedness from one individual region to all others, and the connectedness from all other regions to an individual region, and permits an analysis of how shocks in $PM_{2.5}$ concentrations originate and are transmitted across the greater Santiago metropolitan region. An important feature of the VAR framework used in this paper, known as a generalized variance decomposition, is that the ordering of the time series in the model is not important for establishing the impact of a shock to $PM_{2.5}$ concentration in one location on the concentrations in other locations over the chosen time horizon.[4] Note that this methodology introduces an explicit time dimension to the analysis and is not, therefore, simply accounting for contemporaneous correlations between concentrations at each monitoring location.

[4] For ease of interpretation, the monitoring stations are ordered in the VAR in south-west to north-east order according to their geographic location given that this is the direction of the prevailing wind across Santiago.

The network analysis is based on a VAR model given by

$$\mathbf{Y}_t = \beta_0 + \sum_{j=1}^{6} \beta_j \mathbf{Y}_{t-j} + \sum_{l=2}^{24} \gamma_j \mathbf{I}_j^{HR} + \sum_{k=2}^{12} \theta_k \mathbf{I}_k^{MTH} + \delta \mathbf{T}_t + \lambda \mathbf{WD}_t \mathbf{WV}_t + \varepsilon_t, \qquad \varepsilon_t \sim iid(\mathbf{0}, \mathbf{\Sigma}), \tag{1}$$

where $\mathbf{Y}_t$ is an (11×1) vector containing the hourly observations of $PM_{2.5}$ concentration for each of the stations and ε_t is a disturbance vector which is assumed to be independently and identically distributed. The coefficient matrices, β_j, will capture the strength of the links between the different locations through time without the need for a spatial weighting matrix as required by traditional spatial models. The estimates in β_j also capture the temporal dependence between locations and not merely the contemporaneous effects.

There are a number of additional features in this specification that need explanation. The diurnal pattern in $PM_{2.5}$ is dealt with by an indictor variable, $\mathbf{I}_l^{HR}$ which is a vector of ones when t falls within the l^{th} hour of the day, $l = 2, \ldots, 24$ and zeros otherwise. Similarly, the annual seasonal pattern is accounted for by the indicator $\mathbf{I}_k^{MTH}$ which is a vector of ones when t falls in month k, $k = 2, \ldots, 12$. Therefore, the period, $0:00$ to $1:00$ in January is the base case.

Previous studies have shown that temperature, which influences heating demand, and wind speed and wind direction play an important role in the prediction of $PM_{2.5}$ concentrations given their impact on the atmospheric and ventilation conditions in the Santiago river basin (Kurt and Oktay 2010; Feng, Li, Zhu, Hou, Jin, and Wang 2015; Saide et al. 2016; Moisan et al. 2018). Consequently, $\mathbf{T}_t$, the hourly temperature recorded at each station enters as an exogenous variable. In addition, the prevailing wind direction in Santiago is south-westerly. Given the importance of this direction for ventilation purposes, a dummy variable is included for when the wind is from the south-west as measured at each station, $\mathbf{WD}_t$, and this variable is interacted with wind velocity at each station, $\mathbf{WV}_t$. Temperature and the wind interaction terms are chosen because they are found to give the most effective forecasts, taking into account spatial effects between the measuring stations, (Moisan et al. 2018). If there are any relevant meteorological variables which are omitted, their effects will be accounted for by the sets of dummy variables incorporated in the regressions and the lagged $PM_{2.5}$ terms.

The estimates of connectedness between the $PM_{2.5}$ concentrations at different stations are generated from the shares of forecast error variation in the concentrations at one station that are due to shocks arising in concentrations at other stations. This approach to connectedness is related to the familiar econometric notion of variance decomposition in which the forecast error variance of variable i is decomposed into parts attributable to other variables in the system. Let the fraction of variable $i's$ H-step forecast error variance due to shocks in variable j be denoted d_{ij}^H. The quantity d_{ij}^H takes the form

$$d_{ij}^H = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' \mathbf{A_h} \mathbf{\Sigma} e_j)^2}{\sum_{h=0}^{H-1} (e_i' \mathbf{A_h} \mathbf{\Sigma} \mathbf{A_h}' e_i)}, \tag{2}$$

