CrossMark

Air Quality Index Revisited from a Compositional Point of View

Eusebi Jarauta-Bragulat[1] · **Carme Hervada-Sala**[2] ·
Juan José Egozcue[1]

Abstract The so-called Air Quality Index (AQI), expresses the quality of atmospheric air. The overall AQI is determined from the AQIs of some reference air pollutants, which are calculated by a transform of the respective concentrations. Concentrations of air pollutants are compositional data; they are expressed as part of mass of each pollutant in a total air volume or mass. Therefore, air pollution concentration data, as compositional data, just provide ratio information between concentrations of pollutants. Operations involved in the computation of overall AQI are not admissible operations in the framework of compositional data analysis, as they destroy the original ratio information. Consequently, the standard methodology should be reviewed for such calculations, taking into account the principles and operations of compositional data analysis. The objective of this article is to present a first approach to incorporate compositional perspective to air quality expression. For this, it is proposed to use a balance log-contrast of concentrations expressed in μg/m^3 to define a new kind of air quality indicator. Furthermore, the geometric mean of the concentrations is applied to obtain a new and simple scale air quality index, avoiding definition of piecewise linear interpolations used in the standard AQI computation. As an illustrative example, statistical analysis of atmospheric pollution data series (2004–2013) of the city of Madrid (Spain) has been carried out.

Keywords Air pollution · Air Quality Index · Compositional data analysis · Log-contrast · Balance

✉ Eusebi Jarauta-Bragulat
eusebi.jarauta@upc.edu

[1] Departament de Matemàtica Aplicada III, ETS Enginyers de Camins, Canals i Ports de Barcelona, Universitat Politècnica de Catalunya, Jordi Girona 1-3, C2, 08034 Barcelona, Spain

[2] Departament de Física i Enginyeria Nuclear, Escola d'Enginyeria de Terrassa, Universitat Politècnica de Catalunya, Colom 1, TR1, 08222 Terrassa, Spain

Springer

1 Introduction

In recent years, air quality has become a social concern due to its consequences on population's health. Nowadays, there is a clear society demand to live with reasonable environmental quality, and one of the most important environmental factors is air quality. According to The World Bank (2013), more than half of the human population is living in urban areas. In big cities, where there is a large concentration of population, a reasonable good air quality is then required. On top of this, there is a call for timely and reliable information on air quality, for good data management and measures to protect the health of population.

In this context, air quality must be defined, measured and expressed in an adequate and comprehensible way. Authorities must undertake a proper air quality management and people should be able to understand measures taken in order to keep air quality under thresholds. To quantify air quality, some indices have been defined, all related to pollutants' concentration in air. The most commonly used pollutants for AQI computation are ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), nitrogen dioxide (NO_2) and suspended particles classified by their maximum diameter ($PM_{2.5}$, PM_{10}). The most common air quality indices have been analyzed: AQI defined by the US Environmental Protection Agency (http://www.epa.gov/), the CAQI defined in the European project CITEAIR (van den Elshout et al. 2008) and the revised AQI proposed by Plaia et al. (2013). Plaia and Ruggieri (2011) review air quality indices and Bishoi et al. (2009) do a comparative study of them. Russo and Soares (2014), after a review of concepts on air pollution, introduced a predictive spatiotemporal model. In most cases, the compositional nature of air pollution data is not taken into account neither in the scientific research context nor in the guidelines for environmental authorities. This contribution is aimed at giving methodological suggestions on how compositional data principles may be applied in future developments.

Air quality indices are commonly based on air pollutant concentrations and, thus, functions transforming air pollutant concentrations into indexes are required. Most of air quality index computing methods use a similar algorithm involving piecewise linear functions, which transforms concentrations into normalized indexes. The intervals are defined by breakpoints which are chosen to fit both observed and subjectively evaluated health impacts. The piecewise linear function in this case is

$$
\begin{aligned}
\mathrm{AQI}(P) &= \frac{\mathrm{AQI_{HI}} - \mathrm{AQI_{LO}}}{\mathrm{BP_{HI}} - \mathrm{BP_{LO}}} \left(C(P) - \mathrm{BP_{LO}}\right) + \mathrm{AQI_{LO}}, \\
\mathrm{BP_{LO}} &\leq C(P) < \mathrm{BP_{HI}},
\end{aligned}
\tag{1}
$$

where $\mathrm{AQI}(P)$ is the AQI value for pollutant P; $C(P)$ is the average pollutant P concentration; $\mathrm{BP_{HI}}$ is the breakpoint that is greater than $C(P)$; $\mathrm{BP_{LO}}$ is the breakpoint that is less than or equal to $C(P)$; $\mathrm{AQI_{HI}}$ is the AQI value corresponding to $\mathrm{BP_{HI}}$ and $\mathrm{AQI_{LO}}$ is the AQI value corresponding to $\mathrm{BP_{LO}}$. Such breakpoints are specified by pollutant and differ depending on the country. Air quality index values are divided into different ranges, and each one is assigned a descriptor and a color code (Table 1; Fig. 1).

