Manuscript

Click here to view linked References

Click here to access/download;Manuscript;Manuscript_ACM_rev_MI_IM_cle

Deconstruction of annoyance due to air pollution by multiple correspondence analyses

Milena Machado[a]*, Jane Meri Santos[b], Severine Frere[c], Phillipe Chagnon[c], Valdério Anselmo Reisen[d], Pascal Bondon[f], Márton Ispány[g], Ilias Mavroidis[d], Neyval Costa Reis Junior[b]

[a] Instituto Federal de Ciência e Tecnologia do Espírito Santo, Guarapari, Brazil

[b] Department of Environmental Engineering, Universidade Federal do Espírito Santo, Vitoria, Brazil

[c] Université du Littoral Côte d'Opale, Maison de la Recherche em Science de l´Homme, Dunkerque, France

[d] Hellenic Open University, School of Science and Technology, Greece

[d] Department of Statistics, Universidade Federal do Espírito Santo, Vitoria, Brazil

[f] Laboratoire des Signaux et Systems (L2S), CNRS-Centrale Supélec-Université Paris-Sud, Gif sur Yvette, France

[g] University of Debrecen, Debrecen, Hungary

* Tel: +55 (27) *988527717*, e-mail: milenamm@ifes.edu.br

Abstract

Annoyance caused by air pollution is a matter of public health as it can cause stress and ill-health and affect quality of life, among other burdens. The aim of this study is to apply the multiple correspondence analyses (MCA) technique as a differential tooling to explore relationships between variables that can influence peoples' behaviour concerning annoyance caused by air pollution. Data were collected through a survey on air pollution, environmental issues and quality of life. Face-to-face survey studies were conducted in two industrialized urban areas (Vitoria in Brazil and Dunkirk in France). These two regions were chosen as their inhabitants often report feeling annoyed by air pollution and both regions have similar industrial characteristics. The results showed a progressive correspondence between levels of annoyance and other active variables in the "air pollution" factor group: as the levels of annoyance increased, the levels of the other qualitative variables (importance of air quality, perceived exposure to industrial risk, assessment of air quality, perceived air pollution) also increased. Respondents who reported feeling annoyed by air pollution also thought that air quality was very important and were very concerned about exposure to industrial risks. Furthermore, they often assessed air quality as horrible and they could frequently perceive air pollution by dust, odours and decreased visibility. The results also showed a statistically significant association between occurrence of allergies and high levels of annoyance.

Keywords: *air pollution, perceived annoyance, behaviour, health impacts, Multiple Correspondence Analysis.*

1. Introduction

Air pollution can cause various health impacts, such as respiratory and cardiovascular diseases, hypertension and cancer among others (WHO, 2014; Lercher et al., 1995; Llop et al., 2008; Nascimento et al., 2017; Souza et al., 2018). In addition to the direct health effects caused by the exposure to pollutants, air pollution may also cause annoyance (Amundsen et al., 2008; Machado et al., 2018; Orru et al., 2018b). The concept of annoyance can be complex and extremely subjective as it can be experienced as a perception, an emotion, an attitude or a mixture of these (Berglund, Berglund & Lindvall, 1987). Lindvall and Radford (1973) defined annoyance as "a feeling of displeasure associated with any agent or condition known or believed by individuals or groups to adversely affect them". Annoyance can be associated with negative emotions, such as anger, disappointment, dissatisfaction, helplessness, anxiety and agitation, and behavioural/social changes, such as interference with intended activities (Blanes-Vidal et al., 2012). Furthermore, the World Health Organization (WHO, 2005) defines health as a "state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". Therefore, annoyance caused by air pollution can be considered as a health problem and an ambient stressor that affects the quality of life.

Although a significant number of studies linking air pollution and human health risks exist in the literature (for example Oglesby et al., 2000; Llop et al., 2008; Stenlund et al., 2009; Egondi et al., 2013; Orru et al., 2018), comparatively, there are few studies exploring perceived annoyance caused by air pollution as a health risk and its influencing parameters. According to Jacquemin et al. (2007), there are factors or groups of qualitative variables that can be determinants of the perceived annoyance, such as sociodemographic factors, health symptoms, location of residence and perception of levels of particulate matter.

One way to better understand the relationship among qualitative variables related to annoyance is through the technique of correspondence analysis. Multiple Correspondence Analysis (MCA) is a multivariate analysis technique for categorical data that allows to graphically assess the differences, similarities and relationships between variables and their response categories (Benzécri et al., 1973; Greenacre & Blasius, 2006).

The objective of the present work is to analyse and compare the relationships between qualitative variables determining the nuisance caused by air pollution. Although the MCA technique has not been applied before in relation to air pollution, it is a well-valued technique for exploring qualitative variables related to annoyance caused by air pollution and health risks. The MCA is developed here to evaluate the datasets from two surveys conducted in different urban and industrialized regions, namely Dunkirk (France) and Vitoria (Brazil), which have many characteristics in common.

2. Materials and methods

2.1. Characteristics of the regions

The study was conducted in two distinct urban industrialized regions located in coastal, port and industrial areas (Figure 1 and Figure 2): Dunkirk (France) and Vitoria (Brazil). This allowed for the comparison of the annoyance levels observed in two regions with

similar characteristics and provided further insight into the relationship between annoyance and air pollution effects.

Despite the geographic and socioeconomic differences between these regions, both are exposed to high levels of air pollution and their inhabitants often report annoyance due to air pollution to their local authorities. According to a report concerning the industrial risk perception in Dunkirk (Calvo-Mendieta et al., 2008, Hellequin et al, 2010), air pollution is cited as the first environmental problem by its inhabitants, followed by water and soil pollution. In Vitoria, according to the local environmental agency more than 24% of the complaints refer to air pollution (Machado et al., 2018). In 2009, Dunkirk and Vitoria signed an international cooperation agreement to develop a variety of events, organizing projects in the realm of environment, culture, economy, port activities, urban development and university research (Les ateliers, 2010).

2.1.1. Dunkirk

The metropolitan region of Dunkirk (MRD) has about 210,000 inhabitants, is the third largest port in France and an industrialized region containing steel, food, pharmaceutical and chemical industries, an oil refinery and also a nuclear power station for electricity production (risk's source). The climate is oceanic, that is strongly influenced by the wind: summer breeze sometimes contributes to increase sunshine hours, but there are also episodes of "squalls" accompanied by penetrating winter rains (Boyouk et al., 2011; Salvador et al., 2016). The anthropogenic activities present potential sources of particulate matter, which is the main cause of complaints by the resident population in this region (PPA, 2002). Measurement of air pollutant concentrations are being carried out by air quality monitoring stations distributed according to the French national guidelines (ADEME, 2002), which implement European Union Directives 96/62/EC and 99/30/EC.

According to the Protection Plan of the Atmosphere (PPA, 2002), the emission inventory of MRD clearly shows that the industrial sector is the largest emitter of pollutants. For particulate matter (PM_{10}) the main sources in the region are industries, incineration plants, collective and individual heating, followed by road transport. In PPA (2002) there are measures established for local industries, to prevent and reduce the dispersion of particles, such as spraying water on stock piles, watering paths and storage areas, as well as changing the conditions of discharge.

2.1.2. Vitoria

The metropolitan region of Vitoria (MRV) has about 1,500,000 inhabitants (IBGE, 2010) and is located on the south-eastern coast of Brazil (Figure 2). The MRV climate is classified as tropical, hot and humid, characterized by long summers (usually October to April) and high temperatures, with maximum temperatures occurring usually in December and January. Winter is weak, with average temperature of the coldest month about 18°C, the cold sensation existing occasionally related to the occurrence of cold fronts. The prevailing wind direction is north-easterly which contributes to the dispersion of pollutants emitted from the main large industries in the area towards the city (Andreão et al., 2019).

The MRV comprises a large port system, heavy vehicular traffic and an industrial park that includes, among others, steel production, pelletizing, quarry, cement and food

industries, chemical industries and an asphalt plant, that are all potential sources of air pollution (Santos et al., 2017). In recent years, the region of Vitoria has experienced a process of economic growth and increased industrial production as well as urban development (IJSN, 2015). To monitor the air quality in the Vitoria region, eight air quality monitoring stations, set in different locations in MRV, are managed by the state environmental agency. According to Santos et al., (2017) the major contributor sources of total particles in Vitoria are resuspension due to vehicular traffic in paved and unpaved roads, followed by industrial emissions (mainly the pelletizing and steel industries) and vehicles exhausts. Machado et al. (2018) found that most of the complaints about air pollution are related to nuisance from particles.

2.2. Measured air quality levels

Recent studied have shown that fine (≤2.5 μm, $PM_{2.5}$) and ultrafine particles (≤0.1 μm, $PM_{0.1}$) are associated with diseases in the lower respiratory system (Farfel et al., 2005, Souza et al., 2017). Inhalable particles (≤10 μm, PM_{10}), that also include fine and ultrafine particles, also pose severe public health concerns, such as upper respiratory system disturbances and in some cases persistent nuisance (Vallack and Shillito, 1998; Amundsen et al., 2008; Blanes- Vidal et al., 2012; Machado et al., 2019).

Table 1 shows some descriptive statistics of 24-hour average PM_{10} concentrations measured at the air quality stations in Dunkirk during the year 2008 (Atmo Nord-Pas-de-Calais, 2009). As can be seen from this table, the mean concentration levels differ significantly among stations. The largest standard deviations were observed for the concentrations measured in the stations: Fort Mardyck, St Pol Mer Nord, Grande-Synthe and Mardyck. This is due to the large variability between the minimum and maximum concentration values registered in these stations. The 24-hour average PM_{10} concentrations surpassed in several occasions the respective WHO annual air quality guideline, which is 50 μg/m³ (24h mean) (WHO, 2005).

