

Analysis and optimal design of air quality monitoring networks using a variational approach

By ADOLFO HENRIQUEZ^{1,2*}, AXEL OSSES^{1,2,3}, LAURA GALLARDO^{1,4} and MELISA DIAZ RESQUIN^{1,5,6}, ¹Center for Climate and Resilience Research (CR2), Blanco Encalada 2002, Santiago, Chile; ²Centro de Modelamiento Matemático, Universidad de Chile-CNRS, Blanco Encalada 2120, piso 7, Santiago, Chile; ³Departamento de Ingeniería Matemática, Universidad de Chile, Blanco Encalada 2120, piso 5, Santiago, Chile; ⁴Departamento de Geofísica, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile; ⁵Comisión Nacional de Energía Atómica, Gerencia Química, Av. Gral. Paz 1499 (B1650KNA), San Martín, Pcia. de Buenos Aires, Argentina; ⁶Departamento de Ingeniería Química, Facultad de Ingeniería, Universidad de Buenos Aires, Buenos Aires, Argentina

(Manuscript received 4 July 2015; in final form 11 August 2015)

ABSTRACT

Air quality networks need revision and optimisation as instruments and network requirements, both scientific and societal, evolve over time. Assessing and optimising the information content of a monitoring network is a non-trivial problem. Here, we introduce a methodology formulated in a variational framework using an air quality model to simulate the dispersion of carbon monoxide (CO) as a passive tracer at the city scale. We address the specific case of adding or removing stations, and the more general situation of optimally distributing a given number of stations in a domain taking into account transport patterns and spatial factors such as population density and emission patterns. We consider three quality indicators: precision gain, information gain and degrees of freedom for a signal. These metrics are all functions of the singular values of the sensitivity matrix that links emissions and observations in the variational framework. We illustrate the application of the methodology in the case of Santiago (33.5°S, 70.5°W, 500 m a.s.l.), a city of ca. 7 million inhabitants with significant pollution levels. We deem information gain as the best of the above indicators for this case. We then quantify the actual evolution of Santiago's network and compare it with the optimal configuration suggested by our methodology and with results previously obtained using a statistical approach. The application is restricted to diurnal and summer conditions, for which the dispersion model shows a good agreement with observations. The current method offers advantages in that it allows extending a network to include new sites, and it explicitly considers the effects of dispersion patterns, and desired weighting functions such as emission fluxes and population density. We find that Santiago's air quality has improved two-fold since 1988, regarding CO under diurnal summer conditions. Still, according to our results, the current configuration could be improved by integrating more suburban stations in the southwest of the basin.

Keywords: optimal network design, air pollution, variational framework

1. Introduction

Observing systems, in general, and air quality networks, in particular, need revision and optimisation as monitoring instruments, scientific and societal requirements change over time. The objectives of an air quality monitoring network are multiple, including compliance of air quality standards, evolution of air quality and efficiency of curbing

measures, as well as quantifying impacts on human health, ecosystems and climate. The question of how to best sample airborne pollutants in a monitoring network is non-trivial. The difficulties arise from the questions of defining location, sampling frequency and choice of variable to measure, and so on, in an optimal sense according to different criteria, and the large number of potential sites and times to collect data in a given system. In general, optimal network design is framed as a complex optimisation or decision problem (Caselton and Zidek, 1984; Zidek and Zimmerman, 2010). Currently, urban sustainability in the framework of a changing climate, and the need for combining multiple

*Corresponding author.
email: ahenriquez@dim.uchile.cl

observing platforms, poses new challenges to the optimal design of networks (Lahoz and Schneider, 2014). There is an increasing body of literature addressing optimal network design in connection with air quality problems. Several approaches have been suggested over the last decades. There are those using pure measurements and statistics, and those using pure numerical simulations, as well as combinations of both. Of course, many of these methodologies allow network reduction or extension and weighted criteria under suitable social, economic or technical restrictions (e.g. Langstaff et al., 1967; Husain and Khan, 1983; Shindo et al., 1990; Trujillo–Ventura and Ellis, 1991; Nychka and Saltzman, 1998; Tseng and Chang, 2001; Bocquet, 2005a, 2005b; Kanaroglou et al., 2005; Kao and Hsieh, 2006; Abida et al., 2008; Elkamel et al., 2008; Abida and Bocquet, 2009; Ainslie et al., 2009; Mofarrah and Husain, 2010; Wu et al., 2010; Romary et al., 2011; Saunier et al., 2011; Sturman et al., 2011; Wu and Bocquet, 2011; Ruiz-Cárdenas et al., 2012; Osses et al., 2013; Singh et al., 2013; Wang et al., 2013; Cioaca and Sandu, 2014; Gómez-Losada et al., 2014; Romary et al., 2014; Zidek et al., 2000, 2014).

In this work, we introduce a methodology formulated in a variational framework, that is, based on the optimisation of quality indicators as functions of the analysis covariance matrix in the context of data assimilation. Therefore, the methodology also incorporates transport and mixing in a city by means of the sensitivity matrix calculated using a chemistry transport model (CTM). In this way, we quantify the changes in quality of a monitoring network when adding or removing stations, and when searching for the optimal spatial distribution of a given number of stations in a domain. We use three quality indicators: precision gain, information gain and degrees of freedom for the signal (DFS) (Shannon, 1948; Rodgers, 2000; Fisher, 2003). These metrics are used in optimal network design, but they are seldom put in a common framework. In fact, we show these metrics in terms of singular values of the sensitivity matrix that links emissions and observations in the variational or data assimilation framework. This methodology seems to be new in the context of network design for atmospheric dispersion. Indeed, even if these indicators (degrees of freedom, information content, total precision) are widely used in meteorological data assimilation, they are not extensively used in atmospheric chemistry data assimilation or atmospheric network design. The methodology also allows to weight in emission fields, population density or other spatially varying characteristics in a straightforward manner.

In sum, the present study includes the intercomparison of several network quality indicators, the use of geographical weights to obtain realistic and feasible networks, and a clear methodology to test the robustness analysis of the

proposed optimisation procedure with respect to the initial network.

Furthermore, we apply the proposed methodology at a city scale to the case of Santiago in Chile (33.5°S , 70.5°W , 500 m a.s.l.). The air quality monitoring network of Santiago de Chile, a city of ca. 6 million inhabitants with significant levels of pollution, has evolved as stations have been removed and relocated, and the spatial coverage has grown to include present-day suburban stations, as shown in Fig. 1. These changes have been motivated by the need to monitor the attainment of air quality standards in a city that has rapidly expanded (Romero and Ordenes, 2004), and changed its emission rates and patterns (Gallardo

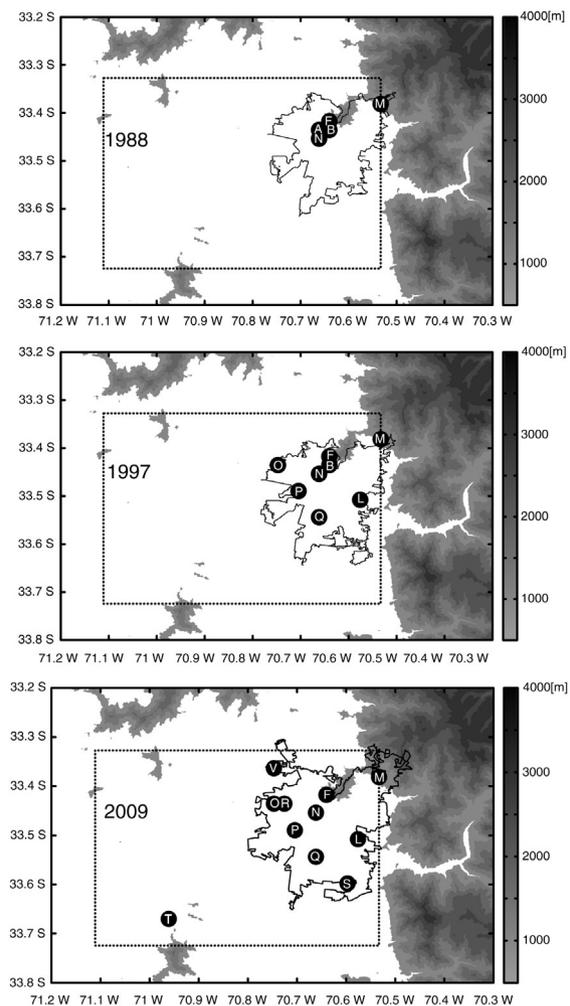


Fig. 1. Santiago’s air monitoring sites (black dots) and stations symbols (white letters) for the years 1988, 1997 and 2009. The main topographic features of the Santiago basin are also shown (shading). The approximate urban–rural limit for the years 1989, 1996 and 2010 is shown in a black line (www.ocuc.cl). The area in which stations can be placed in this study is indicated by a dashed line.

et al., 2012), similar to many other cities in the world (Zhu et al., 2012). For this application, we use a modelling realisation described by Saide et al. (2009) and Saide et al. (2011a). These simulations describe the dispersion of carbon monoxide (CO) over Santiago de Chile under diurnal conditions in a summer month. We further address the evolution of the air quality network in terms of its information gain by taking into account both the number and the location of CO stations. We discuss the pros and cons of the variational approach introduced here and compare it with the results of a statistical approach previously presented by Osses et al. (2013).