Table 1 AQI breakpoints and corresponding color codes

	AQI		
	Low	High	
Good	0	50	
Moderate	51	100	
Unhealthy 1	101	150	
Unhealthy 2	151	200	
Very unhealthy	201	300	
Hazardous	301	400	
Very hazardous	401	500	

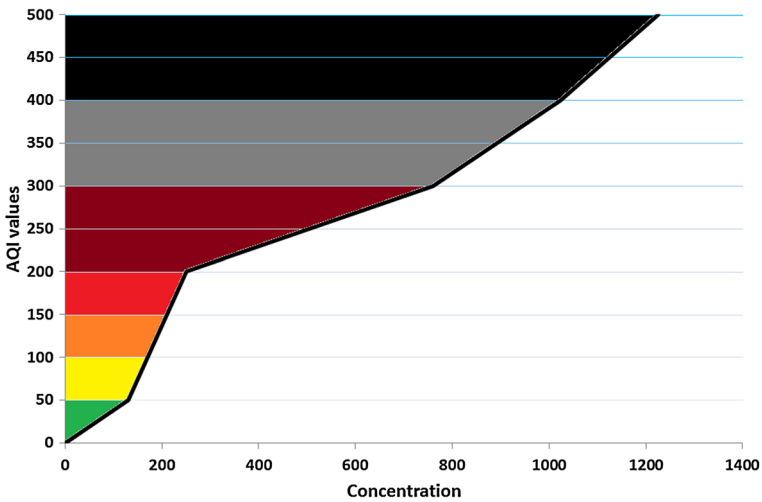**Fig. 1** Piecewise linear function assigning AQI values to concentration of pollutant ($\mu g/m^3$)

Standardized public health alert protocols are associated with each AQI range. The presence of pollutants in air is expressed by their concentration, usually in $\mu g/m^3$ and sometimes in parts per million of mass (ppm) or in parts per billion of mass (ppb). Daily and monthly averages of concentrations are carried out using the arithmetic mean of the corresponding measured values for each element and over the corresponding period of time. These average values are used to assign the corresponding AQI values for each pollutant. The final air quality report usually contains the maximum AQI of different air pollutants, the average of the two greatest ones and the arithmetic mean (average) of all AQIs, the so-called overall AQI. As a comment, note that AQI indexes express the degree of pollution and, strictly speaking, not the quality of the air. Certainly, the larger the AQI, which means the larger the pollution, the worse does the air quality get. Nevertheless, the term "quality" is maintained for the standard air quality indices.

Concentrations of air pollutants are compositional data. They describe the parts in which the whole amount (volume or mass) of air is decomposed. This means that the information provided by this kind of data comes from the ratios between different parts. The analysis of compositional data must be done according to some principles (Aitchison 1986; Aitchison and Egozcue 2005; Egozcue 2009). For instance, scale invariance principle requires that a change of units in concentrations, for example, from proportions to ppm, does not change the result of the analysis; subcompositional coherence principle preconizes that results obtained from the analysis of a composition cannot be contradictory with those obtained from the analysis of a subcomposition. Violation of some of these principles may lead to inconsistent results. Pearson (1897), one of the founders of modern statistics, who pointed out the spurious correlation phenomenon, did the very first reference of such problems.

In the eighties, Aitchison (1986) put forward the so-called log-ratio approach, defining suitable techniques to deal with compositional data. From then on, there have been significant advances in the formal aspects of the analysis that allow greater systematization of his methods. To sum it up, statistical analysis of compositional data can be stated as a three-step process: (1) transforming data in log-ratio coordinates; (2) performing statistical analysis of those coordinates as real variables; (3) interpreting the results on the coordinates or back transforming them into compositional data. However, the usage of this process is not as widespread as it deserves (Egozcue and Pawlowsky-Glahn 2011b).

The main goal of this paper is to contribute to improve analysis, interpretation and management of air quality using compositional data techniques. Particularly, the interest is focused on the formulation of a new air quality index that satisfies the requisites of compositional data analysis. Other advanced aspects of the analysis of air quality are out of the scope of this contribution. For instance, evolution in space using geostatistics (Russo and Soares 2014) or applying differential equations in the simplex (Egozcue and Jarauta-Bragulat 2014). Further research should be necessary to get a complete, coherent and comprehensive air quality approach, considering the perspective of compositional data analysis.