where e_j is a selection vector with j^{th} element unity and zeros elsewhere at time t, $\mathbf{A_h}$ is the coefficient matrix of the h-lagged shock vector in the infinite moving-average representation of the VAR model, $\mathbf{\Sigma}$ is the covariance matrix of the shock vector in the VAR, and σ_{jj} is the j^{th} diagonal element of $\mathbf{\Sigma}$. As shocks are correlated here, sums of forecast error variance contributions are not necessarily unity and therefore d_{ij}^H is normalized to $\tilde{d_{ij}^H}$ by dividing by $\sum_{j=1}^N d_{ij}^H$. Note that in general $d_{ij}^H \neq d_{ji}^H$, meaning there are $11^2 - 11$ separate pairwise directional connectedness measures.

First define total directional connectedness from other stations to station i as

$$C_{i\leftarrow\bullet} = \sum_{j=1, j\neq i}^{N} \tilde{d_{ij}^H}. \tag{3}$$

$C_{i\leftarrow\bullet}$ reflects how $PM_{2.5}$ shocks occurring at other stations influences the concentrations observed at the i^{th} station, identifying the degree of fragility (sensitivity) of concentrations at station i from levels across the region. Next, define total directional connectedness to other stations from station j as:

$$C_{\bullet\leftarrow j} = \sum_{i=1, i\neq j}^{N} \tilde{d_{ij}^H}. \tag{4}$$

$C_{\bullet\leftarrow j}$ reflects how shocks in the concentration at the j^{th} station will influence the concentration at all other stations, identifying the degree of centrality (systemic contribution) of one location to pollution levels across the whole greater Santiago area. A 6-hour ahead forecast error variance decomposition is used to construct these measures.

4 Results

Fig. 2 plots the relative degree of fragility, or sensitivity to shocks from all other locations, for each location (left axis – solid line) and the average level of $PM_{2.5}$ observed at each station (right axis – small dots) as a point of reference. While there is some degree of association between the level of fragility and average $PM_{2.5}$ concentrations, there are also a number of important differences to note. Las Condes and Talagante have the lowest sensitivity to changes in pollution levels from all other locations leading to the lowest average concentrations. This is not surprising as both locations are some distance from the city centre and Las Condes is located at a higher altitude than most of the city. Puente Alto is an interesting case in that it experiences relatively high average $PM_{2.5}$ concentrations but exhibits very low sensitivity to the rest of the city. Most importantly, Parque O'Higgins (the highest), Cerrillos and Quilicura exhibit the greatest fragility of all the other communes. Recall from Table 1 that these locations consume relatively small amounts of firewood, implying that their relatively high concentrations are due to pollution produced in other areas around the city, a result which is consistent with their relatively high degrees of fragility.

Similarly, Fig. 3 shows the estimates of centrality (left axis – solid line) and average concentration levels of $PM_{2.5}$ (right axis – small dots) at each station. This figure reveals the systemic importance of each location in explaining movements in $PM_{2.5}$ across the whole city. It may be deduced from Fig. 3 that El Bosque and Cerro Navia are the two most systemically important locations having the

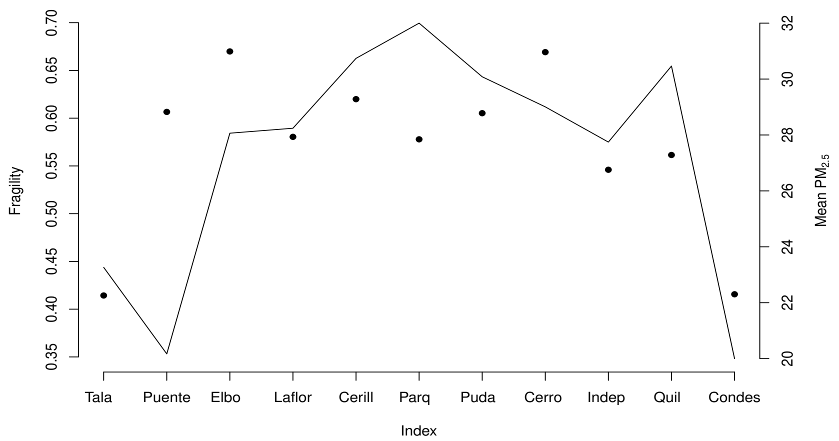

Fig. 2 Measures of fragility (left axis – solid line) for each monitoring station and unconditional average $PM_{2.5}$ concentrations (right axis – small dots) measured in standard units of $\mu g/m^3$.