In the next section, we introduce various metrics to quantify changes in the information content of a network including precision gain (p_g), information gain (i_g) and DFS (d_s). Section 3 covers how these metrics are applied to address the addition and removal of stations, and to the question of optimal design of a monitoring network, including the use of geographical weighting functions such as population density and emission patterns. In Section 4, we apply this methodology to the case of Santiago's air quality network and its evolution. Summary and conclusions are presented in Section 5.

2. Quantifying the quality of a network

In the context of optimal network design using a variational approach, one encounters the use of various metrics to assess the quality of an observing system. Here, we consider three indicators of relative changes in quality: precision gain, information gain and DFS (e.g. Rodgers, 2000). Along these lines, one can evaluate a set of stations in an air quality network to improve the precision of the estimate of an unknown parameter, for example, emissions of a given tracer, here CO. In this particular case, by assuming a linear relationship between emissions and observations of CO, one links emissions \mathbf{x} with measured observations \mathbf{y} at the sites of a given network, using the sensitivity matrix \mathbf{H} (with rank $r > 0$ and an associate error ε

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \varepsilon. \quad (1)$$

This is a reasonable assumption for CO at the city scale since it behaves as a transport tracer (e.g. Saide et al., 2011b). Note that the sensitivity matrix \mathbf{H} is calculated for unitary emissions and evaluated at each observation site. Thus, \mathbf{H} depends on mixing and transport, and the location and timing of the observation network.

Before specifying the selected quality metrics, let us introduce the basic definitions and notations of the data assimilation framework (e.g. Daley, 1993; Bouttier and Courtier, 1997; Kalnay, 2003). Let \mathbf{x}_t be the true n -dimensional state and \mathbf{x}_b its background estimate with error covariance \mathbf{B} , that is, $\mathbf{x}_b \sim \mathcal{N}(\mathbf{x}_t, \mathbf{B})$. Let \mathbf{y}_o be the m -dimensional

measurement observation vector with error covariance \mathbf{R} , that is, $\mathbf{y}_o \sim \mathcal{N}(\mathbf{H}\mathbf{x}_t, \mathbf{R})$. Let us assume that \mathbf{B} and \mathbf{R} are positive definite matrices. The analysis (estimate) of the state \mathbf{x}_a is the unique solution of the optimisation problem

$$\min_{\mathbf{x}} J(\mathbf{x}) := \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}_o\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\mathbf{x} - \mathbf{x}_b\|_{\mathbf{B}^{-1}}^2, \quad (2)$$

where $\|\mathbf{x}\|_{\mathbf{M}} = \mathbf{x}^T \mathbf{M} \mathbf{x}$. The analysis is given by

$$\mathbf{x}_a = \mathbf{x}_b + \Sigma_a \mathbf{H}^T \mathbf{R}^{-1} \mathbf{d} \quad (3)$$

$$\Sigma_a = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}, \quad (4)$$

where Σ_a is the error covariance matrix of the analysis [i.e. $\mathbf{x}_a \sim \mathcal{N}(\mathbf{x}_t, \Sigma_a)$] and $\mathbf{d} = \mathbf{y}_o - \mathbf{H}\mathbf{x}_b$ is the so-called innovation vector. Note that $\Sigma_a^{-1} - \mathbf{B}^{-1} = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \geq 0$, indicating that the assimilation of observations always improves the *precision* of analysis Σ_a^{-1} with respect to the *precision* of the background \mathbf{B}^{-1} .

Let us consider the transformation $\tilde{\mathbf{H}} = \mathbf{R}^{-\frac{1}{2}} \mathbf{H} \mathbf{B}^{-\frac{1}{2}}$, where $\mathbf{B}^{\frac{1}{2}}$ and $\mathbf{R}^{\frac{1}{2}}$ stand for the square roots from the Cholesky decomposition (e.g. Horn and Johnson, 2012) and the singular value decomposition $\tilde{\mathbf{H}} = \tilde{\mathbf{U}} \tilde{\Lambda} \tilde{\mathbf{V}}^T$, with $\tilde{\Lambda}$ the diagonal matrix of non-zero singular values $\tilde{\lambda}_1, \dots, \tilde{\lambda}_r$, and $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ unitary matrices of singular vectors $\tilde{\mathbf{u}}_i$ and $\tilde{\mathbf{v}}_i$, respectively. Also, transforming $\tilde{\mathbf{d}} = \mathbf{R}^{-\frac{1}{2}} \mathbf{d}$ and $\tilde{\mathbf{V}} = \mathbf{B}^{\frac{1}{2}} \tilde{\mathbf{V}}$, eqs. (3) and (4) become

$$\begin{aligned} \mathbf{x}_a &= \mathbf{x}_b + \tilde{\mathbf{V}} (\mathbf{I}_n + \tilde{\Lambda}^2)^{-1} \tilde{\Lambda} \tilde{\mathbf{U}}^T \tilde{\mathbf{d}} \\ &= \mathbf{x}_b + \sum_{i=1}^r \left(\frac{\tilde{\lambda}_i^2}{1 + \tilde{\lambda}_i^2} \right) \frac{\tilde{\mathbf{u}}_i^T \tilde{\mathbf{d}}}{\tilde{\lambda}_i} \tilde{\mathbf{v}}_i \end{aligned} \quad (5)$$

$$\Sigma_a = \mathbf{B}^{\frac{1}{2}} (\mathbf{I}_n + \tilde{\mathbf{H}}^T \tilde{\mathbf{H}})^{-1} \mathbf{B}^{\frac{1}{2}T}. \quad (6)$$

The term in parenthesis is known as the Tychonoff filter factor (Johnson, 2003). This factor ponders the contribution of the i th component of the singular value decomposition in the total analysis. When the corresponding $\tilde{\lambda}_i \gg 1$, this contribution is relevant, whereas when $\tilde{\lambda}_i \ll 1$ the opposite occurs.

In the particular case when the covariances are scalar matrices $\mathbf{B} = \sigma_b^2 \mathbf{I}_n$ and $\mathbf{R} = \sigma_o^2 \mathbf{I}_m$, with ratio $\mu^2 = \sigma_o^2 / \sigma_b^2$, this expression simplifies to

$$\mathbf{x}_a = \mathbf{x}_b + \sum_{i=1}^r \left(\frac{\frac{\lambda_i^2}{\mu^2}}{1 + \frac{\lambda_i^2}{\mu^2}} \right) \frac{\mathbf{u}_i^T \mathbf{d}}{\lambda_i} \mathbf{v}_i \quad (7)$$

$$\Sigma_a = \sigma_b^2 \mathbf{V} \left(\mathbf{I} + \frac{\Lambda^2}{\mu^2} \right)^{-1} \mathbf{V}^T, \quad (8)$$

using in this case the singular value decomposition $\mathbf{H} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T$, with $\mathbf{\Lambda}$ a diagonal matrix of non-zero singular values $\lambda_1, \dots, \lambda_r$, and \mathbf{U} and \mathbf{V} unitary matrices of singular vectors \mathbf{u}_i and \mathbf{v}_i , respectively.

