Int. J. Environ. Sci. Technol. (2012) 9:195–202
DOI 10.1007/s13762-012-0022-1

ORIGINAL PAPER

# Assessing precision and accuracy of atmospheric emission inventories

**J. I. Huertas · M. E. Huertas · J. Díaz**

Received: 30 June 2011 / Revised: 20 August 2011 / Accepted: 17 November 2011 / Published online: 29 February 2012

**Abstract** Assuming that state-of-the-art air quality models are accurate, then the precision and accuracy of their results directly depend on the precision and accuracy of their geographical, meteorological and emission input data. There are important applications, such as open pit mining, in which emission data are the main source of uncertainty. In such cases, historical air quality experimental data are typically available. The present work proposes a backward air quality simulation approach to assess the accuracy of emission inventories for these applications, with the goal of identifying sources that are over or underestimated. This approach consists of finding constants of the linear combination of the estimated emission that maximize $R^2$ and make the slope equal to one in the linear correlation analysis when the results from the air quality model are compared to the experimental measurements of air quality. This methodology was applied to the case of the mining region in northern Colombia. As one of the largest open pit coal mining regions in the world, this region consists of seven independent mines with no relevant additional sources of emission. Use of the proposed methodology allowed quantification of the amount by which companies over or underestimated their emission, as well as quantification of uncertainties due to sources not considered in the model but that locally affect each monitoring station.

**Keywords** AERMOD model · Air quality modeling · Opencast mining

## Introduction

Environmental impact assessment of future or existing atmospheric pollutant sources generally involves air quality modeling. Such studies model the dispersion and concentration of pollutants in time and space at the surface level over an area of interest. By contrasting the obtained concentrations against local air quality standards, researchers can determine the environmental feasibility of implementing a new or continuing an existing project. However, because the political and economic implications of rejecting large projects due to environmental reasons are costly for environmental authorities and for the overseeing company or institution, the accuracy and precision of negative air quality model results are frequently questioned. Accuracy refers to the nearness of estimated pollutant concentrations to the true values, whereas precision refers to the dispersion (lack of reproducibility) of the estimated pollutant concentrations (Neter et al. 2005). To make informed evaluations regarding the accuracy of predicted results in time and space, the results must be compared with measured values in a meaningful, quantitative way. Several metrics, including those shown in Table 1, have been used to achieve this objective. Despite extensive work on the development of acceptable air quality models (USEPA 2005; Holmes and Morawska 2006) and implementation procedures (USEPA 2004), several issues in this matter remain unresolved. Well-accepted models include AERMOD, which is currently recommended by USEPA for modeling point sources of emission (USEPA 2011). This model accurately predicts long-term (∼1 year)

J. I. Huertas (✉) · J. Díaz
Tecnológico de Monterrey, Eduardo Monroy Cárdenas No 2000, Toluca, Mexico
e-mail: jhuertas@itesm.mx

M. E. Huertas
Texas A&M University, College Station, TX 77843, USA

**Table 1** Metrics to evaluate the capability of air quality models to reproduce, in time and space at surface level, experimental observations (Mosca et al. 1998; Walker et al. 1999; ASTM 2000; Li et al. 2011; Zawar-Reza et al. 2005)

| Metric | Definition | Perfect value | Range | Meaning |
|---|---|---|---|---|
| Index of agreement | $d = 1 - \frac{\Sigma(P_i - M_i)^2}{\Sigma\left(\lvert P_i - \overline{M}\rvert + \lvert M_i - \overline{M}\rvert\right)^2}$ | 1 | $0 < d < 1$ | Evaluate the degree the predictions of models are error free |
| Bias | $B = \frac{1}{N}\sum P_i - M_i$ | 0 | $-\infty < B < \infty$ | Average difference between paired predicted and measured values |
| Fractional bias | $\mathrm{FB} = 2\frac{\bar{P} - \bar{M}}{\bar{P} + \bar{M}}$ | 0 | $-2 < \mathrm{FB} < 2$ | Evaluates the sub or over estimation of the model |
| Geometric mean bias | $\mathrm{MG} = \prod \left(\frac{P_i}{M_i}\right)^{1/N}$ | 1 | $0 < \mathrm{MG} < \infty$ | It gives the same weight to pairs showing the same ratio, independently of the absolute value of the data |
| Normalize mean square error | $\mathrm{NMSE} = \frac{1}{N}\sum \frac{(P_i - M_i)^2}{\bar{P}\,\bar{M}}$ | 0 | $\mathrm{NMSE} > 0$ | Provide information on the deviation and not on the over or under estimations |
| Root of the mean square error | $\mathrm{RMSE} = \left(\frac{1}{N}\sum (P_i - M_i)^2\right)^{1/2}$ | 0 | $\mathrm{RMSE} > 0$ | Similar to NMSE. It shows the most striking differences among models |
| Coefficient of determination | $R^2 = 1 - \frac{\sum (P_i - M_i)^2}{\sum (M_i - \bar{M}_i)^2}$ | 1 | $0 < R^2 < 1$ | Measure dispersion. It is the proportion of the variability of the data that is accounted by the model |
| FAC2 | Fraction of data for which $0.5 < P_i/M_i < 2$ | 1 | $0 < \mathrm{FAC2} < 1$ | It measures both dispersion and over estimations |