Table 2 presents the descriptive statistics of 24-hour average PM_{10} concentrations measured during 2011 at the air quality monitoring stations located in MRV, except at Vila Velha-Centre station, which in 2011 did not register enough data (min70%) for analysis. As can be seen, the largest standard deviation (variability between the minimum and maximum concentration values) are measured in the stations: Laranjeiras, Vila Velha-Ines and Cariacica. At all air quality stations, the 24-hour mean is higher than the WHO annual air quality guideline for PM_{10} (WHO, 2005). As in the case of Dunkirk, the PM_{10} data show the occurrence of high concentration peaks during the year.

2.3.The surveys

A survey was conducted at the MRD in 2008 with a representative sample of 518 people (over 18 years old) interviewed using face-to-face questionnaires. The respondents were chosen at random at urban communities grouped into 10 sub-regions: Bourbourg, BrayDunes/Leffrinckoucke, Téteghem, Coudekerque Branche, Gravelines, St Pol sur Mer, Grand Synthe, Petite-Synthe, Dunkirk and Malo/Rosendael. The sample size was proportionally distributed according to gender, geographic location (near or far from industries/port area), and socio-professional category.

A separate survey was conducted in the MRV in 2011; the sample size was determined by using a simple random sampling with proportional allocation method (Cochran, 1977) totalling a representative sample of 515 respondents (over 16 years old), which were distributed proportionally in the sub-regions around the eight air quality monitoring station: Laranjeiras, Carapina, Ibes, Jardim Camburi, Vitoria-Centre, Enseada do Sua, Cariacica and Vila Velha-Centre. It should be noted that the locations of the eight monitoring stations in the MRV were originally planned according to the population density and the proximity to the main industrial sources, vehicular sources and port area.

The questionnaire used in the MRD in 2008 was adapted and applied in the MRV region in 2011. A piloting questionnaire was used for a pre-test to ensure stability over time and internal consistency. The main questions selected, for the analysis in this study, are displayed in Table 3.

The surveys contained questions concerning socioeconomic and demographic factors such as age, level of education, occupation, habits, gender and location of residence (see Table 3). These were closed-ended questions providing answer options based on a Likert scale. To quantify nuisance, categorical scales were prepared and applied according to the context of each question (for example, "*Do you feel annoyed by air pollution?*", with the categorical answer options: *not annoyed, slightly annoyed, very annoyed, extremely annoyed* and "*do not know*".

Categorical scales are often applied in studies on annoyance caused by particles, odour and noise, such as in the studies conducted by Passchier & Passchier (2000) and Klaeboe et al. (2000), for instance. Atari et al. (2009) adopted five-point categorical and numerical scales to measure annoyance caused by odour. Blanes-Vidal et al. (2012) adopted a qualitative scale with five intensity levels (not annoyed, slightly annoyed, moderately annoyed, very annoyed and extremely annoyed) to measure annoyance caused by ammonia. Llop et al. (2008) measure the annoyance caused by air pollution from outdoor sources (such as smoke, gas and particles) by categorizing annoyance through a scale including: not (0), slightly (1-3), moderately (4-7) and extremely (8-10) annoyed. To measure the perceived annoyance, the qualitative answers were recorded in a categorical/ ordinal 4-point scale (1 for not annoyed; 2 for slightly annoyed; 3 for very annoyed; and, 4 for extremely annoyed) to facilitate the interpretation of results by the Multivariate Correspondence Analysis (MCA) technique. Table 3 presents the questions listed in both surveys, the variable from each question and the factor groups of variables. For all questions the answer option "not know" (NK), was also included.

3. Multiple Correspondence Analysis (MCA)

MCA is a multivariate data analysis technique applicable to a large set of categorical variables (Greenacre, 2007). MCA, as the counterpart of PCA for categorical variables, became standard for the analysis of questionnaires (Le Roux & Rouanet, 2010). MCA is used to detect and represent data graphically (by the scatterplot) as a set of points with respect to two perpendicular coordinate axes: the horizontal axis often referred to as the *x*-axis and the vertical one as the *y*-axis. The objective of this technique is to analyse graphically the relationships between variables, response categories and objects, by reducing the dimensionality of the data set (Crivisqui, 1995; Lebart et al., 1984). The graphical representation can be made for individuals, for the variables, for the response categories as well as for the categories and individuals (Le Roux & Rouanet, 2010).

To apply MCA in the present study, data were initially organized in an Excel table containing respondents versus questions: the lines represented the respondents participating in the survey and the columns represented the questions listed in the applied questionnaire (each filled cell is the response category (answer) chosen by each individual for each question). The questions are categorized variables with finite number of response categories. For example, considering the question "Do you feel annoyed with air pollution?", the variable “annoyance” has the following response categories (with their encodings): not annoyed (ANNOY-1); Slightly annoyed (ANNOY-2); very annoyed (ANNOY-3); extremely annoyed (ANNOY-4), as well as the possibility of “no response/not known (NR/NK)” (ANNOY-9/99). Each respondent can choose one and only one response category for each variable or question. If the individual of line 1 answered "very annoyed" to the above question, the variable cell was filled with the "ANNOY-4" category, and so on for all the individuals for each variable.

According to Le Roux & Rouanet (2010), the interpretation of the MCA outcome is based on the observation of the cloud of points, which is defined as a finite set of points in a geometric space. The cloud of points can represent variables, response categories and individuals. In this work, the cloud of points represents the response categories. The great advantage of MCA is the possibility to reduce the multi-dimensional space in an optimal subspace that allows the study of the scatterplot and the consequent analysis and interpretation of results. The generated graphs allow to visually assess whether all the variables of interest have associations among them and how they are associated.

The size of the scatterplot depends on the number of information points in each row or column, minus one. If the number of columns is related to *K* categories of responses of *Q* variables, the maximum dimensionality of the scatterplot of categories is given by:

$$(K_1 - 1) + \cdots (K_q - 1) + \cdots (K_Q - 1) = (K_1 + \cdots + K_q + \cdots + K_Q) + (-1)Q = K - Q \qquad \text{Eq.1}$$

The dimensionality reduction is normally made to R^2, to facilitate interpretation of the cloud of points (scatterplot). Le Roux and Rouanet (2010) define the middle point of the cloud of points in the following way: Let *P* be any point in space and $(M^k)_{(k=1,2,\ldots,K)}$ the points of categories for the scatterplot. The midpoint of the cloud point is the *G* point by the vector $\overrightarrow{PG}$ as:

$$\overrightarrow{PG} = \frac{1}{n}\sum \overrightarrow{PM^k} \qquad \text{Eq.2}$$

Point *G* does not depend on the choice of point *P*, i.e., whatever the chosen point *P* is, point *G* is always the same. Thus, point G is defined as the average of the coordinates of all points given by:

$$G = \frac{1}{n}\sum M^k \qquad \text{Eq.3}$$

The distance between points depends on the different response categories for each variable. The lower the frequencies of the response categories are, the greater the distance

between points. Let n'_{kk} be the number of subjects who chose both categories *k* and *k'*, then the square of the distance between M^k and $M^{k\prime}$ is:

$$d^2\left(M^k M^{k'}\right) = \frac{n_k + n'_k - 2n'_{kk}}{n_k n'_{k/n}} \qquad \text{Eq.4}$$

As more categories *k'* and *k* are chosen for the same individuals, as shorter the distance between M^k and $M^{k'}$ is, and as closer two category points are, the stronger is the association between them. As lower the frequency for category *k* is, the farther from the centre the point M^k is. The less frequent the category of response is, the more it contributes to the overall variance of the cloud of individual points. And the less frequent the standard of responses of an individual is, the more it contributes to the total variance (Le Roux and Rouanet, 2010).

The first principal axis of a cloud of categories can be defined as the line passing through the midpoint of the cloud. The second main axis is perpendicular to the first one and is also passing through the midpoint of the cloud of categories points. The same process is followed to define the third axis, the fourth axis, and so on. There are no set rules for the number of axes to be analysed (Grenacre, 2006). In this study, it appears that the first two axes hold the highest percentage of the total variability of the data. Therefore, scatter plots of categories are formed from the first two axes.

The results of the MCA can be confusing depending on the number of variables. Because of this, the values of the contributions generated from the application of MCA "collaborate" in the interpretation of the axes (Le Roux and Rouanet, 2010). Therefore, in the present work, the contribution of each category as well as the summation are analysed to identify the variable that contributes most to the interpretation of a particular axis.

It is possible to select rows and columns that will generate the active points and also the illustrative points (supplementary points). The active points are responsible for determining the orientation of the principal axes, providing the necessary information for the construction of the optimal cloud of category points. However, it is possible to include more information which is represented by supplementary or illustrative points. The supplementary points may be plotted on the map along with the active points, and they are useful in interpreting features discovered in the primary data, but do not contribute to the construction of the axis (Grenacre, 2007). The supplementary points are used to represent information about the phenomenon under study and invariable information over time, such as sex, race, or information for infrequent categories.

According to Le Roux & Rouanet, (2010), the contribution of a category point to construct a particular axis defines the importance of this point for this same axis. Through the coefficients of this contribution, it is possible to identify which categories (or points) should be considered for the interpretation of each principal axis. The relative contribution constitutes the axis contribution to the variance of the individual point, the quality of representation of a point corresponds to the sum of the squared cosines of axes 1 and 2 (see Table 4 and Table 8). The test values assist the interpretation, but they don't contribute to the total variance and are interpreted as how well represented the supplementary points are (Greenacre, 2007) (see Tables 5, 6, 7 and 9).