Following from this, let \mathbf{Z} be the set of locations of the stations of a city's air quality network

$$\mathbf{Z} = \{z_1, z_2, \dots, z_N\} \quad (9)$$

with $z_i = (x_i, y_i)$ and cardinality $|\mathbf{Z}| = N$. To each set \mathbf{Z} , we associate a network of observation sites, and a corresponding sensitivity matrix \mathbf{H} and covariance matrix $\Sigma_{\mathbf{a}}$ with quality indicator given by $Q(\Sigma_{\mathbf{a}})$

$$\mathbf{Z} \rightarrow \mathbf{H} \rightarrow \Sigma_{\mathbf{a}} \rightarrow Q(\Sigma_{\mathbf{a}}).$$

In order to simplify the notation, we will associate to each network \mathbf{Z} the corresponding quality indicator $Q(\mathbf{Z})$.

$$\mathbf{Z} \rightarrow Q(\mathbf{Z}).$$

2.1. Precision gain

We introduce total precision of a given network \mathbf{Z} characterised by sensitivity matrix \mathbf{H} , as the trace of the precision matrix of the corresponding analysis, that is

$$p(\mathbf{Z}) = \text{Tr}(\Sigma_{\mathbf{a}}^{-1}) = \text{Tr}(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}), \quad (10)$$

which can be simplified in the case of diagonal error covariance matrices \mathbf{B} and \mathbf{R} as follows:

$$p(\mathbf{Z}) = \frac{n}{\sigma_b^2} + \frac{1}{\sigma_b^2} \sum_{i=1}^r \frac{\lambda_i^2}{\mu^2}. \quad (11)$$

Then, *precision gain* is obtained by subtracting the total precision after assimilation of the network observations from the total precision before assimilation:

$$p_g(\mathbf{Z}) = \text{Tr}(\Sigma_{\mathbf{a}}^{-1}) - \text{Tr}(\mathbf{B}^{-1}) = \frac{1}{\sigma_b^2} \sum_{i=1}^r \frac{\lambda_i^2}{\mu^2}. \quad (12)$$

In simple words, this indicator follows the notion that precision increases as errors decrease in a quadratic manner. It is worth noting that this metric is not sensitive to \mathbf{B} .

2.2. Information gain

The *total information* of a random variable \mathbf{x} with values in \mathbb{R}^n normally distributed with covariance $\Sigma_{\mathbf{x}}$ is given by

$$i(\mathbf{x}) = \frac{1}{2} \ln |\Sigma_{\mathbf{x}}^{-1}| - \frac{n}{2} (1 + \ln(2\pi)). \quad (13)$$

Therefore, *information gain* for a network \mathbf{Z} characterised by sensitivity \mathbf{H} is obtained by subtracting the total

information after and before the observations of the network are assimilated:

$$\begin{aligned} i_g(\mathbf{Z}) &= i(\mathbf{x}_{\mathbf{a}}) - i(\mathbf{x}_{\mathbf{b}}) \\ &= \frac{1}{2} \ln |\Sigma_{\mathbf{a}}^{-1}| - \frac{1}{2} \ln |\mathbf{B}^{-1}| \\ &= \frac{1}{2} \sum_{i=1}^r \ln \left(1 + \frac{\lambda_i^2}{\mu^2} \right). \end{aligned} \quad (14)$$

An alternative but equivalent definition uses the Kullback-Liebler distance (Kullback, 1959) as described in Osses et al. (2013). Since μ ponders the relative importance of the errors in observations and emissions, the extreme cases correspond to no knowledge of observations relative to emissions, that is, $\mu \rightarrow +\infty$ so $i_g = 0$, and perfect knowledge of observations relative to emissions, that is, $\mu \rightarrow 0$ so $i_g = +\infty$.

2.3. Degrees of freedom for the signal

The DFS associated with a network \mathbf{Z} characterised by the sensibility matrix \mathbf{H} are given by

$$d_s(\mathbf{Z}) = r - \text{Tr}(\mathbf{B}^{-1} \Sigma_{\mathbf{a}}) = r - \sum_{i=1}^r \left(1 + \frac{\lambda_i^2}{\mu^2} \right)^{-1}, \quad (15)$$

where r is the rank of \mathbf{H} (note that $r \leq n \ll m$ in the application discussed later). DFS represent the number of useful pieces of information present in the noisy measurements of a given network in order to effectively retrieve emissions. It can also be interpreted as the effective rank (but not integer in general) of the sensitivity matrix (see Rodgers, 2000). Extreme cases correspond to no knowledge of observation relative to emissions, that is, $\mu \rightarrow +\infty$ so $d_s = 0$, and perfect knowledge of observations relative to emissions, that is, $\mu \rightarrow 0$ so $d_s = r$.

2.4. Comparison among indicators

All previous quality indicators are functions of the analysis covariance matrix $\Sigma_{\mathbf{a}}$ associated with network \mathbf{Z} , thus non-linear functions of the associated sensitivity matrix \mathbf{H} . In the particular case of scalar covariance matrices for observations (\mathbf{R}) and emissions (\mathbf{B}), these terms can be expressed as the ratio λ_i/μ , where λ_i are the singular values of the sensitivity matrix, and μ is the ratio between observational and background deviations, in the general form $\sum_i f(\lambda_i/\mu)$, where the corresponding values for f are indicated in Table 1. A similar analysis can be done in the non-diagonal case by using expressions (5) and (6). Figure 2 compares the behaviour of f for the quality indicators: p_g , i_g and d_s .

Table 1. Summary of the main quality indicators for measuring network quality

Quality indicator	Q	Definition	$f\left(\frac{\lambda}{\mu}\right)$
Precision gain	p_g	$\text{Tr}(\Sigma_a^{-1}) - \text{Tr}(\mathbf{B}^{-1})$	$\frac{1}{\sigma_a^2} \frac{\lambda^2}{\mu^2}$
Information gain	i_g	$\frac{1}{2} \ln \Sigma_a^{-1} - \frac{1}{2} \ln \mathbf{B}^{-1} $	$\frac{1}{2} \ln\left(1 + \frac{\lambda^2}{\mu^2}\right)$
Degrees of freedom	d_s	$n - \text{Tr}(\mathbf{B}^{-1}\Sigma_a)$	$1 - \left(1 + \frac{\lambda^2}{\mu^2}\right)^{-1}$

See also Fig. 2 for illustration.

Calculating the singular values for the full sensitivity matrix can be cumbersome. For example, in the case of Santiago shown later, the observation operator for the whole domain has 154 560 rows and 9660 columns. Thus, to assert the dependence on singular values, we reviewed a random sample of 250 networks of 11 stations, and for which singular values were calculated. These calculations indicate that, on average, $\bar{\lambda}_{min} \sim 10^{-5}$ and $\bar{\lambda}_{max} \sim 10$. Let us now consider the singular values in the regions $\Lambda_1 = [\bar{\lambda}_{min}, \frac{\mu}{10}]$ and $\Lambda_2 = [\frac{\mu}{10}, \bar{\lambda}_{max}]$. In region Λ_1 , all quality metrics expressed as functions of singular values behave similarly, that is, $\sim \frac{\lambda^2}{\mu^2}$, and show negligible differences if compared with respect to their values in region Λ_2 . In region Λ_2 , the functions behave very differently. In this region, and for the random set of 250 networks, precision gain has the steepest slope. Furthermore, an average of 70% of the spectrum is in region Λ_1 . The remaining 30% of the spectrum is highly determined by singular values in region Λ_2 , since $\bar{f}(\Lambda_1) \sim 10^{-3}$ and $\bar{f}(\Lambda_2) \sim 1$. In order to compare the different quality indicators in the range of all scales of singular value, a logarithmic scale in singular values is needed. Precision gain varies quadratically with singular values in region Λ_2 , whereas DFS show an asymptotic growth, covering the full range of singular values. Information gain shows a linear

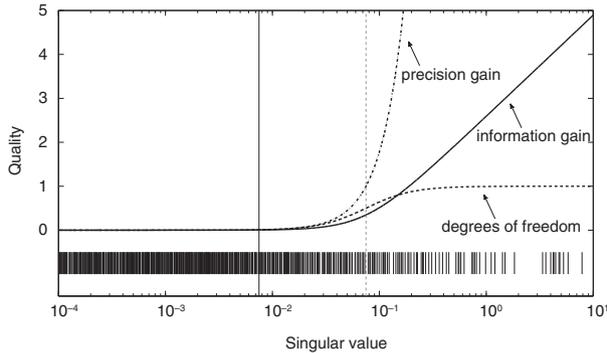


Fig. 2. Comparison of quality indicators as a function of singular values λ . The vertical continuous black line indicates the value of the parameter $\frac{\mu}{10} = 0.0075$, and the vertical dashed grey line indicates the value of the parameter $\mu = 0.075$. The bottom of the figure shows the distribution of the 644 singular values of the sensitivity matrix for Santiago's network in 2009.

growth, in between precision gain and DFS. For these reasons, we adopt information gain and DFS as quality indicators to be applied hereafter in the case of Santiago. Thus, for the simulated meteorological and dispersion conditions used for Santiago, the most robust indicators with respect to the leading singular values correspond to information gain and DFS, even if DFS is less sensitive to changes in singular values than information gain.

3. Methodology

In the previous section, we defined a number of metrics to address information content of a monitoring network. Each quality indicator $Q(\mathbf{Z})$ is associated, via the covariance matrix of the analysis, with some particular network $\mathbf{Z} = \{z_1, z_2, \dots, z_N\}$ characterised by the geographical location of each station $z_i = (x_i, y_i)$.