$P_i$ Predicted pollutant concentration at time or space $i$, $M_i$ Measured pollutant concentration at time or space $i$, $\bar{P}$ Average predicted pollutant concentration, $\bar{M}$ Average measured pollutant concentration, N Number of data points

average concentrations, but has limitations in estimating short-term concentrations (USEPA 2011).

For the purposes of this work, it is assumed that the model used accurately estimates long-term average concentrations; therefore, time is excluded from the evaluation of precision and accuracy. Because these models are deterministic, the results that they produce under unchanged input data have no dispersion; thus, they are precise. Under these circumstances, the accuracy and precision of the results reported by these well-established models depend only on the accuracy and precision of the geographical, meteorological, and emission input data. There are important applications, such as in open pit mining, in which emission data are the main source of uncertainty, due to the fugitive nature of the sources and impossibility of directly measuring the emissions. Geographical and meteorological data are always easy to measure and, therefore, are reliable.

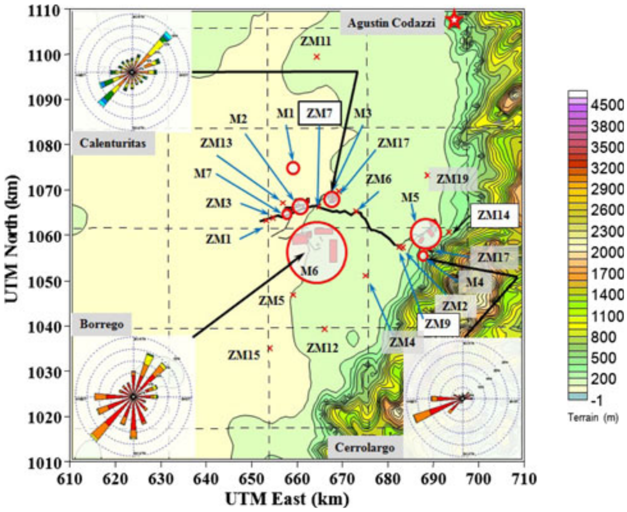

**Fig. 1** Location of mines (M), monitoring stations (ZM) and meteorological stations in the open pit mining region in northern Colombia

## Open pit mining in northern Colombia

In 2010, the mining region in northern Colombia (which is considered one of the largest open pit mining regions in the world) consisted of seven open pit mines with an approximate coal production of 36 Mtons/year. Figure 1 describes the mining region, showing the location of the mines, air quality monitoring stations, and meteorological stations. The major air pollutants arising from the mining operations include total suspended particulate matter (TSP) and particles with an equivalent aerodynamic diameter smaller than 10 µm ($PM_{10}$) (Chakraborty et al. 2002; Sinha and Banerjee 1997). Table 2 lists the total mass of TSP emitted during 2009 at each mine. The TSP and $PM_{10}$ in open pit mining regions reduce air quality and can cause silicosis, black lung (CWP), and increased mortality. They also reduce visibility and affect surrounding flora and fauna (NIOSH 2005; Wheeler et al. 2000).