To generate each MCA stage, active and supplementary variables were defined and the number of factors to compose the factorial plans was set. This decision was based on the analysis of the population composition. There is no set of rules defining how many factorial plans should be scanned in graphics (Le Roux & Rouanet, 2010). To facilitate interpretation, a decision was made to select the first two factorial plans (axis 1 and 2) to compose the correspondence graph or the cloud of points (scatter plot). The coordinates for the axes F1 and F2 are those of the position of each category in the scatter plot. The proportion of the variance of the cloud due to a point is called the contribution of the point to the cloud. The sum of the category contributions for each variable shows which variable contributes the most to the cloud and also to each axis. The representation's quality is expressed by the value of squared cosines (the higher the square cosine value, the higher the quality of representation).

Table 3 presents all questions, the name of each variables to represent these questions and the factors groups represented for each variable/question. The variables selected to all response categories were unified to a single database as the results of both surveys, since the goal is to analyse differences and associations among annoyance caused by air pollution. The sociodemographic and local variables are considered in order to compare differences and similarities between the respondents' opinions in the two study areas. It is important to note that in the first MCA (Figure 3) the active variables selected are nominal/ ordinal while the supplementary variables are nominal.

4. Results and discussion

Tables 4 to 9 present the numerical results obtained using the MCA from a matrix intersection of respondents (rows) and their responses to the five questions (columns) of the questionnaire (factor group named as in Table 3). For each active or supplementary variable, the response options (categories), encoding, frequency and percentages, coordinates of the two axes (F1 and F2), contributions in the construction of the two axes and the squared cosine values are presented.

In Table 4 the sum of the category contributions for each variable shows the active variable "annoyance" contributes the most to the cloud and also to each axis (annoyance contributes 27.9% to the axis F1 and 25.9% to F2). Figure 3 is the correspondence graph (scatterplot) with the coordinates of axes F1 and F2 generated for the active variables shown in Table 4 ("air pollution") and the supplementary variables shown in Table 5 ("local"). The axes F1 (52,67%) and F2 (18,25%) explain about 71% of the variability from the database, considering all active variables simultaneously, which is considered as an excellent performance (Le Roux & Rouanet, 2010).

In Figure 3 analysing the direction from right to left on the F1 axis, there is a progressive tendency for increased levels of annoyance as indicated by the position of the respective categories (ANNOY1-not annoyed, ANNOY2- slightly annoyed, ANNOY3-very annoyed and ANNOY4-extremely annoyed). The same progressive tendency (from categories 1 to 4) can be observed for the variables: importance of air quality (IMP1-not important, IMP2- slightly important, IMP3-very important, IMP4-extremely important); industrial risk perception (RISK1-not exposed, RISK2-slightly exposed, RISK3-very exposed, RISK4-extremely exposed); assessment of air quality (AIRQ1-excellent AIRQ2-good, AIRQ3-bad, and AIRQ4-horrible); and also, to the air pollution perception

variable (PPOL1-never, PPOL2-sometimes, PPOL3-often, PPOL4-always). The axis F1 can be interpreted as defining (from the right to the left) a scale of "perceived annoyance", and the axis F2 appears to oppose moderate response categories (lower side) to both extremely positive and extremely negative responses.

Continuing in Figure 3, the scatter plot can also be interpreted through the parabolic shape of the cloud of points of the chart from the bottom to the centre, setting categories for "slightly" and "very" levels, while the upper right corresponds to the "not" level and the upper left to the categories represented by the level "extremely". Such a pattern of responses suggests what is known in the literature as the "Guttman effect" (Greenacre & Blasius, 2006) or "horseshoe effect" (Van Rijckevorsel, 1987 *in* Greenacre & Blasius, 2006) due to its parabolic shape or arch. This is a structured form of the distribution of the categories of annoyance levels, which are arranged in a hierarchical way, from those who do not report nuisance (upper right), to those who express moderate annoyance (vertex of the parabolic) and arriving at extremely annoyed level (top left).

Also, in Figure 3 the joint progression of annoyance levels and other active categories from right to left in axis F1 indicates that an individual who reported being extremely annoyed due to air pollution also thought that the air quality was extremely important, felt extremely exposed to industrial risks, assessed air quality as horrible and always perceived air pollution by dust/odour/air visibility. Another possible pattern visible in Figure 3 is the "battery effect" which is often observed in survey analysis, where the respondents choose similar answers without necessarily considering the content of the questions. However, these questions were not presented in the same order to the Dunkirk and the Vitoria inhabitants as selected for this analysis. Furthermore, the response options presented for the question concerning the assessment of air quality were ordered in such a way that they express "opposed feelings" in relation to the response to other questions, causing the respondent to give due regard before answering.

Table 5 presents the categories for the sub-regions/areas where the respondents live, the code for each location, the frequencies and percentages of responses, the coordinates for each response category and the test values for the axes F1 and F2 related to the supplementary variables of the "Local" group. The test value is an indication of the significance of the obtained results (5% p-value or 1.96 p-value in absolute terms) (Crivisqui, 1995). The test value was calculated as the distance from each point to the origin of the axes F1 and F2, in numbers of standard deviations.

It is important to observe the negative test values in Table 5, since the negative values in axis F1 correspond to the local variables or the area where people have reported to be very annoyed, while positive test values correspond to the area where people have reported little or no annoyance by air pollution. According to the correspondence graph in Figure 3, the locations where respondents reported being very annoyed due to air pollution correspond to the localities that have negative test values (on the left part of the F1 axis): Grande Synthe (LOCAL-D7) where the location is at -2.699 standard deviations from the mean point (origin) on axis F1; Petite-Synthe (LOCAL-D8) test value = -3.11; Jardim Camburi (LOCAL-V4) test value = -4.138 and Enseada do Sua (LOCAL-V6) test value = -4.812. These regions are geographically close to the main industrial sources in both regions. The residents in these sub-regions reported intense levels of annoyance; they also reported being very exposed to industrial risk, often assessed the air quality as "horrible" and perceived high levels of air pollution due to dust/odour/ visibility in their

neighbourhoods. It is very interesting to note that a comparison between the air quality results shown in Tables 1 and 2, for both urban areas, shows that the PM_{10} concentrations are high, with peaks indicated by the maximum concentration, which suggests that the annoyance is not related only with this variable.

Figure 4 shows the correspondence graph between the active variables in the "air pollution" group and the supplementary variables in the "sociodemographic" group. In this graph, the local variables were removed to facilitate visualization of the correspondence with sociodemographic variables: gender, age, occupation and level of education. Visually, the most important supplementary variables are close to the origin of the graph.

Table 6 shows the results (frequency, coordinates and test value) for the correspondence graph in Figure 3 by each response category of sociodemographic variables. For MRV, it is possible to see that women reported being more annoyed than men, while in MRD this same association was not significantly. Consequently, in MRV women felt more exposed to industrial risk, assessed air quality as more important and perceived air pollution by dust/odour/visibility more than men. According to Fisher et al. (1991), the gender differences are noticeable especially in relation to environmental risks. Explanations are linked to the roles of women in society, which -even in our evolving societies- are most often oriented towards health and children. Gustafson (1998) also discusses this difference between men and women in relation to their roles in society and the power relations that exist between them. For example, women are more sensitive to environmental risks because they take care of their homes and children and clean the house normally more frequently than men, especially in more conservative societies.

Regarding Figure 4, in MRD there was no significant correspondence visible between the age categories and the annoyance categories. However, in MRV it is possible to visualise a progressive relation between age (AGE-V1, AGE-V2, AGE-V3, and AGE-V4) and levels of annoyance. As age increased, the levels of annoyance, the importance of air quality, the perceived exposure risk, the assessment of air quality, and the perceived air pollution also increased. This association can be confirmed considering the test values in Table 6. Respondents older than 34 years (AGE-V3, AGE-V4) are associated to being very or extremely annoyed more than younger respondents (AGE-V1, AGE-V2). Normally, older people are more sensitive to health problems since they belong to the more sensitive population sub-groups (it should be noted that children do not participate in the survey and the responses of women were discussed above). They also experience the effects of air pollution more often, such as when removing dust (particles) for house cleaning, so they report being more concerned with annoyance caused by air pollution. Although this association is not so clear in Dunkirk compared to Vitoria, it should be noted that the results of Lercher et al. (1995) and Klaeboe et al. (2000) suggest that older age is a determinant of perceived air pollution.

Regarding occupation, in MRD the unemployed (OCCUP-D2) and student (OCCUP-D4) are on the right side of the F1 axis, so, they are associated to being slightly or not annoyed by air pollution. In Vitoria, however, the categories associated to being very and extremely annoyed by air pollution are the retired group (OCCUP-V3) and also the unemployed (OCCUP-V2) on the left side of the F1 axis. This association is significant for the retired group (test value = -3.44) in Vitoria and can be justified because generally

they are the group with older age that are also associate to being very and extremely annoyed by air pollution (see Table 6).

Considering the corresponding graph (Figure 4) and the test values for the "level of education" categories in MRV (Table 6), it is possible to see that the university group (EDUC-V4) is on the high levels of annoyance side of the axis F1. This result corroborates with Klaeboe et al. (2000) that suggested that the education level was a determinant of perceived air pollution.

Figure 5 is the correspondence graph between "Air pollution" and "Health" groups. The supplementary variables selected were "health problems" and "health effects". People who report "no" (HEFE-1) occurrence of health problems caused by air pollution are also the ones who are "less annoyed". And people who answered "yes" (HEFE-2) to occurrence of health problems are associated to "being very annoyed". In table 7, the test value = -5.113 for the variable health problems caused by air pollution, indicate the same association, people who answered "no"(HEFE-1), as shown on the right part of the graph, tend not to feel annoyed by air pollution, while those who responded "yes"(HEFE-2), as shown on the left side of the graph, tend to report being extremely annoyed by air pollution.