2 Compositional Approach

The methodological aspects will be demonstrated using a data set granted by the Ayuntamiento de Madrid (Spain). The monthly mean concentrations of air pollutants O_3, $PM_{2.5}$, PM_{10}, CO, SO_2, NO_2 are taken from the years 2004 to 2013, constituting a sample of $n = 117$ monthly data points.

The first step in the analysis was to compare averages of pollutant concentrations and averages of corresponding AQIs. Note that the NO_2 data values are removed for this comparison because this pollutant, for low concentrations, gives null values of the corresponding AQI. As the functions relating concentrations to AQIs are piecewise linear functions of the pollutant concentrations (Eq. 1), it is reasonable to expect a good correlation between the arithmetic average of pollutant concentrations and arithmetic average of AQIs, which is the overall AQI.

Figure 2a shows time evolution of the arithmetic mean of pollutant concentrations and the arithmetic mean of pollutant AQIs. Figure 2b shows the scatterplot of pollutant concentration arithmetic average versus AQI arithmetic average, with the correlation coefficient for these two variables being 0.579, which is a value smaller than expected. Moreover, the shape of the two curves in Fig. 2a is different.
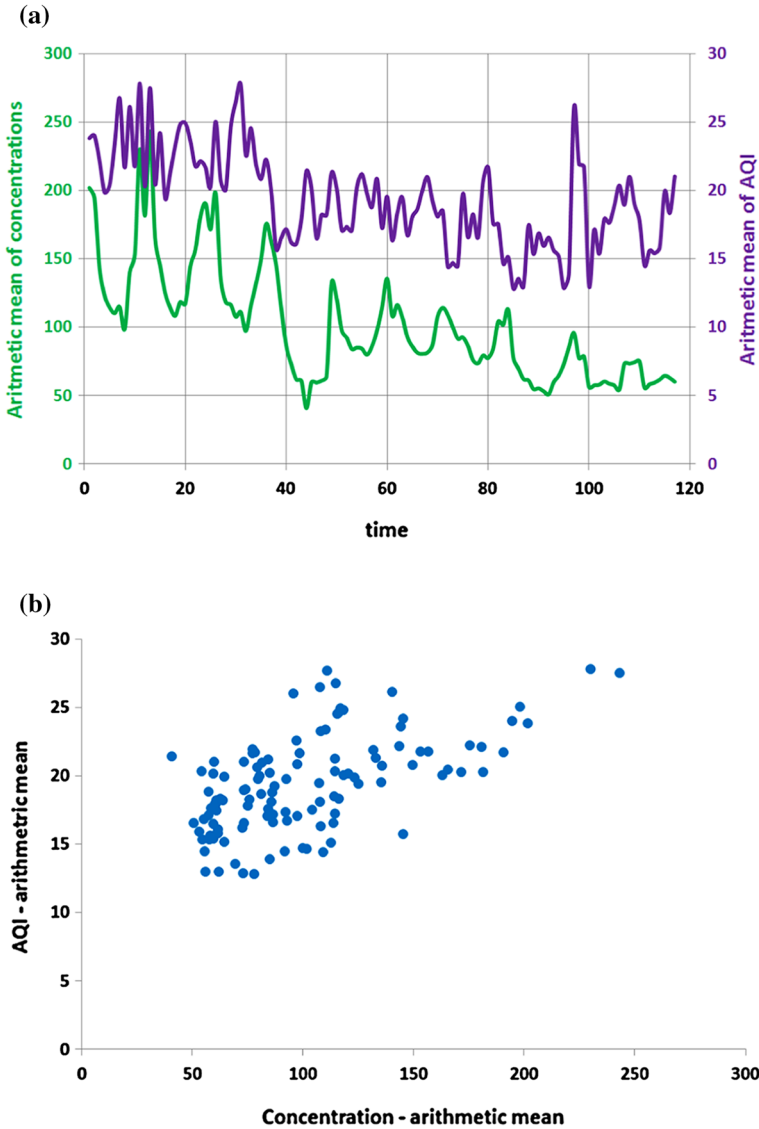

Fig. 2 **a** Pollutant concentration ($\mu g/m^3$) arithmetic mean (*green*) and AQI arithmetic mean (*purple*) as a function of time (months). **b** Scatterplot of pollutant concentration ($\mu g/m^3$) arithmetic mean versus AQI arithmetic mean. Data set Madrid 2004–2013