In 2006, the Colombian environmental authority initiated an air quality network to monitor TSP and $PM_{10}$ in the mining region. Since 2007, this network has been reporting readings that exceed the standards for daily and annual TSP and $PM_{10}$ concentrations in some nearby villages (Huertas et al. 2012a). To assess the impact of mining operations on

**Table 2** Open pit coal mine TSP emissions for 2009

| Sources | TSP emission intensities g/(m$^2$ s) | | | | | | |
|---|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M4 | M5 | M6 | M7 |
| Pit | 5.0681E−05 | 3.4910E−06 | 1.1912E−05 | 9.1367E−04 | 9.8477E−06 | 2.6908E−05 | 1.2887E−05 |
| Dump | 8.6049E−06 | 5.4693E−06 | 5.8058E−06 | 4.3629E−04 | 1.0434E−05 | 6.6505E−06 | 7.0641E−06 |
| Stock | 3.5024E−05 | 2.0435E−04 | 1.3037E−03 | 2.0729E−03 | 2.9833E−03 | 1.4394E−04 | 5.3191E−05 |
| Pit-dump via | 2.1887E−04 | 5.5485E−06 | 1.4318E−04 | 2.1503E−04 | 3.5172E−05 | 7.2359E−05 | 2.6565E−05 |
| Pit-stock via | 4.3951E−05 | 1.0493E−07 | 1.4638E−05 | 2.3989E−07 | 1.1574E−06 | 8.8889E−04 | 9.9652E−06 |
| Beltway | 5.4694E−05 | 1.4430E−07 | 0.0000E + 00 | 9.3328E−07 | 3.1034E−06 | 4.3277E−05 | 3.8399E−06 |
| Total TSP emissions g/(m$^2$ −s) | 3.5713E−04 | 2.1911E−04 | 1.4792E−03 | 3.6390E−03 | 3.0430E−03 | 1.1820E−03 | 1.1351E−04 |
| Total TSP emissions kg/year | 3.9276E+06 | 7.4812E+05 | 2.3139E+06 | 1.6727E+06 | 3.0230E+06 | 1.4819E+07 | 9.0708E+05 |
| Total coal production ton/year | 1.8148E+06 | 1.038E+06 | 5.984E+06 | 6.000E+05 | 4.730E+06 | 1.840E+07 | 1.600E+06 |

Pit, dump, stock and pit-dump via, pit-stock via and beltway are the area sources considered within mine (M’s)

**Table 3** Meteorological information used to simulate TSP and PM$_{10}$ dispersion in the study area

| Name:<br>Type: | Borrego<br>Automatic | Calenturitas<br>Automatic | Cerrolargo<br>Automatic | Agustín Codazzi<br>Observatory |
|---|---|---|---|---|
| Location (UTM km) | | | | |
| *x* | 668.4 | 665.6 | 689.24 | 690.1 |
| *y* | 1,059.5 | 1,067.3 | 1,056.9 | 1,089.3 |
| Primary parameters | | | | |
| Velocity | √ | √ | √ | – |
| Direction | √ | √ | √ | – |
| Max velocity | √ | √ | √ | – |
| Max direction | √ | √ | √ | – |
| Temperature | √ | √ | √ | – |
| Pressure | √ | √ | √ | – |
| Radiation | √ | | √ | – |
| Rainfall | √ | √ | √ | √ |
| Humidity | √ | √ | √ | – |
| Evaporation | √ | √ | √ | – |
| Cloud cover | – | – | – | √ |
| Resolution | | | | |
| Minutes | 30 | 15 | 30 | 3 readings per day |
| % Availability | 98.8 | 99.7 | 99.7 | 99.6 |

regional air quality and to establish an appropriate clean air program for the region, the Colombian environmental authority developed a detailed study of the temporal and spatial variations of the surface levels of TSP and PM$_{10}$ concentrations using AERMOD. Tables 2 and 3 show the emission inventory and the meteorological data used as inputs. The results of the study were used to estimate the contribution of each mine to pollution in each village within the mining region, thereby allowing the environmental authority to determine the appropriate contribution of each mining company towards financing decontamination measures (Huertas et al. 2012a).

The accuracy of the results obtained by the Colombian environmental authority depends directly on the accuracy of the emission data. Masses of emitted pollutants were estimated from the emission inventories, through the following equation (USEPA 2008):

$$E_i = \sum A_j E_{f,ij}\left(1 - \eta_{ij}\right) \qquad (1)$$

where $E_i$ is the mass of pollutant $i$ emitted at each mine, $A_j$ is intensity of the activity or operation $j$ of the productive process being evaluated, $E_{f,ij}$ is the emission factor for pollutant $i$ when it is performing activity $j$, and $\eta_{ij}$ is the efficiency of the control measure to prevent emissions of pollutant $i$ in activity $j$.