For the group of people that answered "yes" (HEFE-2), the main problems reported were lung/respiratory (HPROB-1), allergies (HPROB-3), asthma attacks (HPROB-5) and stress (HPROB-7). However, this correspondence are not significant for all, because only the test value= -2.85 for health effects, indicate that people who reported being very/extremely annoyed by air pollution were associated to the ones who reported allergies (HPROB-3). Although previous epidemiological studies have shown that certain levels of particulate matter concentrations can cause such health effects and increase the number of hospitalizations for respiratory problems (Pope III, 1991; Schwartz, 1991; Braga et al., 2001; Garçon et al., 2006; Llop et al., 2008), in this analysis no correspondence was found with annoyance and health effects.

In order to further explore these survey data, the MCA between the groups of factors "Cause" and "Local" was carried out and the active variables "source", "METEO", "season" and "day/night" as well as the supplementary variables "Dunkirk" and "Vitoria" were selected for analysis. Table 8 presents the summation of the coordinates, contribution and squared cosine values for each response category. It can be observed that the categories grouped under "METEO" and "SEASON" contribute significantly to the F1 axis. And the categories under "SOURCE" contribute significantly to the F2 axis. The summation of the contributions of the other response categories as well as the value of the squared cosines for the F1 and F2 axes can confirm such affirmations.

Figure 6 presents the correspondence graph between the factor groups "Cause" and "Local". Analysing the active variable "METEO", the respondents that answered "no" to the question regarding the influence of meteorological conditions on the perception of air pollution (METEO-2) are located on the right part of the F1 axis, while the left part of the F1 axis indicates the respondents who answered "yes" (METEO-1). Analysing the active variable "SEASON," the left part of the F2 axis corresponds to the yes (METEO-1) category. It is possible to see a progressive tendency from spring (SEASON-4), to

summer (SEASON-1), to autumn (SEASON-2), to winter (SEASON-3). That is, the respondents that reported "yes" to the question about the influence of meteorological conditions on the perception of air pollution also perceived a progressive effect of seasonality from spring and summer to autumn and winter. As shown in other studies (for example Castanho & Artaxo, 2001; Albuquerque et. al., 2012; Andreão et al., 2019) meteorological conditions have a major influence on the suspended particle concentrations, which can suggest that meteorology is another determinant that affects the perception of air pollution.

Regarding the active variable "SOURCE", it is possible to observe the response category "source industry" (SOURCE-2) on the upper-right part of the graph and the categories "building" (SOURCE-4), "suspension of soil" (SOURCE-3), and "vehicle" (SOURCE-1) on the lower-middle of the graph with no association with the variable "METEO". This clearly shows the important role of industry in relation to air pollution in the two examined industrial urban areas, especially as this is perceived by the population. The response categories "yes" (DN-2) and "no" (DN-1) for perceived air pollution changes between day and night are close to the origin and are non-significant.

For the "LOCAL" supplementary variable, the lower part of the graph and on the middle and left of the F1 axis corresponds to the areas in Vitoria showing that the people perceived an influence of weather changes on the quantity of particles/dust (corresponding to METEO-1). The upper part of the graph on the middle and right part of the F1 axis shows that in Dunkirk, there is no perceived association between the meteorological conditions and air pollution (corresponding to METEO-2). This result may be related to the fact that people who live close to industries are already accustomed to pollution and they do not feel the influence of weather changes on the perceived air pollution. Furthermore, as Figure 1 indicates, in Dunkirk industrial pollution sources are more interspersed within the urban area, while in Vitoria they are more at the boundaries of the city (especially northwest, but also partly southeast) and therefore the meteorological conditions – such as wind direction and wind speed – may have a more concrete influence on the perceived air pollution.

Regarding the categories "Source" associated with "Local", the results in Figure 5 suggest that, for Vitoria, the locations far from the main industrial areas, like Cariacica (LOCAL-V7) and Vitoria-centre (LOCAL-V5), were more associated with the construction work/building source (SOURCE-4) and slightly with the vehicular source (SOURCE-1) of air pollution. It should be noted that in these areas, and especially Cariacica, high levels of particulate pollution were measured by the air quality monitoring stations (Table 3) and according to Santos et al., (2017) the main source of particulate matter in Cariacica is construction work. Currently, construction work in Vitoria is developing rapidly in terms of housing construction and paving streets and roads, while the number of vehicles is also increasing: these sources can contribute to increasing air pollution and dust/particles that cause annoyance, especially since these locations are further away from industrial sources. In Dunkirk, the locations close to the main industrial areas, St. Pol Sur Mer (LOCAL-D6), Grande Synthe (LOCAL-D7) and Petite-Synthe (LOCAL-D8) were found to be associated with the "industry" source (SOURCE-2), while the "construction work/building" and "vehicular" sources are not significant, as are the "industrial" sources which are located next to residential areas and visible from the main beaches. St. Pol Sur Mer is also the location where the highest levels of particulate pollution were measured (Table 2) and also very close to the industrial sources.

The above results may also partly be explained by the size of the population in Vitoria, which is larger than in Dunkirk, and the fact that public transport in Dunkirk is well developed; therefore, the number of vehicles circulating in Dunkirk is not as high as in Vitoria. According to Rotko et al. (2002) and Amundsen et al. (2008) heavy traffic is related to annoyance caused by air pollution, so this result can explain why people exposed to heavy traffic in Vitoria perceived vehicular sources more significant than in Dunkirk, especially in areas which are not as influenced by industrial sources. Also, as noted above, currently the construction sector in Vitoria is developing rapidly both in terms of housing construction and paving streets and roads, which can generate dust that may cause annoyance. As suggested by Nikolopoulou et al. (2011), air quality is often considered to be poor at construction sites, which are burdened by higher PM concentrations.

Table 9 shows the coordinates and the test values for the "LOCAL" supplementary categories. Regarding the test values, there is a strong association (not contribution) with the variable that contributes to the F2 axis. The positive test values for the F2 axis are the Dunkirk locations in the upper quadrants of Figure 6, indicating the respondents which perceived urban air pollution from industries sources. The negative test values for the F2 axis are the Vitoria locations in the lower quadrants, which indicating respondents that perceived urban air pollution from construction works and vehicular sources.

5. Conclusions

The purpose of this work was to apply the multiple correspondence analysis (MCA) technique to explore relationships between variables that can influence peoples' behaviour concerning annoyance caused by air pollution. Data were collected using a survey on air pollution, environmental issues and quality of life. The MCA technique was applied as a differential tooling to interpret factors by structural properties and significant, graphical relationships of proximity and opposition between points.

The results showed a progressive correspondence between levels of annoyance and other active variables in the "air pollution" factor group: as the levels of annoyance increased, the levels of the other qualitative variables (importance of air quality, perceived exposure to industrial risk, assessment of air quality, perceived air pollution) also increased. It is possible to conclude that people who reported feeling annoyed by air pollution also thought that air quality was very important, were very concerned about exposure to industrial risks, assessed often air quality as horrible and frequently perceived air pollution by dust/odour/visibility. It is important to emphasize that this result cannot be considered as a "battery effect", since these questions were not applied in the sequence in which the results were analysed. In addition, the summary results of PM_{10} concentration measurements in the two regions showed values above the guidelines established by the World Health Organization for long and short term exposure, which can be an indicator of attention to the possibility of occurrence of health problems, quality of life problems, and complaints about perceived annoyance caused by air pollution. According to the correspondence graph, people who live in areas close to industries, for example in Petite-Synthe and Grande Synthe in Dunkirk and Enseada do Suá and Jardim Camburi in Vitoria, have reported being very annoyed by air pollution. As showed by Machado et al. (2018), the location and the proximity to industrial sources of air pollution play an important role to explain peoples' behaviour.

This study has also explored the association between respondents' perceptions and demographic characteristics. The gender and age variables were associated with perceived annoyance in MRV, while in MRD there was a much lower correspondence. In general, women were more annoyed than men, while people older than 55 years reported feeling more annoyed than people in other age ranges. For MRV, the study showed that women are more annoyed than men and that there is a positive relationship between age and the parameters level of annoyance, importance of air quality, perceived exposure to industrial risk, assessment of air quality and perceived air pollution by dust/odour/visibility.

Kohlhuber et al. (2006) have shown that socio-economic variables are associated with the perception of local air quality, suggesting that these variables may be important determinants of perceived air quality. However, in our study, the relationship between different forms of occupation remains controversial. It was found that retired and unemployed people showed high levels of annoyance in MRV, while in MRD students and unemployed showed low or no level of annoyance. It was observed that people educated to the university level were more concerned about annoyance caused by air pollution than those with lower educational levels in MRV. The importance of socioeconomic factors in the context of air pollution research has been emphasized because they represent underlying aspects that affect susceptibility, exposure, or disease diagnosis and treatment (Bell et al., 2005). Therefore, there is a need for careful choice and interpretation of socioeconomic factors depending on the location, as this can be partly influenced by social/cultural differences or/and by different weather conditions.

Significant association was observed between perceived health risks related to high level of annoyance caused by air pollution, importance of air quality, perceived exposure to industrial risk, assessment of air quality and perceived air pollution by dust/odour/visibility. Perceived occurrence of health problems related to annoyance caused by air pollution was found to be associated with gender (female), age (older), level of education (university level) and the type of occupation (retired), though it varied by the study location. The difference between the two sites could be explained by the difference in the age, occupation and education levels of the residents. There were more people older, employed and with university level in MRD compared to MRV, while in MRV there are more young people, students and with primary level of education as compared to MRD. The main health effects related to air pollution reported by respondents were lung/respiratory, allergies and eye irritation, which are common symptoms for a number of urban air pollutants (WHO, 2005). There is also a confirmed association between occurrence of allergies and high levels of annoyance.