We first present the methodology for removing or adding one station, taking into account the influence of transport fields as well as the original geographical distribution of the network. Thereafter, we consider the problem of optimally distributing in space a given number of stations, taking into account circulation patterns, and weighting functions such as population density and emission fluxes. The inclusion of weighting functions allows a straightforward manner to ponder the importance of different factors, which in turn results in targeted optimal networks.

3.1. Reduction, extension and optimal distribution

- *Reduction.* Consider a quality indicator Q . The reduction of a network \mathbf{Z} by subtracting a station located at $\mathbf{Z}_i = \{(x_i, y_i)\}$ in the discrete spatial domain will correspond to a new network $\mathbf{Z} \setminus \mathbf{Z}_i$. For each station that can be extracted from the network, we calculate the decrease in quality by

$$Q(\mathbf{Z}) - Q(\mathbf{Z} \setminus \mathbf{Z}_i). \quad (16)$$

Thus, the greater the decrease in quality for a particular station, the lower its priority to be removed.

- *Extension.* Consider a quality indicator Q . The extension of a network \mathbf{Z} by adding a new station located at $\mathbf{Z}_0 = \{(x_0, y_0)\}$ in the discrete spatial domain will correspond to the new network $\mathbf{Z} \cup \mathbf{Z}_0$. For each possible new station, we calculate the increase in quality by

$$Q(\mathbf{Z} \cup \mathbf{Z}_0) - Q(\mathbf{Z}). \quad (17)$$

Thus, the greater the increase in quality for a potential new station, the greater its priority to be added.

- *Optimal distribution.* Consider a quality indicator Q and a network size N . We find the optimal network by

maximising the network's quality over all networks \mathbf{Z} of size N in the discrete spatial domain D , according to

$$\max_{\substack{\mathbf{Z} \in D \\ |\mathbf{Z}| = N}} Q(\mathbf{Z}) \quad (18)$$

This optimisation problem has two major difficulties. First, the feasible set is discrete, and second the direct evaluation of all possible values of the objective function is computationally cumbersome. Therefore, we adopt a heuristic optimisation algorithm known as simulated annealing (SA) (e.g. Bertsimas and Tsitsiklis, 1993). This algorithm gives an estimate of the optimal configuration in a manageable computation time even with a personal computer. In the Santiago case, for a network of 11 stations, and a sensitivity matrix of size 2640×644 , corresponding to 11 observation sites over 10 d and 24 hours, and 28×23 emission grids. We simplified the emissions by averaging over time or by choosing a specific hour. We solved this problem in 10 hours using a system of two cores of 2.3 GHz.

3.2. Simulated annealing

The procedure of SA starts from an initial feasible network and, at each iteration, it chooses, according to a given probability, to move to another feasible network in the vicinity of the current network or stay in the current network. The mobility of the search algorithm depends on a global parameter T and it is called *temperature*. Initially SA uses large values of T , allowing a random search. Then, as the values of the objective function increase, the value of T decreases, restricting the search and ensuring convergence to a local optimum.

For a network quality indicator Q , a feasible network \mathbf{Z}_1 , a test feasible network \mathbf{Z}_2 , and $\delta = Q(\mathbf{Z}_2) - Q(\mathbf{Z}_1)$, we define the probabilities

$$\rho = \begin{cases} 1 & \text{if } \delta \geq 0 \\ e^{\frac{\delta}{T}} & \text{if } \delta < 0 \end{cases} \quad (19)$$

$$\rho_a \sim \mathcal{U}([0, 1]). \quad (20)$$

If $\delta \geq 0$, SA moves from \mathbf{Z}_1 to \mathbf{Z}_2 , increasing the value of Q . If $\delta < 0$ and $\rho > \rho_a$, SA moves from \mathbf{Z}_1 to \mathbf{Z}_2 , without increasing the value of Q , but allowing the function to increase its value in future iterations. If $\delta < 0$ and $\rho < \rho_a$, SA stays in \mathbf{Z}_1 and chooses another test network. As T decreases, the latter ensures the convergence of SA near the local optimum.

The max-min network, which is a solution of the problem

$$\max_{\substack{\mathbf{Z} \in D \\ |\mathbf{Z}| = N}} \min_{\substack{z_1, z_2 \in \mathbf{Z} \\ z_1 \neq z_2}} \|z_1 - z_2\| \quad (21)$$

is a network equally spaced along the observation domain. This is an adequate initial condition for SA.

3.3. Geographical weighting

When designing monitoring networks, in practice one has to consider different criteria and objectives, as well as the geographical distribution of information, such as emissions, population density, socioeconomic indexes, among others. To do so, we apply suitable geographical weights to the sensitivity matrix. These weights $\pi = (\pi_i)_i$ are modulated by a penalisation parameter $\beta \geq 0$, according to

$$\mathbf{H}_\beta = \mathbf{H} \Pi_\beta \quad (22)$$

where $\Pi_\beta = \text{diag}(\pi_1^{\frac{\beta}{\beta+1}}, \dots, \pi_n^{\frac{\beta}{\beta+1}})$, so we apply a weight $\pi_j^{\frac{\beta}{\beta+1}}$ to each grid point j . Hence, the previous definitions are slightly modified by changing \mathbf{H} by \mathbf{H}_β . In particular, this implies that the singular values of \mathbf{H}_β depend on the value of β . Recall that for a linear tracer such as CO at the city scale, the sensitivity matrix depends on meteorological parameters. Thus, the inclusion of parameter β alters the sensitivity matrix making it sensitive to geographical weights. This is a simple way to include geographical weights and it is equivalent to a change of variables in the parameter space with the same observation equation: $\mathbf{y} = \mathbf{H}_\beta \mathbf{x} + \varepsilon$, where $\mathbf{x}_\beta = \Pi_\beta^{-1} \mathbf{x}$.

Figure 3 shows two geographical weights used in this work, namely population density as in the census for 2002 (www.ine.cl) and CO emission fluxes by Saide et al. (2011b). The results are strongly modulated by the selection of the penalisation parameter β . This parameter controls the relative importance used to correct the sensitivity matrix given the minima and maxima of a given geographical field. Note that it is not possible to adjust β in order to balance the different quality indicators uniformly with respect to the network since the dependence with respect to β is logarithmic for information quantity and quadratic for total precision. The weight and the corresponding chosen value of β should be defined by the user and it will depend on the specific objective of the network. In our case, we will take a value of β that produces an even quality gain for urban and rural areas when adding a new station (see Section 4.3 and Fig. 5). In any case for $\beta = 0$, the weighting field is not taken into account (uniform weight). In the case of population and emission distributions, we chose β as a trade-off between urban and suburban areas; this corresponds to $\beta = 0.7$ for information gain and $\beta = 1.2$ for DFS.

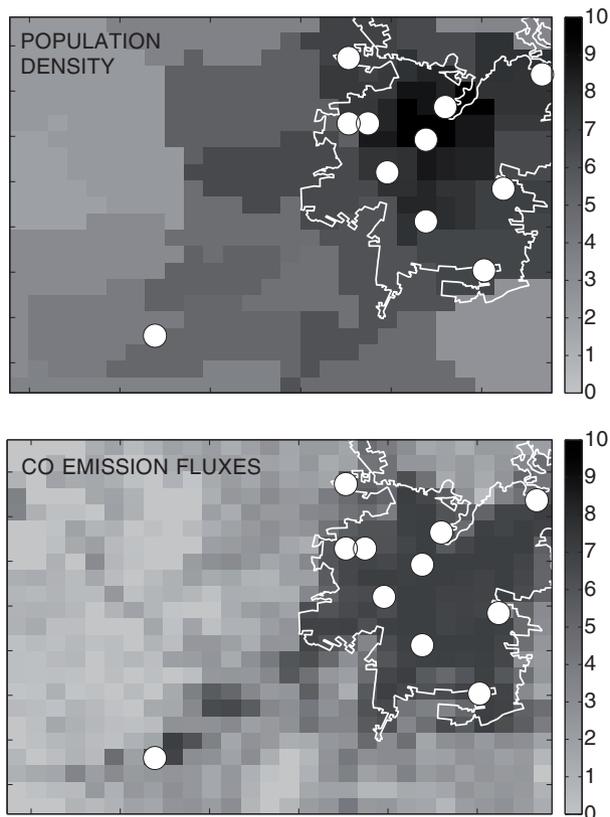


Fig. 3. Weighting functions applied in this work. The upper panel shows logarithm of population density (hab/km^2). The lower panel shows logarithm of normalised CO summer emission fluxes ($\text{mol km}^{-2} \text{h}^{-1}$). The white contour indicates the approximate urban–rural limit of Santiago in 2010. White circles show the location of monitoring stations in 2009.