Institutions like the USEPA have compiled emission factors ($E_{f,ij}$) for most of the activities involved in open pit mining (USEPA 2006a, b, 2008). Such emission factors have been obtained for open pit mines in the western region of the United States (Axetell and Cowherd 1981; Muleski 1990; USEPA 1994; Zeller et al. 1979). However, the applicability of these emission factors has been questioned for open pit mining in tropical areas such as northern Colombia, where the emission rate should be different due to differences in geological and climatic conditions, mining conditions, site practices, and mitigation practices (Ghose 2004). Furthermore, there are no emission factors reported for subactivities in mine operations, such as handling and storage in piles and transportation over unpaved roads. In these cases, emission factors from similar activities are used as an approximation. These circumstances have raised additional questions.

To address these issues, the present paper reviews several alternatives to evaluate the accuracy and precision of emission inventories and proposes an additional alternative. As a case study, the proposed methodology was applied to the mining region in northern Colombia. This research work was developed at Tecnológico de Monterrey (México) and Texas A&M University (USA) during 2010 and 2011. It is out of the scope of the present work to adjust the available emission factors to local conditions.

### Approaches to evaluate atmospheric emission inventories

Previous studies have utilized aggregating uncertainties and bootstrap methods to evaluate the precision of emission inventories. In aggregating uncertainties, the uncertainty of an expression $f = f(x_i)$ is expressed by the error propagation equation (Neter et al. 2005):

$$\Delta f = \sum \frac{\partial f}{\partial x_i} \Delta x_i \tag{2}$$

For the case of Eq. 1, this expression becomes:

$$\Delta E_i = \sum E_{f,ij}(1-\eta_{ij})\Delta A_j + A_j(1-\eta_{ij})\Delta E_{f,ij} + E_{f,ij}A_j\Delta\eta_{ij} \tag{3}$$

where $\Delta E_i$ is uncertainty in the total emissions of pollutant $i$ for a given mine, $\Delta A_j$ is the uncertainty in the intensity of the activity or operation $j$, $\Delta E_{f,ij}$ is the uncertainty in the emission factor for pollutant $i$ when it performs activity $j$, and $\Delta\eta_{ij}$ is the uncertainty in the efficiency of the control measure to prevent emissions of pollutant $i$ in activity $j$.

Alternatively, the following equation can be used (EEA 2009):

$$\Delta E_i = \frac{\left[\prod\left(E_{f,ij}(1-\eta_{ij})\Delta A_j + A_j(1-\eta_{ij})\Delta E_{f,ij} + E_{f,ij}A_j\Delta\eta_{ij}\right)^2\right]^{\frac{1}{2}}}{\sum E_{f,ij}A_j(1-\eta_{ij})} \tag{4}$$

The bootstrap method uses Monte Carlo resampling of $A_j$, $E_{f,ij}$ and $\eta_{ij}$ from their respective distribution functions to estimate the uncertainty of the emission inventory $E_i$. This method requires thousands of samplings. Uncertainty is expressed as the range (confidence interval) in which there is a 95% probability that the actual value of $E_i$ be within that range. It can be used with any type of distribution function for these three input variables (Neter et al. 2005; Ramirez et al. 2008).

Methods to enhance the accuracy of emission inventories include standardizing procedures for preparing emission inventories and the comparative analysis of similar emission inventories. Use of standardizing procedures consists of establishing a unified or standard procedure to prepare and report the emission inventories for a specific application. The most well-known examples are MOVES, which is the USEPA-recommended protocol to estimate emissions from a fleet of vehicles (USEPA 2010), and the guidelines for the elaboration of greenhouse gas emission inventories (EEA 2009). Use of standardizing procedures does not necessarily ensure accurate results, but does make them comparables and reduces uncertainties. For the case of open pit mining, Huertas et al. (2012b) reported a unified methodology for performing TSP and $PM_{10}$ emission inventories.

The comparative analysis of similar emission inventories focuses on comparing the emission inventory results for the same application. This alternative requires an identification of a proper metric of comparison. For the case of open pit coal mining, the proposed metric is the kilograms of TSP or $PM_{10}$ emitted per ton of coal extracted. Small natural variations in this metric are expected, because open pit mining always involves the same operations, and mine companies have implemented similar emission control technologies. Large variations may be attributed to errors in the calculation of the emission inventories. Huertas et al. (2012b) compared emission inventories for the case of the mining region in northern Colombia. They reported that, on average, a mine company generates 0.726 kg of TSP and 0.180 kg of $PM_{10}$, respectively, per Mg of coal produced.

Previous methods have sought to quantify and reduce the uncertainty of emission inventories. However, these methods cannot quantify how inaccurate the data are. Therefore, it is desirable to create a method to quantify the accuracy of atmospheric emission inventories.