The MCA results for the active variable "CAUSE" showed that people perceived that weather conditions and seasonal changes could affect air pollution. This perception was more evident in MRV, where heavy industries are at the boundaries of the city and their effect is more influenced by the prevailing meteorological conditions, such as the wind speed and especially wind direction. In MRD, people identified industrial sources as an important cause of air pollution and did not perceive that air pollution annoyance changes with changes in weather/meteorological conditions. Furthermore, in MRV, the influence of construction works and vehicular sources on the perception of air pollution was evident. Considering the geographic location of these two regions, the weather conditions

669 could influence the perceived annoyance caused by air pollution in different ways, and
670 therefore the geographic location can justify these differences.
671
672 In both regions, the results from PM_{10} concentration measurements show that the
673 perceived annoyance does not only depend on this variable. This conclusion may help to
674 better understand if the answers of the respondents are linked to other variables, such as
675 the source of pollution and the meteorology. It is also important to note that there are
676 differences in air pollution perceptions in the two study areas, which indicates that
677 perceptions in general may depend on an area's overall setting and availability of
678 industries, pollution sources or daily activities.
679
680 The results of this study have shown that Multiple Correspondence Analysis is a very
681 useful tool in providing insight on environmental issues affecting the quality of life, such
682 as the factors affecting the levels of air pollution annoyance of populations living in urban
683 areas. This is a very useful tool to derive and synthesize important information from
684 citizen surveys which can, in turn, complement air quality measurements, in order to
685 define the best mix of actions to address air quality issues, for instance, national or
686 regional emission reduction policies in order to meet the air quality objectives in
687 background locations, as well as to decide on more site-specific and short-term measures
688 to address air pollution. Such combination of measures is often necessary for the
689 protection of the public health and the improvement of the quality of life of citizens.
690
691 **6. Declarations**
692
693 The ethics committee approved this study was from the University Federal of Espírito
694 Santo with the 142.581 committee's reference number. This manuscript does not contain
695 any individual /specific person's data.
696
697 The datasets used and/or analysed during the current study are available from the
698 corresponding author on reasonable request.
699
700 The authors declare that they have no competing interests.
701
702 This study is part of the results from the research developed in the PhD thesis of the first
703 author (MM) under the supervision of the authors JS and VR who made the necessary
704 corrections. All authors declare to have contributed equally by reading and improving the
705 analyses and the text.
706
707 These authors are indebted to CentraleSupélec and Université Paris-Sud for their financial
708 supports. This research was also partially supported by the iCODE Institute, research
709 project of the IDEX Paris-Saclay, and by the Hadamard Mathematics LabEx (LMH)
710 through the grant number ANR-11-LABX-0056-LMH in the Programme des
711 Investissements d'Avenir.
712
713 The authors would like to acknowledge the support of two Brazilian research funding
714 agencies: FAPES (The Research Funding Agency in the State of Espirito Santo) and
715 CAPES (The Federal Funding Agency for the Development of Higher Education). They
716 would also like to acknowledge COFECUB, Irénée Zwarterook and a research group
717 studying industrial risk and the urban environment in the TVES laboratory, Université du

718 Littoral Côte d'Opale in France. Finally, we thank Vinicius de Martin (UFES' phd
719 student) for helping with graphical design.
720
721 **7. References**
722
723 Albuquerque, T. T. de A.; Andrade, M. de F.; Ynoue, R. Y. Characterization of
724 atmospheric aerosols in the city of São Paulo, Brazil: comparisons between polluted and
725 unpolluted periods. Environ. Monit. Assess. 2012, 184, 969–984, doi:10.1007/s10661-
726 011-2013-y.

727 ADEME (2002). Classification et critères d'implantation des stations de surveillance de
728 la qualité de l'air. Retrieved on 28 July 2013 from:
729 http://www.oramip.org/pdf/ademe_typologies.pdf

730 Amundsen, A. H.; Klæboe, R.; Fyhri, A. Annoyance from vehicular air pollution:
731 Exposure–response relationships for Norway. Atmos. Environ. 2008, 42, 7679–7688,
732 doi:10.1016/j.atmosenv.2008.05.026.

733 Andreão, W. L., Trindade, B. T., Nascimento, A. P., Reis Júnior, N. C., Andrade, M. D.
734 F., & Albuquerque, T. T. D. A. (2019). Influence of Meteorology on Fine Particles
735 Concentration in Vitória Metropolitan Region During Wintertime. Revista Brasileira de
736 Meteorologia, 34(4), 459-470.

737 Atari, D. O.; Luginaah, I. N.; Fung, K. The Relationship between Odour Annoyance
738 Scores and Modelled Ambient Air Pollution in Sarnia, "Chemical Valley", Ontario. Int.
739 J. Environ. Res. Public Health 2009, 6, 2655–2675, doi:10.3390/ijerph6102655.

740 Atmo Nord-Pas-de-Calais (2009). Bilan 2008 des poussières sédimentables sur le
741 Dunkerquois. Rapport d'études 01 – 2009- LC. Available at: http://www.atmo-npdc.fr

742 Barnett, V. Sample survey: principles and methods; 3rd ed.; Wiley: Chichester, 2002;
743 ISBN 9780470685907.

744 Bell, M.L.; O'Neill, M.S.; Cifuentes, L.A.; Braga, A.L.F.; Green, C.; Nweke, A.; Rogat,
745 J.; Sibold, K. Challenges and recommendations for the study of socioeconomic factors
746 and air pollution health effects. Environ. Sci. Policy 2005, 8, 525–533

747 Benzécri, J.P. (1969) Statistical analysis as a tool to make patterns emerge from data. In
748 S. Watanabe (ed.), Methodologies of pattem regonition.New York: Academic press.

749 Benzécri, J.P. (1973) L'Analyse des Données. Vol. 2: L'Analyse des Correspondances.
750 Paris: Dunod.

751 Benzécri, J.P. (1992) Correspondence analysis handbook. New York: Dekker.

752 Berglund, B.; Berglund, U.; Lindvall, T. Measurement and control of annoyance. In
753 Environmental annoyance: characterization, measurement and control; Koelga, H. S.,
754 Ed.; Elsevier Sci B.V.: Amstredam, 1987; Vol. 15, pp. 29–43.

Blanes-Vidal, V.; Suh, H.; Nadimi, E. S.; Løfstrøm, P.; Ellermann, T.; Andersen, H. V.; Schwartz, J. Residential exposure to outdoor air pollution from livestock operations and perceived annoyance among citizens. Environ. Int. 2012, 40, 44–50.

Boyouk, N., Léon, J. F., Delbarre, H., Augustin, P., & Fourmentin, M. (2011). Impact of sea breeze on vertical structure of aerosol optical properties in Dunkerque, France. Atmospheric research, 101(4), 902-910.

Braga A.L.F.; Saldiva, P.H.N.; Pereira, L.A.A.; Menezes, J.J.C.; Conceição, G.M.S.; Lin, C.L.; Zanobetti, A.; Shwartz, J.; Dockery, D.W. (2001). Health effects or air pollution exposure on children and adolescents in São Paulo, Brasil. Pediatr. Pulmonol., 31, 106-113, 2001.

Calvo-Mendieta, I.; Flaquart, H.; Frere, S.; Gonthier, F.; Hellequin, A. P.; Blanc, A. Le Perception du risque industriel par les populations du Dunkerquois, Rapport intermédiaire.; Dunkerque, 2008.

Castanho, A. D. A., & Artaxo, P. (2001). Wintertime and summertime São Paulo aerosol source apportionment study. Atmospheric Environment, 35, 4889–4902.

Cochran, W. G. Sampling techniques - A Wiley publication in applied statistics; 3rd ed.; Wiley: New York, 1977; ISBN 047116240X.

Crivisqui, E. (1995). The correspondence analysis as instrument to analyse the profile of the worker in the tomato crop in the Goiás state, Brazil. Programme de Recherche et D'Enseignement en Statistique Appliquée. PRESTA, Belgique: Université Libre de Bruxelles.

Egondi T.; Kyobutungi C.; Ng N.; Muindi K.; Oti S.; Vijver S. V. De; Ettarh R.; Rocklov J.(2013) Comunity perceptions of air pollution and related health risks in Nairobi Slums. Int. J. Environ. Res. Public Health, 10, 4851-4868.

Farfel, M.R., Orlova, A.O., Lees, P.S.J., Rohde, C., Ashley, P.J., Julian Chisolm, J., 2005. A study of urban housing demolition as a source of lead in ambient dust on sidewalks, streets, and alleys. Environ. Res. 99, 204–213. doi:10.1016/j.envres.2004.10.005

Fischer G.W., Morgan M.G., Fischoff B., Nair I, and Lave L.B. (1991), “What risk are people concerned about?”, Risk Analysis, 11, 303-314.

Garçon G., Dagher Z., Zerimech F., Ledoux F., Courcot D., Aboukais A., Puskaric E., Shirali P. (2006) Dunkerque city air pollution particulate matter-induced cytotoxicity, oxidative stress and inflammation in human epithelial lung cells (L132) in culture. Toxicology in Vitro, 20, 519-528.

Grenacre, M. J. (1984) Theory and Applications of correspondence analysis. New York: Academic press.

Greenacre, M. & Blasius, J. (2006). Multiple correspondence analysis and related methods. Chapman & Hall/CRC Press, London.

792 Greenacre, M. (2007). Correspondence Analysis in Practice, Second Edition. London:
793 Chapman & Hall/CRC.

794 Gustafson E. (1998). Gender differences in Risk perception: theoretical and
795 methodological Perspectives, Risk Analysis, 18(6), 805-811.

796 Hellequin, A. P.; Zwarterook, I. Mauvais air: vivre au quotidien près d'usines polluantes
797 et dangereuse. In IIeme colloque international UMR 5600-ENTPE "Le risque industriel:
798 une question de sciences humaines"; Lyon, 2010.