Figure 3a shows time evolution of the geometric mean of pollutant concentrations and the geometric mean of pollutant AQIs. Figure 3b shows the scatterplot of pollutant concentration geometric mean versus AQI geometric mean. The correlation coefficient for these two variables is 0.999, and the shape of the two curves fits quite well. These examples, presented in Fig. 3a, b, suggest that geometric mean is a better centrality estimator than arithmetic mean when dealing with concentrations for this data set.
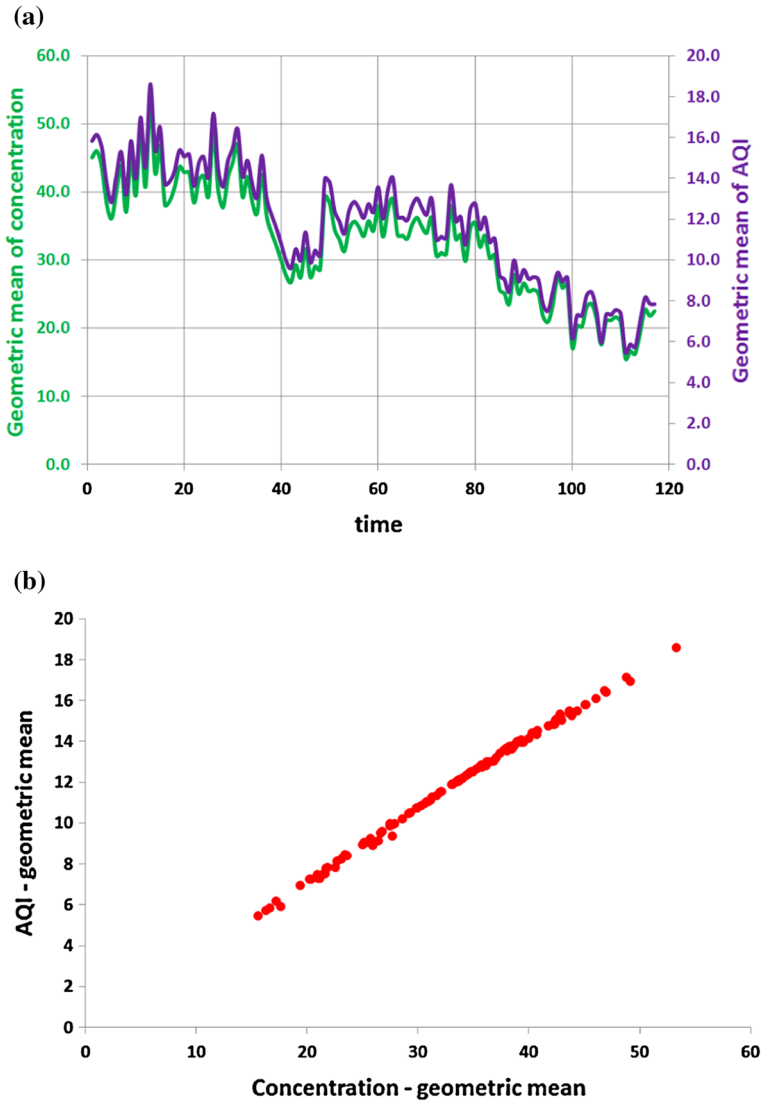

Fig. 3 **a** Pollutant concentration ($\mu g/m^3$) geometric mean (*green*) and AQI geometric mean (*purple*) as a function of time (months). **b** Scatterplot of pollutant concentration ($\mu g/m^3$) geometric mean versus AQI geometric mean. Data set Madrid 2004–2013

Table 2 Correlation matrices of air pollutants in Madrid showing the effect of spurious correlation

(a)	O_3	$PM_{2.5}$	PM_{10}	CO	SO_2
O_3	1				
$PM_{2.5}$	-0.361	1			
PM_{10}	-0.330	0.179	1		
CO	-0.714	0.344	0.583	1	
SO_2	-0.641	0.060	0.605	0.814	1

(b)	O_3	$PM_{2.5}$	PM_{10}	CO	SO_2
O_3	1				
$PM_{2.5}$	0.443	1			
PM_{10}	0.353	0.073	1		
CO	-0.924	-0.569	-0.624	1	
SO_2	-0.164	-0.379	0.486	0.003	1

(a) When concentrations of pollutants are expressed in ppb of mass; (b) when concentrations of pollutants are expressed of proportions in mass in the subcomposition O_3, $PM_{2.5}$, PM_{10}, CO, SO_2

This is in accordance with well-known facts in compositional data analysis (Aitchison 1986; Aitchison and Egozcue 2005).