## Materials and methods

The proposed methodology to assess the accuracy of emission inventories of a given pollutant compares experimental data of pollutant concentration with concentrations obtained through an air quality model that use emissions inventory in a backward modeling approach.

Let $n$ be the number of emission sources in the region of interest. This method consists on affecting the estimated emission of each source $E_i$ by a constant $A_i$, where $A_i > 0$. The resulting emission inventory $\left(E = \sum_i^n A_i E_i\right)$ is used as data input of a well-established air quality model to obtain the pollutant concentration at the $j$ points where that pollutant has been measured for a long period of time. Then, the correlation between the numerical results ($P_{t,j}$) and experimental values ($M_{t,j}$) was statistically evaluated in terms of long-term averages ($\overline{M}_j$ vs. $\overline{P}_j$). The sets of constants $\{A_i\}$ are changed systematically to determine the set of constants $\{A_i\}_{\max}$ that maximizes $R^2$ and simultaneously makes the slope $= 1$.

The obtained set of constants $\{A_i\}_{\max}$ measures the inaccuracy of each emission inventory. $A$ value of $A_i > 1$ indicates that the corresponding emission inventory was underestimated by that amount. The expression $1 - R^2$ quantifies the relative no systematic contributions to the measured pollutant concentration of the sources of pollutants that were not included in the model. The difference between the measured background concentration ($M_{\mathrm{B}}$) and the constant of the linear regression ($b$) determines the systematic contribution to the measured pollutant concentration of all sources of pollutants that were not included in the model.

This methodology makes the following assumptions: (1) all of the relevant sources of emission in the region are included in the emission inventory; (2) historical air quality experimental data exist for at least 1 year at several points (>3) within the regions of interest; and (3) the air quality model estimates correctly the dispersion of the pollutant under consideration.

The proposed methodology differs from the receptor models in that receptor models do not use pollutant emission, meteorological data, or chemical transformation mechanisms to estimate the contribution of sources to receptor concentrations. Instead, receptor models use the chemical and physical characteristics of gases and particles measured at the source and receptor to identify the presence of and to quantify source contributions to receptor concentrations (USEPA 2011).

## Results and discussion

The empty blue circles in Fig. 2 show the base case scenario, which compares the measured annual geometric mean TSP concentration against the estimated TSP concentration at 9 points within the mining region. Estimated TSP concentrations were obtained using AERMOD with meteorological data for the year 2009, as reported by the Calenturitas station. Additionally, this case used emission inventories reported by each of the seven mines operating in the region for the same year. The unified methodology described in Huertas et al. (2012b) was used, except for the cases of mines 1 and 6, for which a third methodology was used.

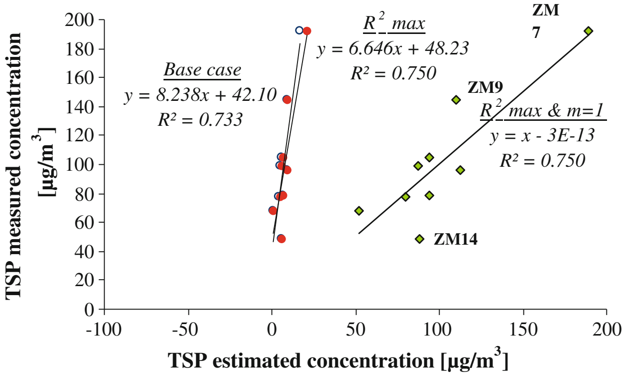

**Fig. 2** Linear correlation of measured versus estimated annual mean geometric TSP concentrations for the base case and optimized scenarios

As shown in Table 4, a correlation between the experimental and estimated TSP concentrations exists for the base case scenario, with a confidence $(1 - p) = 99.7\%$. The level of this correlation is $r = 0.856$. The slope of the linear correlation indicates that the estimated TSP concentration is underestimated by $m = 8.238$. The coefficient of determination $R^2 = 0.733$ measures the dispersion of the data around the line of tendency, and quantifies the mean error between each data point and the line of tendency. According to these results, $1 - R^2 = 26.7\%$ of the measured pollutant concentration cannot be explained by the model. Assuming that the model and the geographical and meteorological data are reliable, then the underestimation and dispersion of the estimated results with respect to the measured TSP concentrations could be due to inaccuracies of the emission inventories and to the existence of additional sources of pollution that were not considered in the model (e.g., existence of nearby agricultural activities, unpaved roads, or uncovered land exposed to wind erosion).