799 IBGE Sensus 2010; Instituto Brasileiro de Geografia e Estatística: Brasilia, 2010.

800 (IJSN) INSTITUTO JONES DOS SANTOS NEVES –. Regional Profile 2009: Região
801 Metropolitana da Grande Vitória. Disponível em: http://www.ijsn.es.gov.br. Acesso em
802 09 de junho de 2015.

803 Jacquemin B., Sunyer J., Forsberg B., Gotschi T., Oglesby L., Ackermann-Liebrich U.,
804 De Marco R., Heinrich J., Jarvis D., Toren K., Kunzli N. 2007. Annoyance due to air
805 pollution in Europe. International Journal of Epidemiology, 36, 809–820.

806 Kohlhuber, M.; Mielck, A.; Weiland, S.K.; Bolte, G. (2006). Social inequality in
807 perceived environmental exposures in relation to housing conditions in Germany.
808 Environ. Res., 101, 246–255.

809 Kim M., Yi O., Kim H. (2012). The role differences in individual and community
810 attributes in perceived air quality. Science of the Total Environment, 425, 20-26.

811 Klæboe, R., Kolbenstvedt, M., Clench-Aas, J., Bartonova, A. (2000). Oslo traffic study
812 part 1: an integrated approach to assess the combined effects of noise and air pollution on
813 annoyance. Atmospheric Environment, 34, 4727-4736.

814 Lebart L, Morineau A, Warwick K. (1984). Multivariate descriptive statistical analysis.
815 Chichester, UK: Wiley.

816 Lercher, P., Schmitzberger, R., Kofler, W. (1995). Perceived traffic air pollution,
817 associated behavior and health in an alpine area. Science of the Total Environment 169
818 (1–3), 71–74.

819 Le Roux, B. and Rouanet, H. (2004). Geometric Data Analysis, From Correspondence
820 Analysis to Structured Data Analysis. Dordrecht. Kluwer Academic Publishers.

821 Le Roux, B. et Rouanet, H. (2010). Multiple Correspondence Analysis, SAGE, Series:
822 Quantitative Applications in the Social Sciences, CA:Thousand Oaks Paris.

823 Les ateliers. Atelier international de maîtrise d'œuvre urbaine. Dunkerque, une
824 agglomération maritime transfrontalière actrice d'un développement métropolitain.
825 Cahier de session - 2ème edition. Du 3 au 17 juillet 2010.

826 Lindvall, T., Radford, E.P., 1973. Measurement of annoyance due to exposure to
827 environmental factors. Environ. Res. 6, 1–36. doi:10.1016/0013-9351(73)90014-5.

828 Llop S., Ballester F., Estarlich M., Esplugues A., Fernández-Patier R., Ramón R., Marco
829 A., Aguirre A., Sunyer J., Iñiguez C., on behalf of INMA-Valencia cohort. Ambient air
830 pollution and annoyance responses from pregnant women. Atmospheric Environment,
831 42, 2982-2992, 2008.

832 Machado, M., Santos, J. M., Reisen, V. A., Reis Jr, N. C., Mavroidis, I., & Lima, A. T.
833 (2018). A new methodology to derive settleable particulate matter guidelines to assist
834 policy-makers on reducing public nuisance. Atmospheric Environment, 182, 242-251.

835 Machado, M.; Reisen, V. A.; Santos, J. M.; Reis, N. C.; Frère S.; Bondon P.; Ispány M.;
836 Cotta H. H.A. Use of multivariate time series techniques to estimate the impact of
837 particulate matter on the perceived annoyance. Atmos. Environ. 2019, 117080.
838 DOI: 10.1016/j.atmosenv.2019.117080.

839 Nascimento A.P.; Santos, J. M.; Mill J.G.; Souza J.B.; Reis, N. C.; Reisen, V. A. (2017).
840 Association between the concentration of fine particles in the atmosphere and acute
841 respiratory diseases in children. Revista de Saúde Pública, 51, 3. Epub January 12, 2017.

842 Nikolopoulou M., Kleissl J., Linden P.F., Lykoudis S. (2011). Pedestrians' perception of
843 environmental stimuli through field surveys: Focus on particulate pollution. Science of
844 the Total Environment, 409(13), 2493-202.

845 Oglesby, L.; Künzli, N.; Monn, C.; Schindler, C.; Ackermann-Liebrich, U.; Leuenberger,
846 P. Validity of annoyance scores for estimation of long-term air pollution exposure in
847 epidemiologic studies: the Swiss Study on Air Pollution and Lung Diseases in Adults
848 (SAPALDIA). Am. J. Epidemiol. 2000, 152, 75–83,
849 doi:10.1080/10473289.2000.10464156.

850 Orru, H.; Idavain, J.; Pindus, M.; Orru, K.; Kesanurm, K.; Lang, A.; Tomasova, J.
851 Residents' Self-Reported Health Effects and Annoyance in Relation to Air Pollution
852 Exposure in an Industrial Area in Eastern-Estonia. Int. J. Environ. Res. Public Health
853 2018, 15, 252, doi:10.3390/ijerph15020252.

854 Orru, K.; Nordin, S.; Harzia, H.; Orru, H. The role of perceived air pollution and health
855 risk perception in health symptoms and disease: a population-based study combined with
856 modelled levels of PM10. Int. Arch. Occup. Environ. Health 2018b, 91, 581–589,
857 doi:10.1007/s00420-018-1303-x.

858 Passchier-Vermeer, Willy, and Wim F. Passchier. "Noise exposure and public
859 health." Environmental health perspectives 108.suppl 1 (2000): 123-131.

860 Pope C.A. III, (1991). Respiratory hospital admissions associated with PM10 pollution
861 in Utah, Salt Lake and Cache valleys. Arch. Environ. Health, 46, 90-97.

862 PPA (2002). Plan de protection de l'atmosphère de l'aglomeration Dunkerquoise.
863 Retrieved on 20 April 2013 2013 from: http://www.nord-pas-de-calais.developpement-
864 durable.gouv.fr/IMG/pdf/ppa-dunkerque.pdf

865 Rotko T., Oglesby L., Kunzli N., Carrer P., Nieuwenhuijsen M.J., Jantunen M. (2002).
866 Determinants of perceived air pollution annoyance and association between annoyance

867 scores and air pollution (PM2.5. NO2) concentrations in the European EXPOLIS study.
868 Atmospheric Environment, 36, 4593–4602.

869 Salvador, N., Reis, N. C., Santos, J. M., de Almeida Albuquerque, T. T., Loriato, A. G.,
870 Delbarre, H., ... & Moreira, D. M. (2016). Evaluation of weather research and forecasting
871 model parameterizations under sea-breeze conditions in a North Sea coastal
872 environment. Journal of Meteorological Research, 30(6), 998-1018

873 Santos, J. M.; Reis Jr, N. C.; Galvão, E. S.; Silveira, A.; Goulart, E. V.; Lima, A. T.; Reis,
874 N. C.; Galvão, E. S.; Silveira, A.; Goulart, E. V.; Lima, A. T. Source apportionment of
875 settleable particles in a mining-impacted urban and industrialized region in Brazil.
876 Environ. Sci. Pollut. Res. 2017, doi:10.1007/s11356-017-9677-y.

877 Souza, J. B.; Reisen, V. A.; Franco, G. C.; Ispány, M.; Bondon, P.; Santos, J. M.
878 Generalized additive models with principal component analysis: an application to time
879 series of respiratory disease and air pollution data. J. R. Stat. Soc. Ser. C (Applied Stat.
880 2018, 67, 453–480, doi:10.1111/rssc.12239.s

881 Stenlund, T.; Lidén, E.; Andersson, K.; Garvill, J.; Nordin, S. Annoyance and health
882 symptoms and their influencing factors: A population-based air pollution intervention
883 study. Public Health 2009, 123, 339–345, doi:10.1016/j.puhe.2008.12.021.

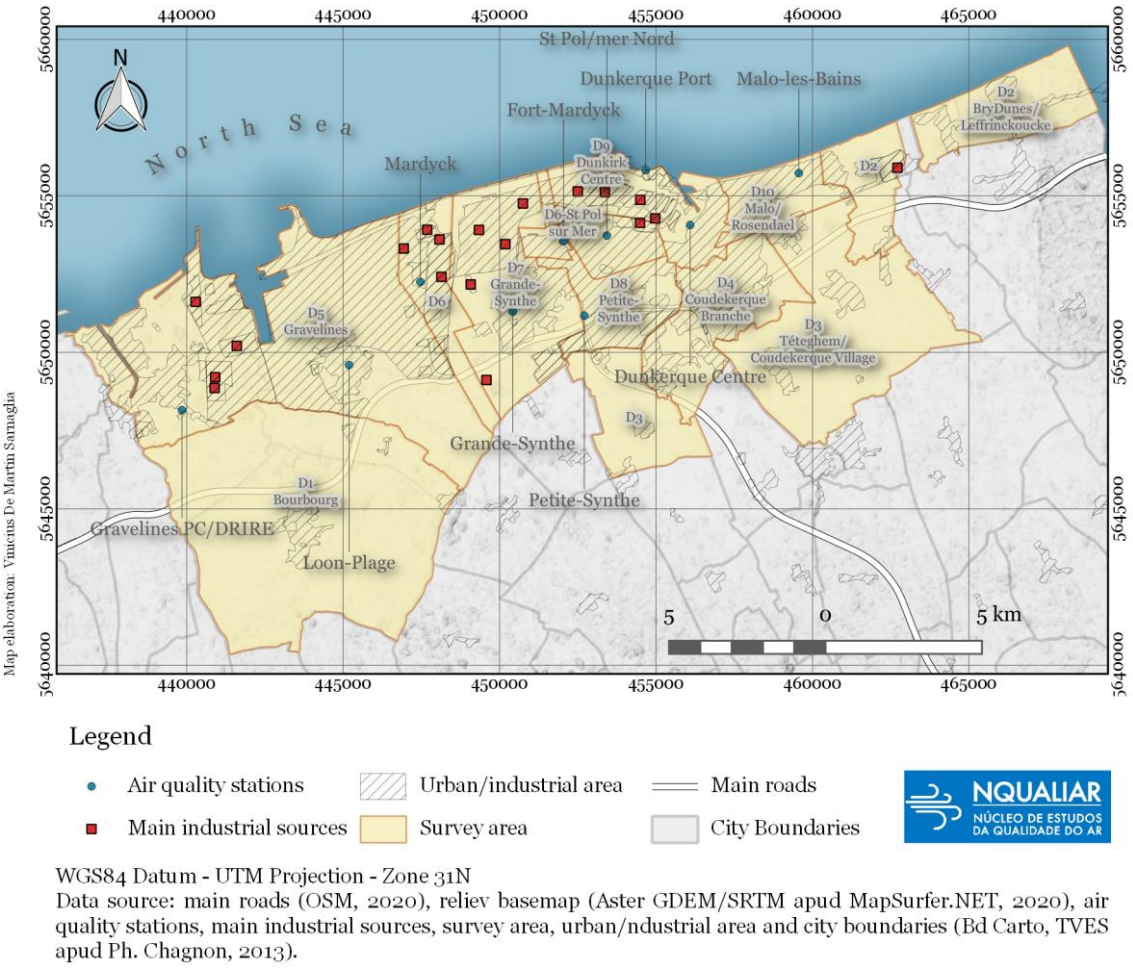
884 Schwartz, J. (1991). Particulate air pollution and daily mortality in Detroit.
885 Environmental Research, 56(2), 204-213.