greatest impact on $PM_{2.5}$ across Santiago. These locations experience the highest average concentrations while at the same time do not exhibit relatively high degrees of fragility, leading to the conclusion that it is locally produced $PM_{2.5}$ that is responsible for the high concentrations at these locations. Consideration of the summary statistics in Table 1 reveals that these high levels of pollution are not due to the burning of firewood, but perhaps are related to high rates of traffic flow in these communes, a function of their very high population densities. Similarly, Puente Alto exhibits a relatively high degree of centrality and hence a significant impact on the rest of the city, but its high concentration levels are solely the result of locally generated $PM_{2.5}$ given that its degree of fragility is very low. This observation is consistent with its location to the south of central Santiago as pollution will move in a northerly direction on the prevailing wind. At the opposite end of the spectrum are Las Condes and Talagante which are the least centrally important and have the smallest impact on the rest of the city.

These results allow the following conclusions to be drawn about the spatial behaviour of particulate matter in Santiago. The two locations with the highest average $PM_{2.5}$ concentrations, namely El Bosque and Cerro Navia, exhibit the highest centrality. Changes in concentration levels at these locations have the largest, and widest impact on pollution levels across the city. Importantly however, from a public policy perspective, is that Parque O'Higgins and Cerrillos (and to lesser degree, Independencia and Quilicura) have quite low degrees of centrality. Also recall the earlier result that the degrees of fragility at these stations are the highest across the city. The combination of low centrality, high fragility and low local firewood consumption means that the high average concentration levels experienced at these locations are not locally produced and are the result of $PM_{2.5}$ generated at other locations.

To give an easily interpretable summary of the differences in centrality, Fig. 4 shows generalized impulse response functions (GIRFs) which illustrate the effect of

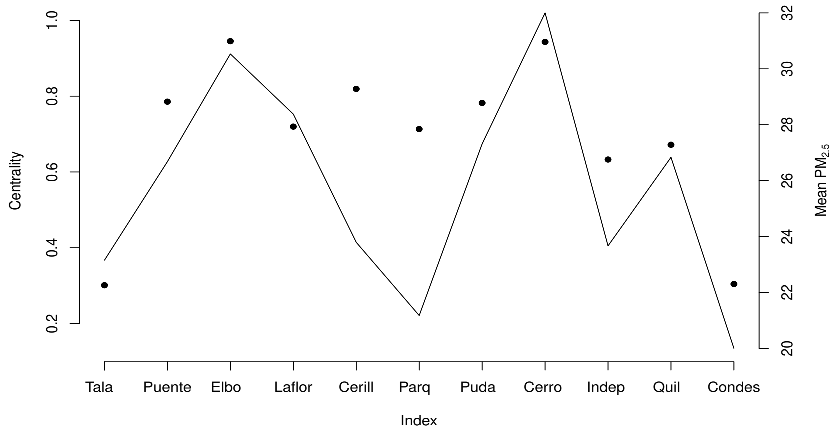

Fig. 3 Measures of centrality (solid line left axis) for each monitoring station and unconditional average $PM_{2.5}$ concentrations (dots from right axis) measured in standard units of $\mu g/m^3$.

shocks at four selected stations on all the other locations. The GIRFs are based on the VAR coefficients from Eq. (1) and are computed for shocks to $PM_{2.5}$ at Parque O' Higgins and Independencia, representing communes with low centrality, and El Bosque and Cerro Navia, representing communes with high centrality. The impact of differences in centrality are immediately evident in Fig. 4. In comparison to Parque O' Higgins and Independencia, shocks to concentrations at El Bosque and Cerro Navia have much larger impacts across most locations for up to 6-9 hours ahead. Well beyond just experiencing higher average concentrations, changes in concentrations at El Bosque and Cerro Navia have proportionally a much bigger impact on all other locations in greater Santiago for many hours into the future.

These results provide new insights in the spatial importance of the different locations across Santiago. Osses et al. (2013) show that from a measurement perspective, the information gain is greatest from downtown locations such as Parque O'Higgins, Independencia and El Bosque meaning that they provide the most accurate information regarding pollution levels across the entire city. However from a public policy perspective, the current results show that there are significant differences between the importance of these locations. Changes in pollution levels at the most centrally important locations, El Bosque and Cerro Navia have the greatest impact on the spatio-temporal evolution of pollution across many parts of the city. It is at these locations where official efforts should be focused to help reduce $PM_{2.5}$ pollution.