4. Results

As described earlier, we consider information gain to be the best metric for assessing the quality of a network for the case of Santiago de Chile under summer conditions. Therefore, hereafter we show results using this indicator. We performed several tests and found similar results using DFS (not shown).

4.1. Experimental setting

For simulating CO dispersion over Santiago, we adopt the same modelling tool used by Saide et al. (2009) and Saide et al. (2011a), currently used by the Chilean Weather Office for forecasting ozone in summer. This system uses an air quality platform consisting of the CTM Polyphemus (Mallet et al., 2007) fed with meteorological fields calculated with the MM5 model (Grell et al., 1995). The simulations were performed for the 15–25 January 2002 period using four nested grids with horizontal resolution of 54, 18,

6 and 2 km, respectively. The 2 km resolution domain, centred in the city of Santiago, was fed into Polyphemus. The Polyphemus domain is also centred at Santiago, with 70×63 grid cells of 2 km resolution, and 12 vertical levels: from 13 to 6000 m above the surface, with finer resolution in the boundary layer. These simulations used zero initial and boundary conditions. Given the relatively clean Pacific air that arrives in Santiago, these assumptions appear as reasonable (e.g. Allen et al., 2011).

The sensitivity matrix (\mathbf{H}) was calculated using unitary emissions in a domain of 28×23 cells depicted in Fig. 1. For building (\mathbf{H}), we followed the same procedure as described by Saide et al. (2011a) for diurnal conditions between 6:00 and 20:00 hours, local time. Following these authors, for simplicity and lack of better knowledge, \mathbf{B} and \mathbf{R} are treated as scalar matrices, being the error in emissions 20%, and the error in observations 5%. Note that the null boundary conditions for CO are far away from the emission grid, since the CTM simulations were performed in a bigger and nested mesh. We point out that for the geographical weights considered in this study, and for the different initial networks used, the final optimal networks sometimes presented few stations located near the boundary of the emission grid (see Figs. 6 and 9) but not affecting the main orientation or distribution of the networks identified in Fig. 7. Therefore, for the purposes of this study, that is, influence of meteorology in the orientation of the network, influence of the geographical weight, robustness of the optimisation procedure, and so on, the chosen emission grid was adequate.

4.2. Reduction of the network

According to our calculations, if one were to remove a monitoring station from Santiago’s network as in 2009, using population density as the relevant weighting function, one would conclude that the removal of station O would result in the least decrease in quality of the network (Fig. 4, left panel). In contrast, one would not suggest removing stations Q, N, M, L or F, removal of which would result in a significant loss in network quality. If instead, the factor under consideration was CO emissions, one would still remove O and suggest keeping S, Q, P, V, M and T. Thus, depending on the network’s objectives, one would prioritise to keep different sets of stations.

It is interesting to note that the potential removal of stations O and R has the least impact on the quality of the network for assessing CO, either considering population density or emission fluxes. This is independent of their close geographical location. In fact, if we remove either station or if we merge the two stations by averaging their sensitivity values, the final quality decrease map is similar to the one shown in Fig. 4. This means that O and R share

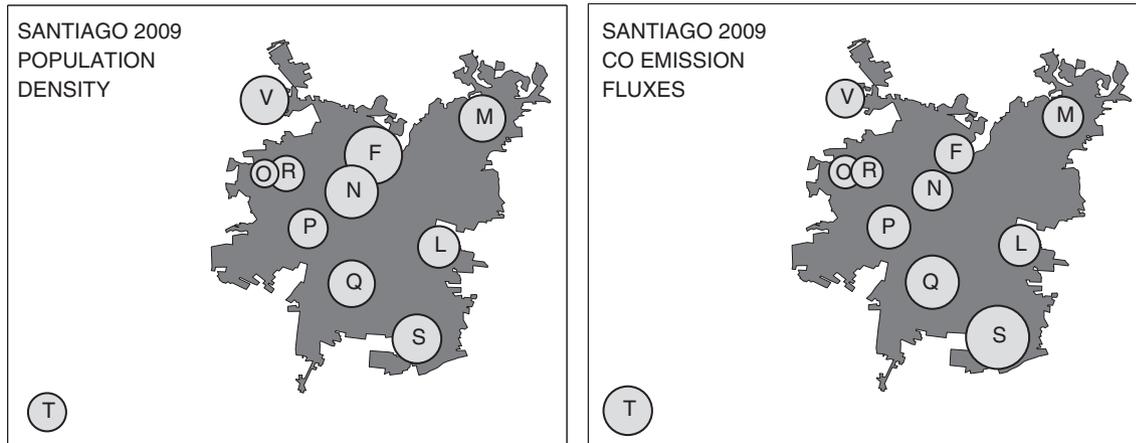


Fig. 4. Decreases in network's quality after removing stations. The centre of the circles indicates the position of the stations and each radius is a linear factor of the decrease in quality. The left (right) panel corresponds to the use of population density (emission fluxes) as weighting factors.

information and that the information provided by these stations is less relevant to the network's quality. This is firstly because the summer circulation in Santiago shows a clear southwest to northeast axis (following the predominant wind direction) and secondly because neither population nor emission density is distinct over this area. The former results in a relatively larger sensitivity of stations along the circulation axis and the latter results in too small a weighting factor to change the sensitivity for O and R. In the case of station S in the southeast of Santiago, the weighting factor is significant enough to counteract the fact that this station is off the predominant wind direction. Station T, which is located along the air circulation (SW to NE), increases its importance when emission criteria are prioritized because it is located close to a major highway with relatively large emissions. To verify these results, we calculated our metrics for the case $\beta = 0$, finding that stations located in areas with the lowest density of stations

(e.g. T) are those that have the largest impact on the quality of the network. It must be pointed out that these results use a sensitivity matrix restricted to diurnal conditions, when circulation and stability patterns are better captured by the dispersion model (e.g. Saide et al., 2011a). Thus, by no means do we question the location of stations O and R in the west of the city, which in fact capture the nocturnal maxima in pollution conditions (e.g. Saide et al., 2011b).

All in all, our results for summer conditions suggest that downtown stations are important from the point of view of total information content regarding CO. A previous study that considered multiple species using statistical information indexes by Osses et al. (2013) also suggested retention of downtown stations. These stations were in fact deemed to be the most representative of dispersion patterns for the Santiago basin, considering all species and all hours of the day, and all seasons of the year.

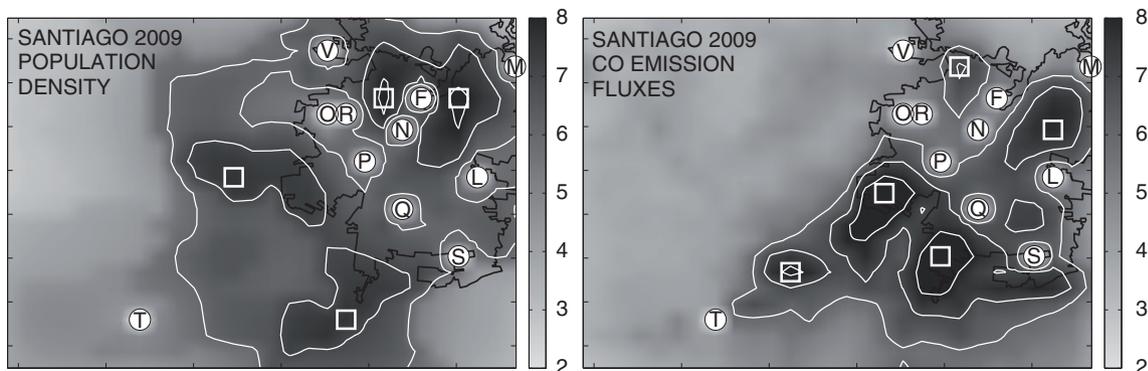


Fig. 5. Spatial distribution of increases in quality (shaded contours) for the 2009 network (white circles), after adding stations. White squares suggest the potential location of new stations coinciding with maxima in information gain. The left (right) panel shows the corresponding results using population density (emission fluxes) as weighting functions.