The concentrations of air pollutants constitute a compositional vector, a composition for short. Compositions represent parts of a whole, which, in this case, are total micrograms in a cubic meter of air. The information conveyed by a composition is summarized in the ratios between different parts. Therefore, when multiplying a composition by a constant, information does not change and the two vectors are equivalent as compositions. For instance, expressing air pollutants in $\mu g/m^3$ is equivalent to expressing them in other units, for instance in g/m^3.

Another characteristic of compositions is the subcompositional coherence: analysis made in a subset of parts of the original composition should not produce different results referred to common parts. For instance, the analysis without taking into account NO_2 must be coherent with the one obtained including NO_2. An example of violation of this principle is shown in Table 2 which shows two correlation matrices of air pollution data set computed on the whole composition and computed on the subcomposition containing the pollutants O_3, $PM_{2.5}$, PM_{10}, CO, SO_2. See that correlation CO-PM_{10} in case (a) is 0.583, meaning that when CO increases PM_{10} also increases. In case (b) the correlation is -0.624, meaning that when CO increases PM_{10} decreases, an incoherent interpretation.

A proper compositional data analysis should be performed through functions of concentrations that are invariant under change of units or multiplication by constants (Egozcue 2009). On the other hand, the scale of ratios is commonly treated taking logarithms, as proposed in Aitchison (1986). Compositional data can be properly analyzed and represented using the expressions called log-contrasts. For a composition $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, a log-contrast (Egozcue and Pawlowsky-Glahn 2011a) is a linear combination of logarithms of parts where the sum of the coefficients is zero, that is

$$\sum_{i=1}^{n} a_i \log(x_i) = \log\left(\prod_{i=1}^{n} x_i^{a_i}\right), \quad \sum_{i=1}^{n} a_i = 0, \tag{2}$$

which is a scale invariant function. Some of the a_i are positive, call them b_i, and the rest are negative, call them $-c_i$, or null, call them $-d_i$. Then, Eq. (2) can be written as

$$\log\left(\prod_{i=1}^{n} x_i^{a_i}\right) = \log\left(\prod_{j=1}^{p} x_j^{b_j} \prod_{k=p+1}^{n} x_k^{-c_k}\right) = \log\left(\frac{x_1^{b_1} \cdots x_p^{b_p}}{x_{p+1}^{c_{p+1}} \cdots x_n^{c_n}}\right),$$
$$b_1 + \cdots + b_p = c_{p+1} + \cdots + c_n.$$

A very important and interesting property of any log-contrast is that, if a change of units is applied to one or more components of the numerator or denominator, the only resulting change in the log-contrast is the addition of a known constant. That is

$$\log\left(\frac{(\alpha x_1)^{b_1} \cdots x_p^{b_p}}{(\beta x_{p+1})^{c_{p+1}} \cdots x_n^{c_n}}\right) = \log\left(\frac{\alpha^{b_1}}{\beta^{c_{p+1}}}\right) + \log\left(\frac{x_1^{b_1} \cdots x_p^{b_p}}{x_{p+1}^{c_{p+1}} \cdots x_n^{c_n}}\right).$$

When working with given air pollutant concentrations, the original composition is made of concentrations of pollutants and a fill-up value, consisting of other air components. If they are expressed in ppb, they add up to 10^9, as it was done in Table 2a. Calling the fill-up value x_n, the log-contrast

$$\log\left(\frac{x_n^{b_n}}{x_1^{c_1} \cdots x_{n-1}^{c_{n-1}}}\right), \tag{3}$$

can be used as a index of air pollution or, as in the AQI terminology, an air quality index for given coefficients b_n, c_i, $i = 1, 2, \ldots, n-1$ satisfying $b_n = \sum c_i$. These coefficients can be used to weight the impact of each considered pollutant on population health. As a first approach, the log-contrast

$$b_{AQ}(\mathbf{x}) = \log\left(\frac{x_n}{x_1^{1/(n-1)} \cdots x_{n-1}^{1/(n-1)}}\right) = \log\left(\frac{x_n}{g_{n-1}(\hat{\mathbf{x}})}\right)$$
$$= \log(x_n) - \log(g_{n-1}(\hat{\mathbf{x}})), \quad \hat{\mathbf{x}} = (x_1, x_2, \ldots, x_{n-1}), \tag{4}$$