5 Conclusion

In the Chilean capital of Santiago air pollution is a major issue. This paper examines how pollution levels, measured in terms of the concentration of $PM_{2.5}$, at different locations across the city interact and seeks to identify the spatio-temporal

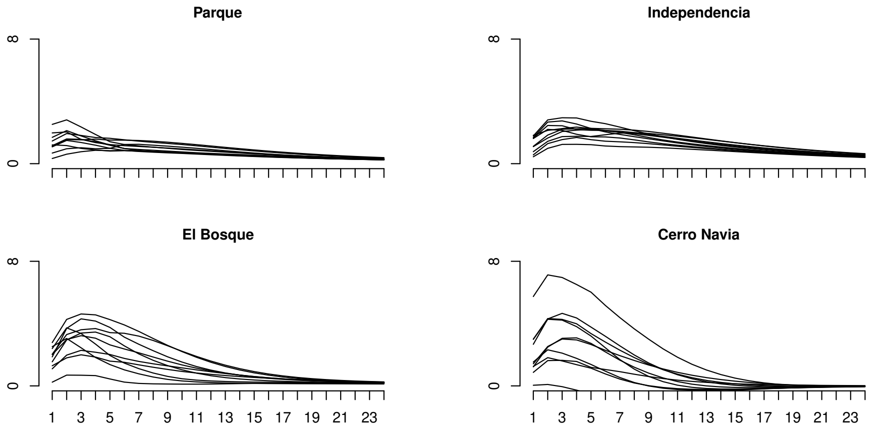

Fig. 4 Generalized impulse response functions given shocks to $PM_{2.5}$ at the four named stations 24 hours ahead. Each subplot represents the change in $PM_{2.5}$ concentrations (measured in $\mu g/m^3$) at all other stations resulting from a shock to $PM_{2.5}$ concentration at the station named in each subplot.

links between pollution levels at the various monitoring locations. These links are then used to construct the sensitivity of each location to pollution generated elsewhere, and also the systemic importance of each location to the pollution levels of other locations.

While many of the monitoring stations in central Santiago experience relatively high average concentration levels, the driving forces behind these readings differ. A number of important stations in the centre of the city experience high concentration levels that are generated primarily by heavy traffic (and not burning firewood) and these locations turn out to be crucial in transmitting pollution to other parts of the city. On the other hand, there are a number of other heavily polluted areas (with lower populations and firewood consumption) where pollution levels are mainly driven by their sensitivity to neighbouring locations.

Identifying locations of systemic importance provides a powerful tool for policymakers. Pinpointing the most systemically important areas of the city may open up further options for a more targeted approach to controlling $PM_{2.5}$ levels. These options should involve a closer examination of the factors leading to the production of $PM_{2.5}$ at these locations. The results reported here indicate that it is not only the consumption of firewood which is responsible for the high levels of particulate matter pollution in areas with high centrality but also possibly traffic and industrial activity.

References

Acharya A, Blackwell M, Sen M (2016) Explaining causal findings without bias: Detecting and assessing direct effects. The American Political Science Review 110(3):512–529, http://dx.doi.org/10.1017/S0003055416000216

Cifuentes L (2010) Relación de la Norma de Calidad Primaria PM2.5 con la Norma de Calidad Primaria de PM10. Tech. rep., Comision Nacional del Medio Ambiente, Region Metropolitana de Santiago

Díaz-Robles L, Ortega J, Fu J, Reed G, Chow J, Watson J, Moncada J (2008) A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of temuco, Chile. Atmospheric Environment 42(35):8331–8340, https://doi.org/10.1016/j.atmosenv.2008.07.020

Diebold F, Yilmaz K (2014) On the network topology of variance decompositions: Measuring the connectedness of financial firms. Journal of Econometrics 182:119–134, https://doi.org/10.1016/j.jeconom.2014.04.012

Feng X, Li Q, Zhu Y, Hou J, Jin L, Wang J (2015) Artificial neural networks forecasting of PM 2.5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmospheric Environment 107:118–128, https://doi.org/10.1016/j.atmosenv.2015.02.030