To evaluate the accuracy of each independent emission inventory, the methodology described in the previous section was implemented. The TSP emission inventory of each mine was affected by a constant. As a first step, each constant was individually varied to find a set of values that maximized $R^2$. Emissions from mines 4 and 5 were kept constant since previous studies indicated that, due to their location, their contribution to the pollution in the area of interest is negligible.

The solid red circles in Fig. 2 show the results obtained. Figure 3 shows that by affecting, one at a time, the emission inventory by a constant, there is a value for that

**Table 4** Results of the metrics to evaluate the accuracy of air quality model results when applied to the case of the open pit mining in northern Colombia in 2009, using AERMOD, Calenturitas meteorology, air quality data from nine stations and results expressed as annual mean geometric TSP concentrations

| | Base | M1 | M2 | M3 | M6 | M7 | $R^2$ max | $R^2$ max and $m = 1$ |
|---|---|---|---|---|---|---|---|---|
| $A$ | N/A | 6.65 | 0.66 | 13.29 | 6.65 | 9.97 | N/A | N/A |
| $1 - p$ | 99.7 | 99.7 | 99.7 | 99.7 | 99.7 | 99.7 | 99.7 | 99.7 |
| $m$ | 8.24 | 1.24 | 1.29 | 0.99 | 1.24 | 1.20 | 6.65 | 1.00 |
| $b$ | 42.10 | −17.68 | −19.60 | 0.98 | −17.68 | −16.66 | 48.23 | 0.00 |
| $R^2$ | 0.7336 | 0.7336 | 0.7357 | 0.7437 | 0.7336 | 0.7352 | 0.7501 | 0.7501 |
| $r$ | 0.8565 | 0.8565 | 0.8577 | 0.8624 | 0.8565 | 0.8574 | 0.8661 | 0.8661 |
| $d$ | 0.26 | 0.87 | 0.86 | 0.91 | 0.87 | 0.88 | 0.26 | 0.92 |
| Bias | −93.79 | −5.24 | −7.20 | −0.17 | −5.24 | −3.25 | −93.00 | 0.00 |
| FB | −1.74 | −0.05 | −0.07 | 0.00 | −0.05 | −0.03 | −1.71 | 0.00 |
| MG | 0.06 | 0.98 | 0.96 | 1.02 | 0.98 | 1.00 | 0.06 | 1.02 |
| NMSE | 14.15 | 0.05 | 0.06 | 0.04 | 0.05 | 0.05 | 12.47 | 0.04 |
| RMSE | 100.97 | 22.82 | 23.60 | 20.75 | 22.82 | 22.15 | 99.89 | 20.49 |
| FAC2 (%) | 0 | 100 | 100 | 100 | 100 | 100 | 0 | 100 |

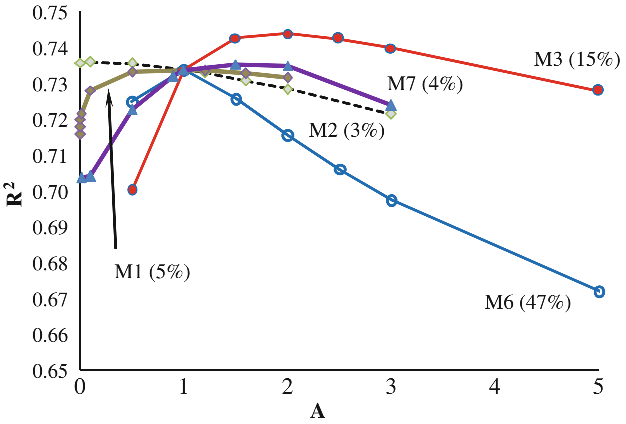

**Fig. 3** Variations of the metrics of correlation as function of Ai

constant which maximizes $R^2$ and, for all of the cases, $R^2$ decreases as the constant moves away from that point. There are no local maximums. These facts indicate that $R^2$ is a convex function with respect to these constants. Therefore, there exists a unique set of constants that will allow $R^2$ to reach its global maximum. However, that $R^2$ max is unlikely to be 1, because there will always be unpredicted or unconsidered situations affecting each monitoring station.

Figure 3 and Table 4 also show that, when the condition for $R^2$ maximum is reached, the constants affecting mines 1 and 6 are both 1. The emission inventories of these 2 mines were elaborated by the same consulting firm. Mine 6 is, by far, the largest source of emission. Thus, when the condition of maximum $R^2$ is reached, the set of constants obtained indicates the relative inaccuracy of each source relative to the dominant source of emissions.