886 Vallack, H.; Shillito, D. Suggested guidelines for deposited ambient dust. Atmos.
887 Environ. 1998, 32, 2737–2744, doi:10.1016/S1352-2310(98)00037-5.

888 WHO (2005). WHO air quality guidelines global update 2005. Report on a WHO
889 Working Group. Bonn, Germany, 18-20 October 2005. Retrieved on 28 July 2014 from:
890 http://www.euro.who.int/Document/E87950.pdf.

891 WHO (2014). Burden of disease from the joint effects of Household and Ambient Air
892 Pollution for 2012. WHO: Geneva. Retrieved on 28 July 2014 from:
893 http://www.who.int/phe/health_topics/outdoorair/databases/FINAL_HAP_AAP_BoD_2
894 4March2014.pdf?ua=1

Figure



WGS84 Datum - UTM Projection - Zone 31N
Data source: main roads (OSM, 2020), reliev basemap (Aster GDEM/SRTM apud MapSurfer.NET, 2020), air quality stations, main industrial sources, survey area, urban/ndustrial area and city boundaries (Bd Carto, TVES apud Ph. Chagnon, 2013).

Figure 1. Sub-regions of interest and the main emission sources in Dunkirk.

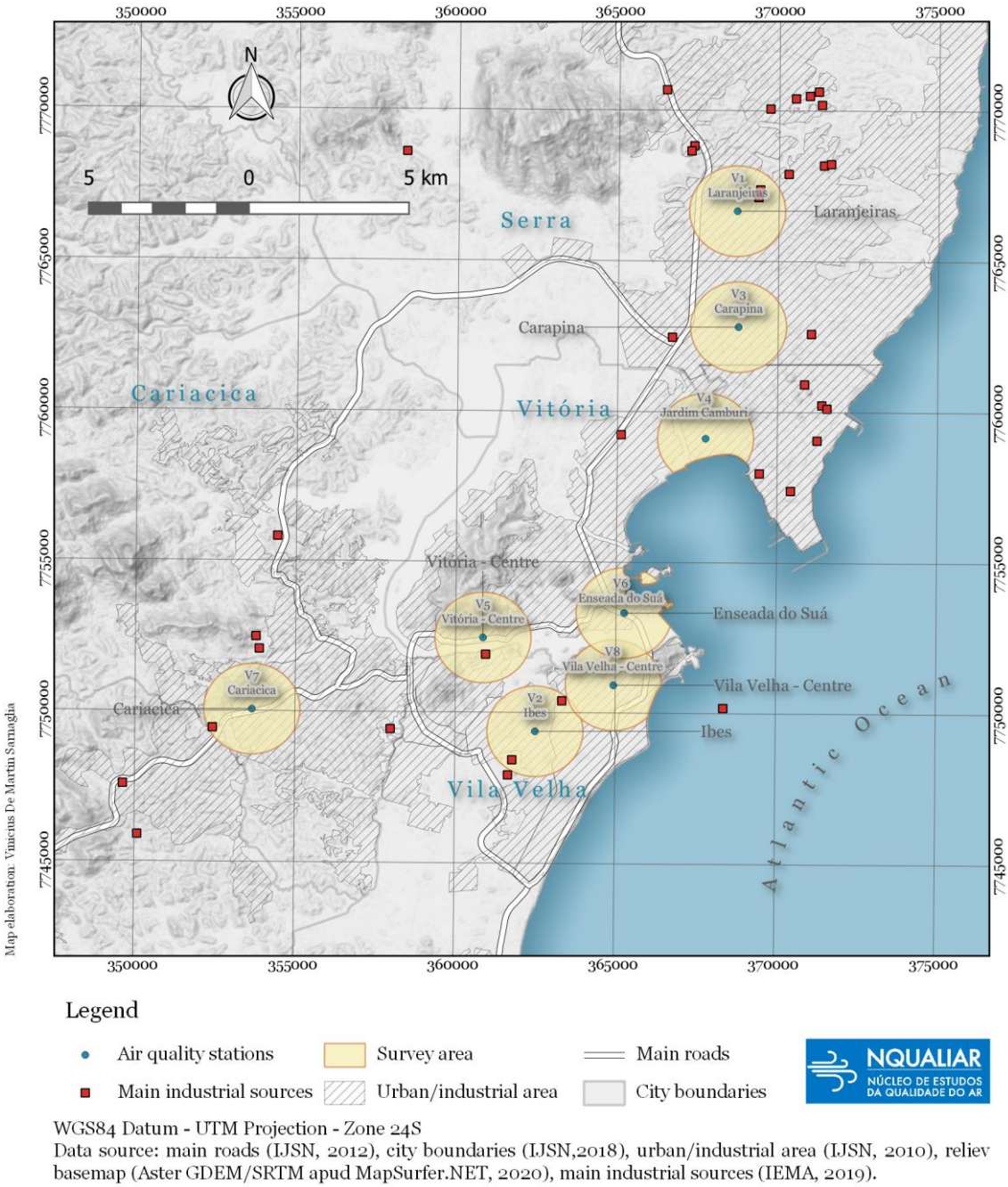


Figure 2. Sub-regions of interest (survey area) and the main emission sources in Vitoria.

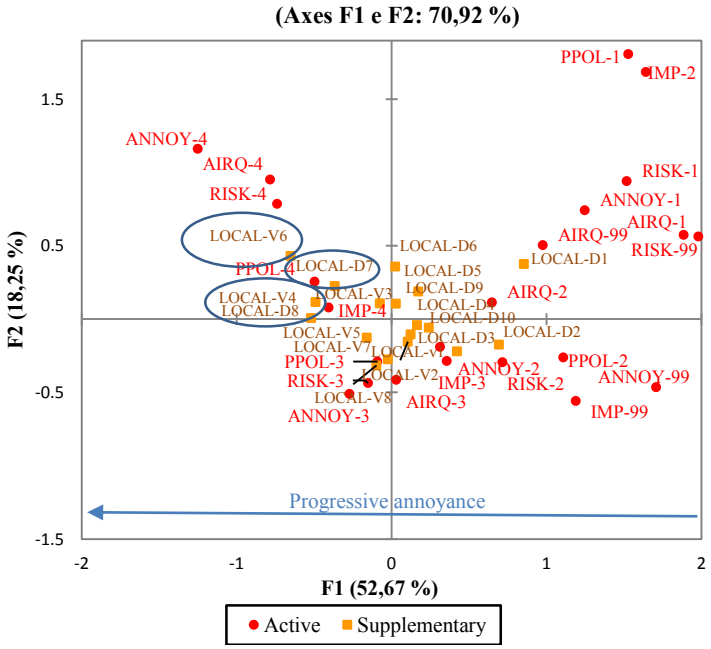


Figure 3. MCA results for the "air pollution" group of active variables and the "housing location" group of supplementary variables.

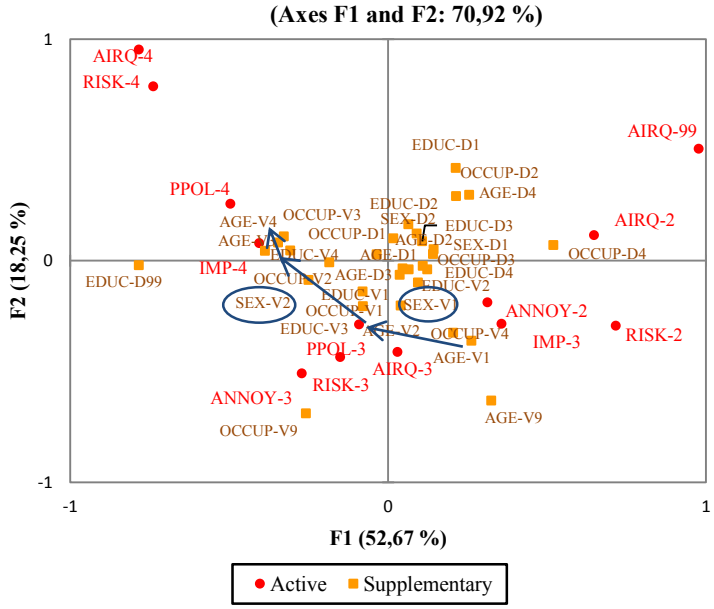


Figure 4. MCA results for the "air pollution" group of active variables and the "sociodemographic" group of supplementary variables.