Gajardo S (2016) Pobreza y distribución del ingreso en la región metropolitana de santiago: Resultados encuesta casen 2015. Seremi de Desarrollo Social Metropolitana

Gramsch E (2014) Actualización y sistematización del inventario de emisiones de contaminantes atmosféricos en la región metropolitana. Ministerio del Medio Ambiente

Henriquez A, Osses A, Gallardo L, Diaz Resquin M (2015) Analysis and optimal design of air quality monitoring networks using a variational approach. Tellus B: Chemical and Physical Meteorology 67(1):25385 `https://doi.org/10.3402/tellusb.v67.25385`

INE (2015) Census 2015. National Statistics Institute (INE) Chile

Kurt A, Oktay A (2010) Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. Expert Systems with Applications 37(12):7986–7992, https://doi.org/10.1016/j.eswa.2010.05.093

Liang Z, Zhou C, Yang F, Che L, Wang B, Sun D (2019) Spatio-temporal evolution and the influencing factors of $PM_{2.5}$ in China between 2000 and 2015. Journal of Geographical Sciences (2):253–270, https://doi.org/10.1007/s11442-019-1595-0

Maciejewska K (2020) Short-term impact of $PM_{2.5}$, PM_{10}, and PM_c on mortality and morbidity in the agglomeration of Warsaw, Poland. Air Quality, Atmosphere & Health https://doi.org/10.1007/s11869-020-00831-9

Martin V, Hurn A, Harris D (2013) Econometric Modelling with Time Series. Cambridge University Press, New York

MMA (2011) Norma primaria de calidad ambiental para material particulado fino respirable PM 2.5. Reporte Oficial. Tech. rep., Ministerio del Medio Ambiente

Moisan S, Herrera R, Clements A (2018) A dynamic multiple equation approach for forecasting $PM_{2.5}$ pollution in Santiago, Chile. International Journal of Forecasting 34:566–581, https://doi.org/10.1016/j.ijforecast.2018.03.007

Molina C, Toro A, Morales R, Manzano C, Leiva-Guzman M (2017) Particulate matter in urban areas of south-central Chile exceeds air quality standards. Air Quality, Atmosphere & Health 10(5), https://doi.org/10.1007/s11869-017-0459-y

Nicolis O, Díaz M, Sahu S, Marín J (2019) Bayesian spatio-temporal modeling for estimating short-term exposure to air pollution in Santiago de Chile. Environmetrics 30(7):e2574, https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2574

Osses A, Gallardo L, Faundez T (2013) Analysis and evolution of air quality monitoring networks using combined statistical information indexes. Tellus B: Chemical and Physical Meteorology 65(1):19822, 10.3402/tellusb.v65i0.19822

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

Perez P, Reyes J (2006) An integrated neural network model for PM10 forecasting. Atmospheric Environment 40(16):2845–2851, https://doi.org/10.1016/j.atmosenv.2006.01.010

Pino-Cortés E, Díaz-Robles L, Campos V, Vallejo F, Gómez J, Cereceda-Balic F, Fu J, Figueroa J (2020) Effect of socioeconomic status on the relationship between short-term exposure to $PM_{2.5}$ and cardiorespiratory mortality and morbidity in a megacity: the case of Santiago de Chile. Air Quality, Atmosphere & Health 13(5), https://doi.org/10.1007/s11869-020-00818-6

Sahu S, Gelfand A, Holland D (2006) Spatio-temporal modeling of fine particulate matter. Journal of Agricultural, Biological, and Environmental Statistics 11:61–86, https://doi.org/10.1198/108571106X95746

Saide P, Carmichael G, Spak S, Gallardo L, Osses A, Mena M, Pagowski M (2011) Forecasting urban PM 10 and PM 2.5 pollution episodes in very stable nocturnal conditions and complex terrain using wrf–chem co tracer model. Atmospheric Environment 45(16):2769–2780, https://doi.org/10.1016/j.atmosenv.2011.02.001

Saide P, Mena M, Tolvett S, Hernandez P, Carmichael G (2016) Air quality forecasting for winter-time PM 2.5 episodes occurring in multiple cities in central and southern Chile. Journal of Geophysical Research: Atmospheres 121(1):558–575, https://doi.org/10.1002/2015JD023949

SINIA (2010) Análisis del impacto económico y social de la norma de calidad primaria de material particulado 2.5. Tech. rep., Sistema Nacional de Información Ambiental