can be taken as an index of air quality, denoted b_{AQ}, as it is a balance (Egozcue and Pawlowsky-Glahn 2005) between non-polluted air and considered air pollutants equally weighted. In Eq. (4), $g_{n-1}(\cdot)$ is the geometric mean of $n-1$ first components, in the example $n - 1 = 6$ pollutants. Note that b_{AQ} is dimensionless, that is, it does not depend on the units in which concentrations are expressed. However, both x_n and $g_{n-1}(\hat{\mathbf{x}})$ have the units of concentrations, for instance, $\mu g/m^3$, ppm or ppb. In practice, $\log(x_n)$ is almost constant across the sample, and therefore, b_{AQ} can be approached as $-\log g_{n-1}(\hat{\mathbf{x}})$ up to an additive constant. The quantity $-\log g_{n-1}(\hat{\mathbf{x}})$ is not dimensionless and is only useful to compare different pollutant concentration values, as the additive constant cancels out. Figure 4 shows how b_{AQ} and $-\log g_{n-1}(\hat{\mathbf{x}})$ differ approximately in a constant term.
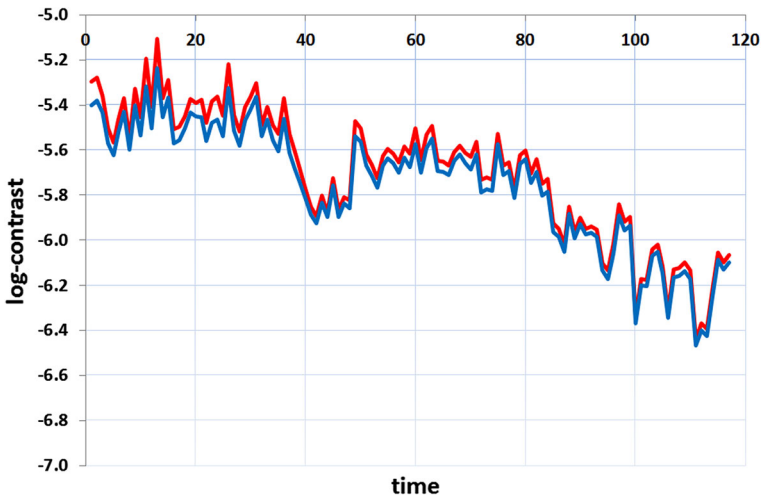
Fig. 4 Log-contrast b_{AQ} (*blue*) and $\log(10{,}000) - \log g_{n-1}(\hat{\mathbf{x}})$ (*red*). Small and almost constant differences are shown between the two curves. Data set Madrid 2004–2013

From the point of view of compositional data analysis, the use of log-contrasts is much more suitable and coherent than using concentrations or even log-concentrations. On top of this, air quality indexes can be redefined using the log-contrast b_{AQ} or its approximation by geometric mean of pollutant concentrations. The main advantage of this approach is that b_{AQ} is dimensionless and does not depend on the units of concentrations. Moreover, changing units of concentrations, modifying the value of $-\log g_{n-1}(\hat{\mathbf{x}})$ gets even adding a constant.

3 A New Kind of Air Quality Indexes

The common practice in air quality evaluation consists of computing AQI values for each pollutant and these values are then combined to obtain a global AQI: for instance, the arithmetic average of all pollutant-AQI, the maximum pollutant-AQI and the average of the two maximum pollutant-AQI values. Some authors (Bruno and Cocchi 2002; Plaia et al. 2013) have suggested defining the global air quality indexes as functions of pollutant concentrations and not through the pollutant AQIs. In order to define a new kind of global air quality index directly based on the air pollutant concentrations, the log-contrast b_{AQ} is a natural choice from the compositional point of view, once the pollutants are given. However, the fill-up value on the composition (non-polluted air) is almost never reported and so b_{AQ} is not easily computed. Instead, it is proposed to use the geometric mean of pollutant concentrations $g_{n-1}(\hat{\mathbf{x}})$. This new kind of air quality index, here denoted AQI*, is proposed to be proportional to that geometric mean of pollutant concentrations, that is

$$\text{AQI}^*(\mathbf{x}) = k^* g_{n-1}(\hat{\mathbf{x}}), \tag{5}$$

where the coefficient k^* value depends on the units of concentrations and also of the considered pollutants. Once units of concentration are fixed, for example $\mu g/m^3$ in this case, the value of the constant k^* also depends on the desired scale for AQI values.