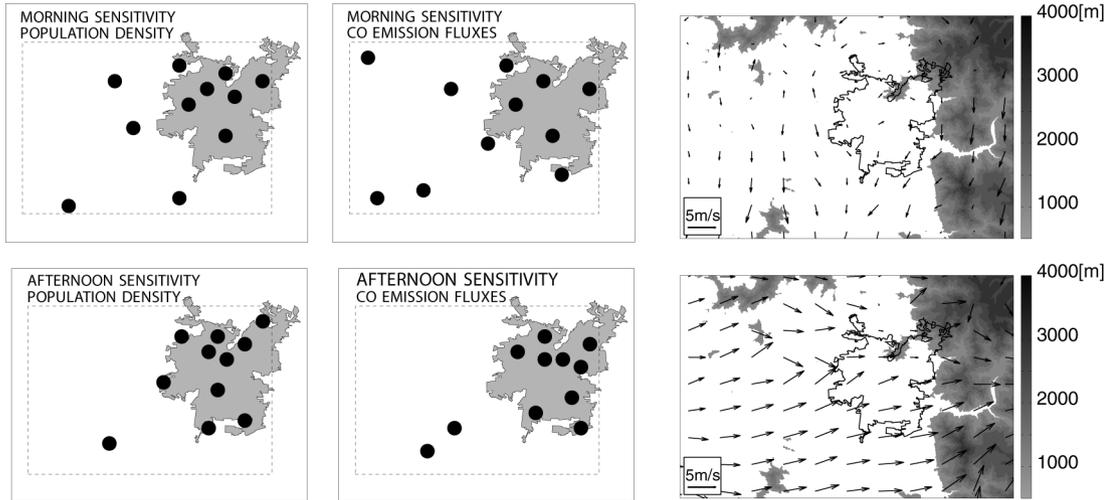


Fig. 6. Optimal networks according to wind pattern. Upper row: morning sensitivity weighted by population density (left), by CO emission fluxes (centre), and morning winds (right). Lower row: the same but for the afternoon sensitivity and afternoon winds.

4.3. Extension of the network

Here, we address the question of where to add stations so that the quality of the network increases by growing information content. In other words, we calculate the information gain provided by adding a monitoring station in every grid box lacking monitoring at present. According to the values of the sensitivity matrix and the weighting function used, one may determine the change in

information gain. In this way, we can build an information gain map as shown in Fig. 5. Both weighting functions suggest the addition of stations in the urban north and northeast areas of the basin, and in the suburban areas to the west and south of the city. This follows from the coincident geographical patterns of population and emission fields. In fact, the main difference between the suggested new sites is in a location southwest of the city, which appears when

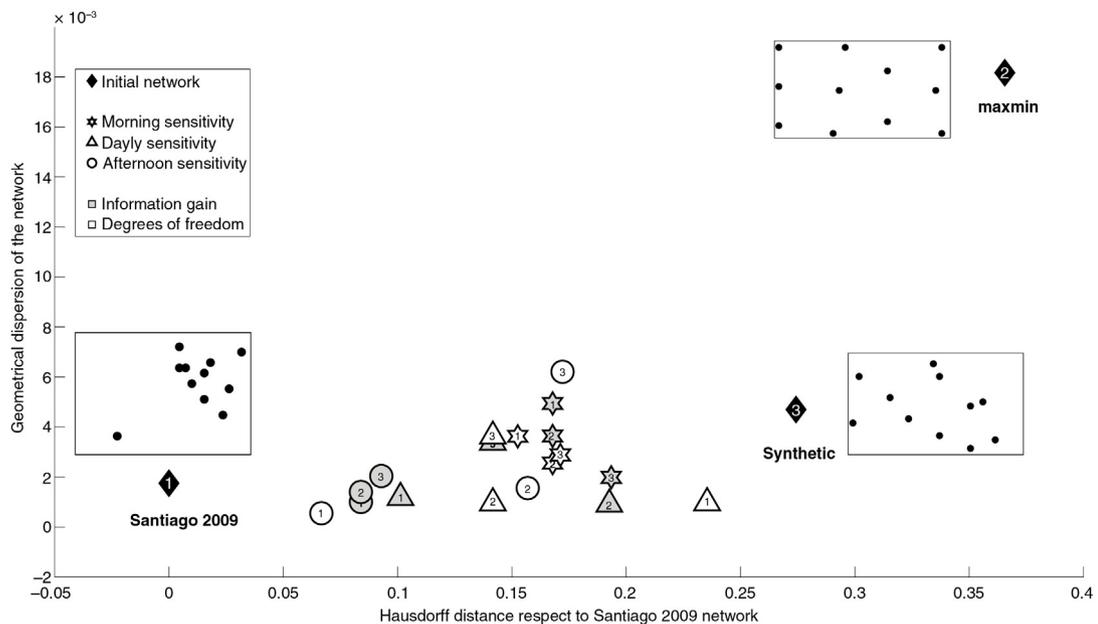


Fig. 7. Comparison among optimal networks using Hausdorff distance (in $^{\circ}$), and geometrical dispersion (in $^{\circ}$) for a set of 18 optimal networks for different hours of the day, initial conditions for SA and quality indicators. Black diamonds indicate the initial configurations used in the optimisation procedure: 1) Santiago 2009, 2) max-min and 3) synthetic. A sketch of each network is also shown. Stars indicate the optimal networks for morning hours, circles correspond to afternoon hours and triangles to average diurnal conditions. Networks obtained using information gain (degrees of freedom for signal) are shown by grey shaded (white) symbols.

weighting the sensitivity matrix by emissions along the highway, as was mentioned earlier. Interestingly, the extension of the network that took place in 2009 did add station T in the southwest of the basin, a conclusion that is supported by evaluating the information gain for the 2005 network. This is discussed in the following section.

4.4. Optimal network

Next, we address the question of designing the optimal network for a given sensitivity matrix, that is, we search for the network that maximises its quality indicator, in this case information gain. To do so, we start with a network of 11 stations evenly distributed in the domain using a max–min approach. We then optimise the network using simulated annealing. As in the previous section, the resulting optimal network follows from both the dispersion patterns embedded in the sensitivity matrix and the weighting function used. If we use \mathbf{H} with no weighting or a radial weighting centred in the middle of the domain, the optimal network shows an axis along the prevailing wind direction, with a more even distribution in the morning hours than in the afternoon hours (not shown). In the morning, the low wind speeds (≤ 3 m/s) and inhibited vertical mixing result in a larger sensitivity to the impact of neighbouring grid boxes compared to the impact of grid boxes farther away. Under these conditions, the resulting optimal distribution shows a larger spatial spread, that is, it privileges the coverage of the domain. In the afternoon, high wind speeds and vigorous vertical mixing result in a less geographically spread distribution. But since no weight is used, the stations tend to be located outside the city. Considering the modified sensitivity matrix by weighting functions, the corresponding optimal networks have similar characteristics, but the coverage of the domain is now complemented with stations in the city according to the weighting functions. These results are illustrated in Fig. 6.

4.5. Intercomparison of optimal networks

A network $\mathbf{Z} = \{z_1, z_2, \dots, z_N\}$ can be characterised by two independent geometrical parameters. One corresponds to the Hausdorff distance with respect to a reference network \mathbf{Z}' (Huttenlocher et al., 1993), defined as:

$$d_H(\mathbf{Z}, \mathbf{Z}') = \max \left\{ \sup_{z \in \mathbf{Z}} \inf_{z' \in \mathbf{Z}'} \|z - z'\|, \sup_{z' \in \mathbf{Z}'} \inf_{z \in \mathbf{Z}} \|z - z'\| \right\}, \quad (23)$$

which can be loosely defined as the maximum distance between two separated points but still close enough to be

members of the same set of points (network). A second parameter is the spread or spatial dispersion of the network:

$$\sqrt{\frac{1}{N} \sum_{z \in \mathbf{Z}} \|z - \bar{z}\|^2}, \quad (24)$$

where \bar{z} corresponds to the centroid of the network. Other shape indicators such as centroid, skewness, main direction and eccentricity may also be used. In Fig. 7, we show the corresponding Hausdorff distance and geometrical dispersion for a set of 18 optimal networks according to different hours of the day, initial networks for SA and quality indicators. These optimal networks were generated starting from three very different configurations, namely, the actual Santiago network in 2009, the maximum coverage configuration (max–min approach) and a synthetic network. The synthetic configuration has an orientation roughly perpendicular to Santiago’s network. We determine optimal networks using both information gain and DFS as quality indicators. Furthermore, to characterise the influence of wind fields, we present the results for morning and afternoon hours, and also for the average diurnal conditions. The Hausdorff distance is calculated with respect to Santiago’s network in 2009. In all cases, we use population density as weighting function. Independent of the starting network, hour of the day and quality indicator, all optimal networks have more spatial coverage (dispersion) and are distant from the current configuration of Santiago’s network. In other words, according to our calculations, the current configuration is suboptimal. Depending on the wind and dispersion conditions, the optimal networks may require more (morning) or less (afternoon) spatial coverage as discussed in Section 4.4. Note that both information gain and DFS are robust quality criteria with respect to the choice of the initial network and converge to optimal networks with similar characteristics as shown in Fig. 7.