Table 4 shows that $m = 6.65$ for the previous results. Therefore, as a second step, all of the emission sources must be affected by the same constant, in such a way that $m = 1$. This step does not change $R^2$. According to the sensitivity analysis reported by Huertas et al. (2012a), this constant should be approximately 6.65. In this step, the background concentration is also considered, such that $M_B - b \approx 0$.

Table 4 lists the set of constants that make $R^2$ maximum ($R^2 = 0.75$), $m = 1$, and $b = 0$ (when including the background concentration). Those constants specify the degree of under or overestimation of each emission source. The results show that mine 2 overestimated its emission by $1/0.66 = 1.52$ and the other mines underestimated their emissions by a factor between 6.65 and 13.29.

The air quality network uses station ZM11 to track the background concentration in the region. In 2009, this station reported an annual mean concentration of 48 μg/m$^3$. Therefore, the result of $M_B - b = 0$ indicates that the model accurately predicted the background concentration, and that the main sources of pollutants systematically affecting all of the monitoring stations were included in the model.

The result that $1 - R^2 = 25.0\%$ indicates that there are relevant circumstances that locally affect each monitoring station and that they were not included in the model. As stated before, these circumstances could be additional sources of emissions located nearby the monitoring station or calibration problems. Stations ZM9 and ZM14 exhibited the largest dispersion.

The use of $R^2$ and $m$ as metrics to evaluate accuracy of the emission inventories has a drawback: leverage data points may cause large changes in the linear regression coefficients when they are not included in the analysis. Leverage points, which occur when a monitoring system is located near a large source of emissions, can be identified using the partial leverage metric (Neter et al. 2005). When a leverage point exists, the measuring station reads high levels of pollution, which are primarily influenced by a single source. Leverage points dilute the participation of other data points and make $R^2$ insensible to variations in the emissions of other sources. In the present case, station ZM7 (which is located in the middle of the mines) had the highest pollutant concentrations. However, this station did not become a leverage point in any of the correlation analyses.

Table 4 also presents the results obtained using the metrics specified in Table 1 to evaluate air quality results. Bias and FB evaluate differences between paired predicted and measured values. By adjusting all of the emission sources by the same constant such that $m = 1$ in the linear correlation analysis, Bias and FB reach their optimal values of zero. NMSE and RMSE evaluate the dispersion of paired data; by adjusting each independent emission source by a different constant such that $R^2$ in the linear correlation analysis becomes a maximum, one can ensure that these 2 metrics reach their optimal minimum values. FAC2, MG, and $d$ are affected by the average deviation and dispersion. $R^2$ must be maximized and $m$ should be equal to 1 to ensure that these metrics reach their optimum values.

Any combination of the metrics specified in Table 1 can be used with the proposed methodology to evaluate the accuracy of the emission inventories. However, it was preferred to use linear correlation analysis, because it is commonly used in many fields, there are several ready-to-use tools to perform it, and the physical meaning of each of its parameters is well understood.

## Conclusion

In this study, it was attempted to quantify the accuracy and precision of emission inventories for cases in which the main sources of pollutants were well identified and there were air quality data from several points (>3) over more than 1 year. A backward air quality simulation approach was proposed, in which each emission inventory was weighted by a constant. A linear correlation was made between the air quality results and experimental measurements of pollutant concentration. The set of constants that maximized $R^2$ was identified and the slope of the correlation equal to 1 was rendered.

When this methodology was applied to the case of the mining region in northern Colombia, $R^2$ was determined to be an $n$-dimensional convex function with respect to the set of $n$ constants that affect each emission source; thus, there existed a unique set of constants for which $R^2$ was a global maximum. $R^2$ moved from 0.73 in the base case scenario to 0.75 in the optimum scenario. Mine 2 was found to overestimate its emission by 1/0.66, and the other mines underestimated their emissions by a factor between 6.65 and 13.29. The result of $1 - R^2 = 25.0\%$ indicated that there were relevant circumstances that locally affected each monitoring station and were not included in the air quality model. Stations ZM9 and ZM14 exhibited the largest dispersion. Although various metrics are available to evaluate differences between air quality model results and measured data, we chose to use a linear correlation analysis because of its commonality and ease of use.