Table 3 shows the breakpoint values used to compute the US Environmental Protection Agency index. It also shows the AQI values corresponding to breakpoints of each pollutant, the geometric mean of pollutant concentrations (NO_2 excluded) and the proposed new AQI*. In order to give a reasonable value to k^*, Table 3 is revisited. The AQIs are computed for each pollutant, and the maximum AQI, called conditional AQI, is the one corresponding to the so-called conditional pollutant. It results that the conditional AQI is not sensitive to those pollutant AQIs that are not the maximum one. For instance a PM_{10} concentration of $154.0\,\mu g/m^3$ (Table 3, second row) and values for other pollutants, as in the breakpoints in the first row of Table 3, correspond to a conditional AQI of 100 and the conditional pollutant is PM_{10}. However, when all pollutants have concentrations as specified in the second row of Table 3, the conditional AQI is still 100, although the quality of the air has decreased considerably. This undesirable behavior of the conditional AQI is compensated by considering two additional indexes: the average of the two maximum pollutant AQIs and the average of all (positive) single pollutant AQIs. As discussed previously the AQI (arithmetic) average has an unclear relation with the arithmetic mean of pollutant concentrations (Fig. 2) and represents a poor subjective correction to the conditional AQI. Similar arguments are valid for the average of the two worst pollutant AQIs.

The seventh row of Table 3, corresponding to 500 AQI values, gives concentrations of pollutants that, when appearing simultaneously, correspond to an almost incredible pollution of the air. Let us take these concentrations in $\mu g/m^3$, and their geometric mean $2268.7\,\mu g/m^3$, as a reference of 100 for a suitable AQI*. Accordingly, if concentrations of pollutants are expressed in $\mu g/m^3$, a choice of an index AQI* ranging from 0 to 100 requires the choice of $k^* = 100/2268.7$, so that

$$\text{AQI}^*(\mathbf{x}) = \frac{100}{2268.7} g_{n-1}(\hat{\mathbf{x}}). \tag{6}$$

Table 3 Columns from O_3 to NO_2, concentration pollutants breakpoints ($\mu g/m^3$)

O_3	$PM_{2.5}$	PM_{10}	CO	SO_2	NO_2	AQI	GM	AQI*	
129.9	15.4	54.0	5210.4	92.0	0.0	50	139.0	6	
170.5	40.4	154.0	11131.4	389.8	0.0	100	340.9	15	
211.1	65.4	254.0	14684.0	606.3	0.0	150	499.9	22	
251.7	150.4	354.0	18236.6	822.8	1044.6	200	725.6	32	
759.2	250.4	424.0	35999.4	1634.9	1992.8	300	1365.3	60	
1023.1	350.4	504.0	47841.4	2176.2	2635.7	400	1798.4	79	
1226.1	500.4	604.0	59683.3	2717.6	3278.5	500	2268.7	100	

Traditional AQI values corresponding to breakpoints. *GM* geometric mean of pollutant concentrations excluding NO_2. AQI*, scaled from 0 to 100
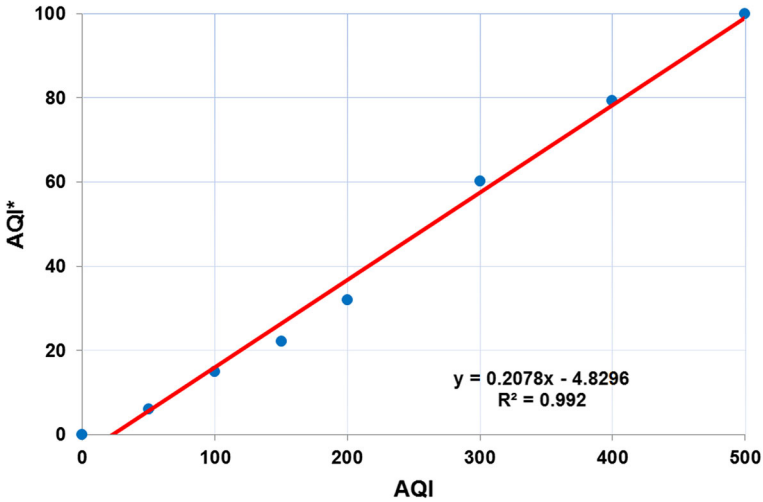
Fig. 5 AQI* values vs AQI values at the traditional breakpoints (*blue dots*). Fitted line (*red*)
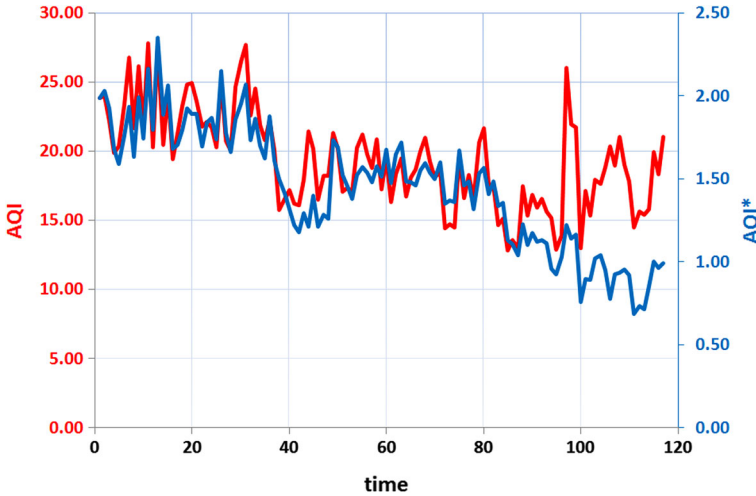