4.6. Evolution of a network

Finally, we compare the actual evolution of Santiago’s air monitoring network with that of the optimal network according to our calculations for summer conditions. We use information gain weighted by emissions, starting each time from the actual distribution and adding/removing the number of stations actually added/removed over time in Santiago. Except in 2002, when stations were removed, the quality of the network has improved over time, with today’s network being roughly twice as informative as the network of 1988 (Fig. 8).

Note that the methodology to obtain an optimal network of 11 stations is different from the methodology to evolve a network by adding stations ‘one to one’ in an optimal way. In principle, one could obtain different results. Here we noted that the resulting networks are similar in shape and quality.

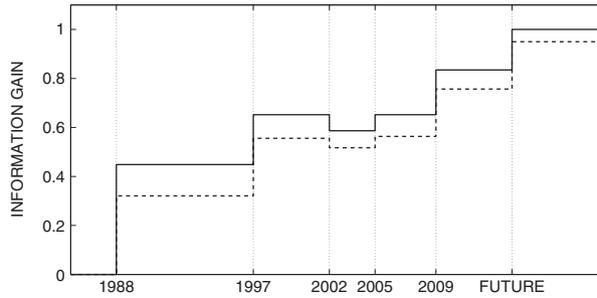


Fig. 8. Historical (dashed line) and optimal (continuous line) evolution of Santiago's air quality network using information gain and CO emission fluxes as weighting function. All values are normalised with respect to the optimal future scenario.

Figure 9 illustrates the spatial distribution of stations according to the actual evolution and the corresponding optimal evolution, considering information gain and a sensitivity matrix pondered by emissions. Each optimisation contemplates the same number of added stations as in the actual evolution for the transitions between 1988 and 1997, and between 1997 and 2009. We include a future scenario considering the addition of three stations, both starting with the 2009 configuration. Our results suggest adding stations in the southwest and northeast of the city. In fact, this example shows how to construct and justify a feasible change in the actual real network by mixing the different methodologies of the previous sections (adding, optimising, evolving).

5. Summary and conclusions

We have presented a methodology to evaluate the quality of an air pollutant monitoring network. The methodology

uses quality metrics calculated on the basis of a sensitivity matrix estimated by means of dispersion modelling, that is, it is formulated in a variational framework as emissions and observations are connected by means of a sensitivity matrix calculated with a CTM. The metrics evaluated are precision gain, information gain and DFS. These metrics are expressed in a common framework defined by the singular values of the sensitivity matrix and the covariance matrices that characterise the errors in emissions and observations. The methodology allows the assessment of adding or removing stations in a network and the more general situation of optimally distributing a given number of stations in a domain taking into account transport patterns in a city. Compared with statistical approaches addressing the optimal design of monitoring networks, the use of a variational framework is advantageous in that it explicitly factors in air circulation patterns and it allows the assessment of future scenarios. However, it requires the use of CTMs to calculate sensitivity matrices, albeit for linear tracers such as CO at the city scale.

We applied the methodology to evaluate the evolution and optimal design of Santiago de Chile's air quality network. We show evidence of the importance of considering multiple factors when designing and modifying a network. In particular, we address the impacts of transport patterns and of geographically varying characteristics such as population density and emissions. For instance, downtown stations in Santiago are key to the quality of the network, a finding that is consistent with previous studies (Osses et al., 2013). This follows from their location along the main axis of air circulation in a relatively flat area of the basin, and the high emission and population densities. Also, as the spatial coverage of the network has increased

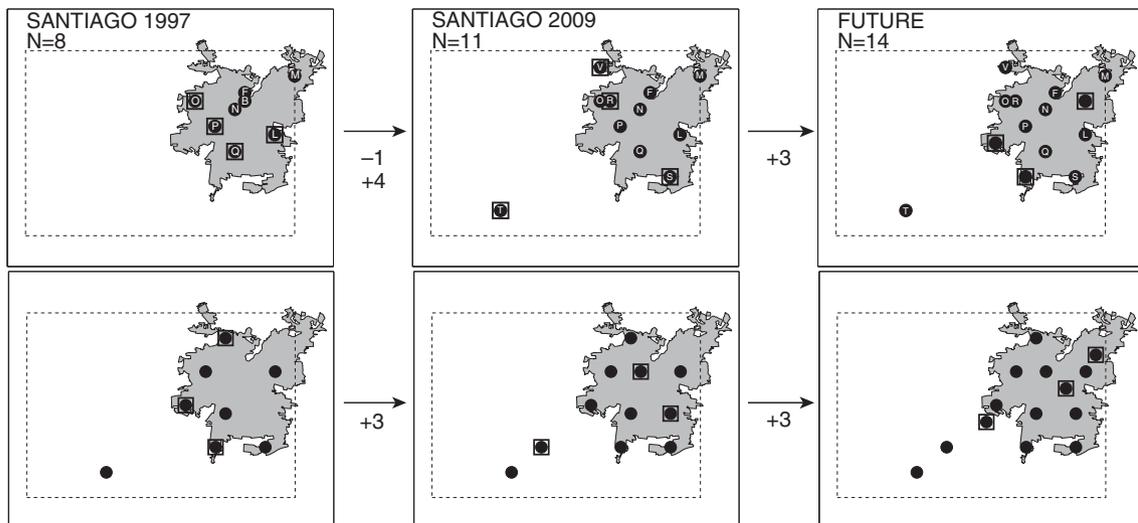


Fig. 9. Evolution of Santiago's air quality network. The upper (lower) panel shows the actual (optimal) evolution. New stations are highlighted by a black square.

by adding peripheral stations (e.g. T, V, S), so has its quality increased as determined by information gain and other metrics. In fact, according to our results, the quality of the network has increased over time, with the present configuration being twice as informative for CO as it was in 1998. Still, the evolution has been suboptimal, and it can be significantly improved by integrating more suburban stations in the southwest of the basin. Again, these conclusions refer to diurnal conditions in summer. The purpose of this work was nevertheless to review the use of various quality metrics for characterising monitoring networks and to illustrate its use in a coherent framework. Future works should address the development of more accurate scenarios, consider nocturnal conditions and factor in the seasonal variability in circulation patterns, as well as other design objectives such as identifying maximum impact areas in terms of air pollution.

6. Acknowledgements

This work was initially carried out with the aid of a grant from the Inter-American Institute for Global Change Research (IAI) CRN II 2017 which is supported by the US National Science Foundation (Grant GEO-0452325), and it was completed with support by Conicyt/FONDAP/15110009. The authors greatly appreciate comments and suggestions from Dr. N. Huneus, Dr. K. Pistone and Dr. P. Saide. The authors also appreciate the comments and suggestions by two anonymous referees. The second author acknowledges the partial support of Fondecyt 1151512 and ACT 1106 Conicyt grants.