**Acknowledgments** Authors acknowledge the contributions of M.Sc. Sara Vera and Natalia Navarrete to the present work.

## References

ASTM (2000) Standard guide for statistical evaluation of atmospheric dispersion model performance. American Society for Testing and Materials. D 6589-00

Axetell K, Cowherd C (1981) Improved emission factors for fugitive dust from western surface coal mining sources, vol 2. EPA Contract No. 68-0302924. US Environmental Protection Agency

Chakraborty MK, Ahmad M, Pal D, Bandopadhyay C, Chaulya SK (2002) Determination of the emission rate from various opencast mining operations. Environ Modell Softw. 17:467–480

EEA-European Environment Agency (2009) EMEP/EEA emission inventory guidebook 2009. Technical guidance to prepare national emissions inventories. Technical report No 9. ISSN 1725-2237

Ghose MK (2004) Emission factors for the quantification of dust in India coal mines. J Sci Ind Res 63:763–768

Holmes NS, Morawska L (2006) A review of dispersion modeling and its application to the dispersion of particles: an overview of different dispersion models available. Atmos Environ 40(30):5902–5928

Huertas JI, Huertas ME, Izquierdo S, González ED (2012a) Air quality impact assessment of multiple open pit coal mines in Northern Colombia. J Environ Manag 93:121–129

Huertas JI, Camacho D, Huertas ME (2012b) Standardized emissions inventory methodology for open pit mining areas. Environ Sci Pollut Res. doi: 10.1007/s11356-012-0778-3

Li L, Chen CH, Fu JS, Huang C, Streets DG, Huang HY, Zhang GF, Wang YJ, Jang CJ, Wang HL, Chen RY, Fu JM (2011) Air quality and emissions in the Yangtze River Delta. China Atmos Chem Phys 11:1621–1639

Mosca S, Graziani G, Klug W, Bellasio R, Bianconi R (1998) A statistical methodology for the evaluation of long-range dispersion models: an application to the ETEX exercise. Atmos Environ 32(24):4307–4324

Muleski GE (1990) Update of AP-42 emission factors for western surface coal mines and related sections. US Environmental Protection Agency

Neter J, Kutner M, Wasserman W, Nachtsheim C (2005) Applied linear statistical models. 5th edn. McGraw-Hill, New York

NIOSH-National Institute for Occupational Safety and Health (2005) Significant dust dispersion models for mining operations. IC 9478

Ramirez A, Keizer C, Van der Sluijs J, Olivier J, Brandes L (2008) Monte Carlo analysis of uncertainties in the Netherlands greenhouse gas emission inventory for 1990–2004. Atmos Environ 42:8263–8272

Sinha S, Banerjee SP (1997) Characterization of haul road in Indian open cast iron ore mine. Atmos Environ 31:2809–2814

USEPA (1994) Surface coal mine emission factor field study. EPA-454/R-95-010

USEPA (2004) User's guide for the AERMOD. US Environmental Protection Agency

USEPA (2005) Revision to the guideline on air quality models: adoption of a preferred general purpose (Flat and complex terrain) dispersion model and other revisions; Final Rule 40 CFR Part 51

USEPA (2006a) Revision of emission factors for AP-42. Miscellaneous source, Chap 13. Section 13.2.2: Unpaved Roads (Fugitive Dust Sources)

USEPA (2006b) Revision of emission factors for AP-42. Miscellaneous source, Chap 13. Section 13.2.4: Aggregate handling and storage piles (Fugitive dust sources)

USEPA (2008) Revision of emission factors for AP-42. Mineral products industry, Chap 11. Section 11.9: Western surface coal mining

USEPA (2010) Motor vehicle emission simulator (MOVES). User's guide

USEPA (2011) Technology transfer network. Support center for regulatory atmospheric modeling. http://www.epa.gov/ttn/scram/receptorindex.htm

Walker S, Slordal L, Guerreiro C, Gram F, Gronskei K (1999) Air pollution exposure monitoring and estimation. Part II. Model evaluation and population exposure. J Environ Monit 1:321–326

Wheeler AJ, Williams I, Beaumont RA, Manilton RS (2000) Characterization of particulate matter sampled during a study of children's personal exposure to airborne particulate matter in a UK urban environment. Environ Monit Assess 65:69–77

Zawar-Reza P, Kingham S, Pearce J (2005) Evaluation of year-long dispersion modeling of $PM_{10}$ using mesoscale model TAPM for Christchurch, NewZealand. Sci Total Environ 349:249–259

Zeller KF, Fox GF, Marlatt WE (1979) Estimating dust production from surface mining. Report No. EPA-600/7-79-182