Fig. 6 AQI values (*red*) and AQI* values (*blue*) in time (months). Data set Madrid 2004–2013

Note that NO_2 was not used for assessing the scale of AQI*, but the value 2268.7 μg/m^3 including NO_2 in the geometric mean is an almost incredible air pollution situation.

Equation (5) is maintained as the definition of AQI* although the geometric mean includes NO_2. Figure 6 shows a comparison between AQI and AQI* values computed for the reference data set (Madrid, 2004–2013); a secondary scale is used, as the corresponding scales are different although they are almost linearly related (Fig. 5). Figure 6 also shows that air pollution in Madrid was in general very low (AQI roughly ranges from 10 to 30 and AQI* from 0.7 to 2.5). The AQI* ranging from 0 to approximately 100 is proposed, so that the AQI* is easily interpreted. For instance, for the Madrid data set it is said that the AQI* values are in the interval (0.7,2.5) pollution over 100.

4 Conclusions

Measures of air quality are based on concentrations of some given pollutants and concentrations of suspended particles classified by size. Those concentrations are compositional data and, consequently, they are properly analyzed using log-ratio approaches. These analyses are based on log-contrasts, as they are scale invariant.

The log-contrast between concentration of non-pollutant components over pollutant components is taken as a proper dimensionless index of air pollution. A good approximation of this index is the logarithm of the geometric mean of concentrations of air pollutants. The scale of this geometric mean is comparable to the traditional AQI scale.

A first approach to a new air quality index ranging from 0 to 100 is proposed as $AQI^*(\mathbf{x}) = k^* g_{n-1}(\hat{\mathbf{x}})$, where k^* is a constant depending on the units of pollutant concentrations and $g_{n-1}(\cdot)$ is the geometric mean of air pollutants concentrations.

Further modifications of the proposed index would require to determine weighting coefficients of the proposed log-contrast according to health indicators.

References

Aitchison J (1986) The statistical analysis of compositional data. Monographs on statistics and applied probability. Chapman & Hall Ltd., London

Aitchison J, Egozcue JJ (2005) Compositional data analysis: where are we and where should we be heading? Math Geol 37(7):829–850

Bishoi B, Prakash A, Jain V (2009) A comparative study of air quality index based on factor analysis and US-EPA methods for an urban environment. Aerosol Air Qual Res 9:1–17

Bruno F, Cocchi D (2002) A unified strategy for building simple air quality indices. Environmetrics 13:243–261

Egozcue JJ (2009) Reply to "On the Harker variation diagrams;.." by J. A. Cortés. Math Geosci 41(7):829–834

Egozcue JJ, Jarauta-Bragulat E (2014) Differential models for evolutionary compositions. Math Geosci 46(4):381–410

Egozcue JJ, Pawlowsky-Glahn V (2005) Groups of parts and their balances in compositional data analysis. Math Geol 37(7):795–828

Egozcue JJ, Pawlowsky-Glahn V (2011a) Análisis composicional de datos en ciencias geoambientales. Boletín Geológico y Minero 122(4):439–452

Egozcue JJ, Pawlowsky-Glahn V (2011b) Basic concepts and procedures. In: Pawlowsky-Glahn V, Buccianti A (eds) Compositional data analysis. Theory and applications. Wiley, New York

Pearson K (1897) Mathematical contributions to the theory of evolution on a form of spurious correlation which may arise when indices are used in the measurement of organs. In: Proceedings of the Royal Society of London LX, pp 489–502

Plaia A, Di Salvo F, Ruggieri M, Agró G (2013) A multisite-multipollutant air quality index. Atmos Environ 70:387–391

Plaia A, Ruggieri M (2011) Air quality indices: a review. Rev Environ Sci Biotechnol 10:165–179

Russo A, Soares A (2014) Hybrid model for urban air pollution forecasting: a stochastic spatio-temporal approach. Math Geosci 46:75–93

The World Bank (2013) Report. on line. data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS
van den Elshout S, Léger K, Nussio F (2008) Comparing urban air quality in europe in real time. A review of
    existing air quality indices and the proposal of a common alternative. DCMR Environ Int 34:720–726