References

- Abida, R. and Bocquet, M. 2009. Targeting of observations for accidental atmospheric release monitoring. *Atmos. Environ.* **43**, 6312–6327.
- Abida, R., Bocquet, M., Vercauteren, N. and Isnard, O. 2008. Design of a monitoring network over France in case of a radiological accidental release. *Atmos. Environ.* **42**(21), 5205–5219.
- Ainslie, B., Reuten, C., Steyn, D. G., Le, N. D. and Zidek, J. V. 2009. Application of an entropy-based Bayesian optimization technique to the redesign of an existing monitoring network for single air pollutants. *J. Environ. Manage.* **90**(8), 2715–2729.
- Allen, G., Coe, H., Clarke, A., Bretherton, C. and Wood, R. 2011. South East Pacific atmospheric composition and variability sampled along 20 S during VOCALS-Rex. *Stat. Sci.* **8**(1), 10–15.
- Bertsimas, D. and Tsitsiklis, J. 1993. Simulated annealing. *Atmos. Chem. Phys.* **11**, 5237–5262.
- Bocquet, M. 2005a. Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: theory. *Q. J. Roy. Meteorol. Soc.* **131**, 2191–2208.
- Bocquet, M. 2005b. Reconstruction of an atmospheric tracer source using the principle of maximum entropy. II: applications. *Q. J. Roy. Meteorol. Soc.* **131**, 2209–2223.
- Bouttier, P. and Courtier, F. 1997. Data assimilation: concepts and methods. *Training Course Notes of ECMWF*, Vol. 53, Reading, UK.
- Caseltan, W. and Zidek, J. 1984. Optimal monitoring network designs. *Stat. Probab. Lett.* **2**, 223–227.
- Cioaca, A. and Sandu, A. 2014. Low-rank approximations for computing observation impact in 4D-Var data assimilation. *Comp. Math. Appl.* **67**, 2112–2126.
- Daley, R. 1993. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge.
- Elkamel, A., Fatehifar, E., Taheri, M., Al-Rashidi, M. S. and Lohi, A. 2008. A heuristic optimization approach for Air Quality Monitoring Network design with the simultaneous consideration of multiple pollutants. *J. Environ. Manage.* **88**, 507–516.
- Fisher, M. 2003. Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. *ECMWF Tech. Memo. No. 397*, 18 pp. Online at: http://old.ecmwf.int/publications/library/ecpublications/_pdf/tm/301-400/tm397.pdf
- Gallardo, L., Dawidowski, L., Rojas, N. J., Andrade, M. F. and Osses, M. 2012. Evaluation of vehicle emission inventories for carbon monoxide and nitrogen oxides for Bogota, Buenos Aires, Santiago, and Sao Paulo. *Atmos. Environ.* **47**, 12–19.
- Gomez-Losada, A., Lozano-Garca, A., Pino-Mejas, R. and Contreras-Gonzalez, J. 2014. Finite mixture models to characterize and refine air quality monitoring networks. *Sci. Total Environ.* **485–486**, 292–299.
- Grell, G. A., Dudhia, J. and Stauffe, D. R. 1995. *A Description of the Fifth Generation Penn State/NCAR Mesoscale Model (MM5)*. NCAR Tech. Note TN-398+STR, 1995.
- Horn, R. and Johnson, C. 2012. *Matrix Analysis*. Cambridge University Press, Cambridge.
- Husain, T. and Khan, H. 1983. Shannon’s entropy concept in optimum air monitoring network design. *Sci. Total Environ.* **30**, 181–190.
- Huttenlocher, D., Klanderman, G. and Rucklidge, W. 1993. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(9), 850–863.
- Johnson, C. 2003. *Information Content of Observations in Variational Data Assimilation*. PhD Thesis. Department of Meteorology, The University of Reading, UK.
- Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge.
- Kanaroglou, P. S., Jerret, M., Morrison, J., Beckerman, B., Arain, M. A. and co-authors. 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location–allocation approach. *Atmos. Environ.* **39**, 2399–2409.
- Kao, J. J. and Hsieh, M. R. 2006. Utilizing multiobjective analysis to determine an air quality monitoring network in an industrial district. *Atmos. Environ.* **40**, 1092–1103.
- Kullback, S. 1959. *Information Theory and Statistics*. Wiley, New York.
- Lahoz, W. A. and Schneider, P. 2014. Data assimilation: making sense of Earth Observation. *Front. Environ. Sci.* **2**, 16.
- Langstaff, J., Seigneur, C., Mei-Kao, L., Behar, J. and McElroy, J. L. 1967. Design of an optimum air monitoring network for exposure assessments. *Atmos. Environ.* **21**, 1393–1410.

- Mallet, V., Quélo, D., Sportisse, B., Ahmed de Biasi, M., Debry, E. and co-authors. 2007. Technical note: the air quality modeling system Polyphemus. *Atmos. Chem. Phys. Discuss.* **7**, 6459–6486.
- Mofarrah, A. and Husain, T. 2010. A holistic approach for optimal design of air quality monitoring network expansion in an urban area. *Atmos. Environ.* **44**(3), 432–440.
- Nychka, D. and Saltzman, N. 1998. Design of air-quality monitoring networks. *Case Stud. Environ. Stat.* **132**, 51–76.
- Osses, A., Gallardo, L. and Faundez, T. 2013. Analysis and evolution of air quality monitoring networks using combined statistical information indexes. *Tellus B.* **65**, 19822, DOI: <http://dx.doi.org/10.3402/tellusb.v65i0.19822>
- Rodgers, C. 2000. Inverse methods of atmospheric sounding, theory and practice. *Atmos. Ocean. Planet. Phys.* **2**, 12–41.
- Romary, T., De Fouquet, C. and Malherbe, L. 2011. Sampling design for air quality measurement surveys: an optimization approach. *Atmos. Environ.* **45**, 3613–3620.
- Romary, T., Malherbe, L. and Fouquet, C. 2014. Optimal spatial design for air quality measurement surveys. *Environmetrics.* **25**, 16–28.
- Romero, H. and Ordenes, F. 2004. Emerging urbanization in the Southern Andes: environmental impacts of urban sprawl in Santiago de Chile on the Andean piedmont. *Mt. Res. Dev.* **24**, 197–201.
- Ruiz-Cárdenas, R., Ferreira, M. and Schmidt, A. 2012. Evolutionary Markov chain Monte Carlo algorithms for optimal monitoring network designs. *Stat. Methodol.* **9**(1–2), 185–194.
- Saïde, P., Bocquet, M., Osses, A. and Gallardo, L. 2011a. Constraining surface emissions of air pollutants using inverse modeling: method inter-comparison and a new two-step multi-scale approach. *Tellus B.* **63**, 360–370.
- Saïde, P., Carmichael, G., Spak, S., Gallardo, L., Osses, A. and co-authors. 2011b. Forecasting urban PM₁₀ and PM_{2.5} pollution episodes in very stable nocturnal conditions and complex terrain using WRF-Chem CO tracer model. *Atmos. Environ.* **45**, 2769–2780.
- Saïde, P., Osses, A., Gallardo, L. and Osses, M. 2009. Adjoint inverse modeling of a CO emission inventory at the city scale: Santiago de Chile case. *Atmos. Chem. Phys. Discuss.* **9**, 6325–6361.
- Saunier, O., Bocquet, M., Mathieu, A. and Isnard, O. 2011. Model reduction via principal component truncation for the optimal design of atmospheric monitoring networks. *Atmos. Environ.* **43**, 4940–4950.
- Shannon, C. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3), 379–423.
- Shindo, J., Oi, K. and Matsumoto, Y. 1990. Considerations on air pollution monitoring network design in the light of spatio-temporal variations of data. *Atmos. Environ. B Urban Atmos.* **24**, 335–342.
- Singh, K., Sandu, A., Jardak, M., Bowman, K. and Lee, M. 2013. A practical method to estimate information content in the context of 4D-var data assimilation. *SIAM/ASA J. Uncertain. Quantif.* **1**, 106–138.
- Sturman, A., Titov, M. and Zawar-Reza, P. 2011. Selecting optimal monitoring site locations for peak ambient particulate material concentrations using the MM5-CAMx4 numerical modelling system. *Sci. Total Environ.* **409**, 810–821.
- Trujillo-Ventura, A. and Ellis, J. H. 1991. Multiobjective air pollution monitoring network design. *Atmos. Environ. A Gen Topics.* **25**, 469–479.
- Tseng, C. and Chang, N. B. 2001. Assessing relocation strategies of urban air quality monitoring stations by GA-based compromise programming. *Environ. Int.* **26**, 523–541.
- Wang, K., Ding, Y., Zhao, H., Hou, L. and Sun, F. 2013. Optimization of air pollutant monitoring stations based on genetic algorithm. In: *Proceedings of the Emerging Intelligent Data and Web Technologies (EIDWT) Fourth International Conference on 2013*, IEEE, 680–684.
- Wu, L. and Bocquet, M. 2011. Optimal redistribution of the background ozone monitoring stations over France. *Atmos. Environ.* **45**(3), 772–783.
- Wu, L., Bocquet, M. and Chevallier, M. 2010. Optimal reduction of the ozone monitoring network over France. *Atmos. Environ.* **44**(25), 3071–3083.
- Zhu, T., Melamed, M. L., Parrish, D., Gauss, M., Gallardo, L. and co-authors. 2012. WMO/IGAC impacts of megacities on air pollution and climate. (GAW Report No. 205, 205), World Meteorological Organization, Geneva.
- Zidek, J. V., Shaddick, G. and Taylor, C. G. 2014. Reducing estimation bias in adaptively changing monitoring networks with preferential site selection. *Ann. Appl. Stat.* **8**, 1640–1670.
- Zidek, J. V., Sun, W. and Le, N. D. 2000. Designing and integrating composite networks for monitoring multivariate Gaussian pollution. *Appl. Stat.* **49**(1), 63–79.
- Zidek, J. V. and Zimmerman, D. L. 2010. Chapter 10. Monitoring network design. In: *Handbook of Spatial Statistics* (eds. A. E. Gelfand, P. J. Diggle, M. Fuentes and P. Guttorp), CRC Press, Taylor and Francis Group, New York, 131–148.