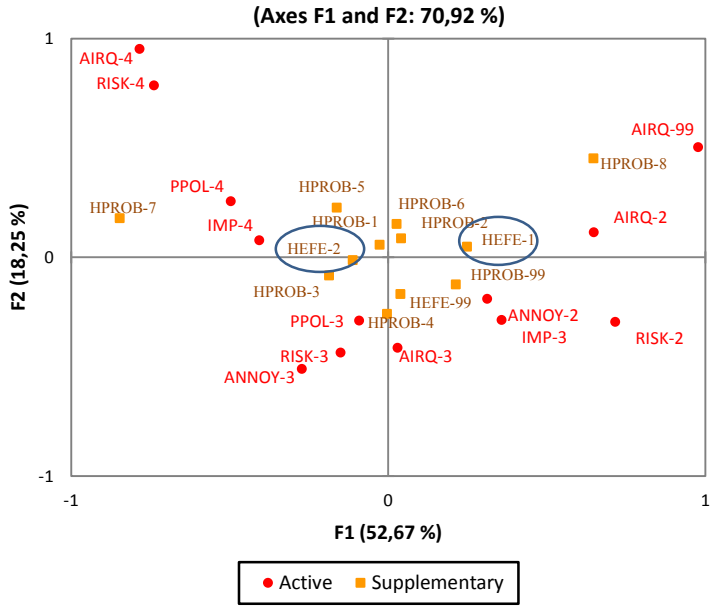


Figure 5. MCA results for the "air pollution" group of active variables and the "health problems" group of supplementary variables

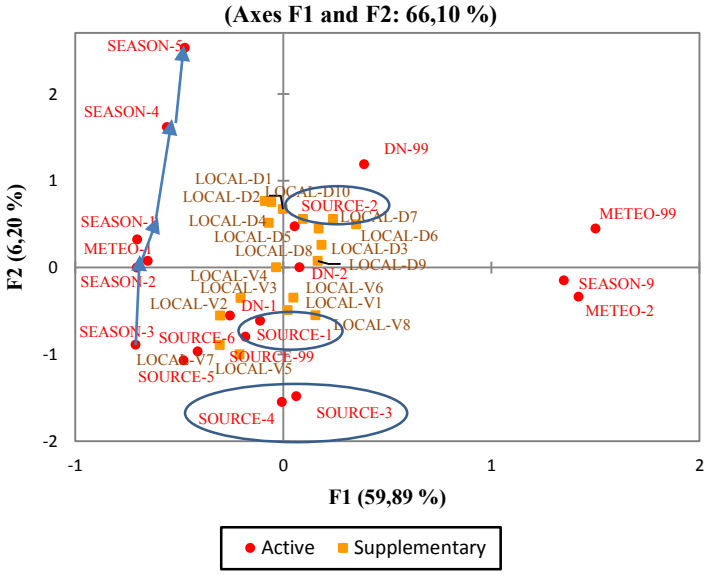


Figure 6. MCA results for the "cause" group of active variables and the "local" group of supplementary variables

Table 1. Descriptive statistics of PM_{10} concentrations (μg/m³) at the Dunkirk monitoring stations in 2008

Stations	Minimum	Maximum	Percentile 90%	Percentile 95%	Mean	Std. deviation
Dunkerque Centre	6.47	84.87	31.76	34.98	21.51	8.95
Malo-les-Bains	6.04	63.20	27.66	31.44	19.10	8.03
Fort-Mardyck	6.62	158.81	50.68	56.56	31.59	16.89
StPol/merNord	5.65	143.08	54.32	64.40	34.19	17.02
Petite-Synthe	7.29	80.25	31.43	36.48	21.44	9.70
Grande-Synthe	8.17	114.25	46.60	57.49	27.69	16.71
Mardyck	5.87	101.14	47.41	61.16	25.39	16.99

Table 2. Descriptive statistics of PM_{10} concentrations (μg/m³) at the Vitoria monitoring stations

Stations	Minimum	Maximum	Percentile 90%	Percentile 95%	Mean	Std. deviation
Laranjeiras	8.29	85.58	50.58	57.08	33.60	12.60
Carapina	2.88	84.76	31.01	35.41	21.51	7.86
Jardim Camburi	7.75	76.66	29.12	32.38	21.44	7.17
Enseada do Suá	9.70	71.28	41.14	46.13	29.44	8.77
Vitória Centro	11.13	83.12	41.45	47.45	29.26	9.31
Vila Velha-Ibes	6.41	76.75	46.87	52.29	31.22	12.09
Cariacica	11.16	115.20	73.04	85.01	46.67	19.4

Table 3. The questions from the two surveys and the factor groups for the variables represented by these questions

Questions	Variable	Factor groups
Do you feel annoyed by air pollution?	Annoyance	Air pollution
Is air quality important to you?	Importance or air quality	
How do you feel regarding industrial risk	Industrial risk perception	
How do you rate air quality in your	Assessment of air quality	
How often do you perceive air pollution due	Air pollution perception	
Where do you live?	Dunkirk-locals	Local
	Vitoria- locals	
What is your gender?	Gender	Socio-demographic
Who old are you?	Age	
What is your occupation?	Occupation	
What is your level of education?	Level of education	
Do you or someone in your household have	Health problems	Health
What are the main health problems caused by	Health effects	
What is the main source of settled dust in your	Source	Cause
Do you think that air pollution/settled	Meteorological conditions	
If yes, in which season is air pollution/settled	Season	
Do you think air pollution is worst during the	Day/ Night	

Table 4. MCA results of the active variables grouped under the "Air pollution" factor group

Active variables	Categories	Code	Freq. (%)		Coordinates		Contributions		Squared cosines	
					Axis F1	Axis F2	Axis F1	Axis F2	Axis F1	Axis F2
Annoyance	Not annoyed	ANNOY 1	167	(16,16)	1.248	0.743	0.132	0.058	0.300	0.106
	Slightly ann.	ANNOY 2	265	(25,65)	0.313	-0.190	0.013	0.006	0.034	0.012
	Very Annoyed	ANNOY 3	458	(44,33)	-0.271	-0.510	0.017	0.075	0.059	0.207
	Extremely ann.	ANNOY 4	139	(13.45)	-1.251	1.162	0.111	0.119	0.243	0.210
	NK	ANNOY 9/99	4	(0.38)	1.709	-0.463	0.006	0.001	0.011	0.001
Importance of air quality	Not important	IMP 1	11	(1.06)	2.390	2.932	0.032	0.060	0.062	0.093
	Slightly imp.	IMP 2	28	(2.7)	1.643	1.684	0.038	0.050	0.075	0.079
	Very important	IMP 3	423	(40.94)	0.358	-0.286	0.028	0.022	0.089	0.057
	Extremely imp.	IMP 4	566	(54.79)	-0.406	0.078	0.047	0.002	0.199	0.007
	NK	IMP 9/99	5	(0.48)	1.189	-0.558	0.004	0.001	0.007	0.002
Industrial risk perception	Not exposed	RISK 1	76	(7.3)	1.518	0.941	0.089	0.042	0.183	0.070
	Slightly exp.	RISK 2	186	(18.0)	0.716	-0.295	0.049	0.010	0.113	0.019
	Very exposed	RISK 3	508	(49.17)	-0.150	-0.436	0.006	0.061	0.022	0.184
	Extremely exp.	RISK 4	255	(24.68)	-0.738	0.785	0.071	0.099	0.179	0.202
	NK	RISK 9/99	8	(0.77)	1.980	0.564	0.016	0.002	0.031	0.002
Assessment of air quality	Excellent	AIRQ 1	10	(0.96)	1.886	0.575	0.018	0.002	0.035	0.003
	Good	AIRQ 2	178	(17.23)	0.649	0.114	0.038	0.001	0.088	0.003
	Bad	AIRQ 3	601	(58.18)	0.030	-0.414	0.000	0.065	0.001	0.238
	Horrible	AIRQ 4	222	(21.49)	-0.783	0.952	0.069	0.127	0.168	0.248
	NK	AIRQ 9/99	22	(2.1)	0.978	0.504	0.011	0.004	0.021	0.006
Air pollution perception	Never	PPOL 1	20	(1.9)	1.529	1.808	0.024	0.041	0.046	0.065
	Sometimes	PPOL 2	146	(14.13)	1.110	-0.261	0.091	0.006	0.203	0.011
	Often	PPOL 3	474	(45.88)	-0.091	-0.289	0.002	0.025	0.007	0.071
	Always	PPOL 4	375	(36.3)	-0.495	0.256	0.047	0.016	0.140	0.037
	NR	PPOL 9	3	(0.29)	0.512	-1.801	0.000	0.006	0.001	0.009
	NK	PPOL99	15	(1.45)	2.324	3.219	0.041	0.098	0.080	0.153

Table 5. MCA results of the supplementary variables grouped under the "Local" factor group

Supplementary Variables	Category	Code	Freq. (%)	Coordinates		Test value	
				Axis F1	Axis F2	Axis F1	Axis F2
Dunkirk	Bourbourg	LOCAL-D1	24 (2.32)	0.856	0.376	4.240	1.864
	BrayDunes/Leffrinckoucke	LOCAL-D2	33 (3.19)	0.694	-0.175	4.048	-1.020
	Téteghem/Coudekerque/Village	LOCAL-D3	31 (3.00)	0.124	-0.105	0.702	-0.592
	Coudekerque Branche	LOCAL-D4	82 (7.93)	0.242	-0.058	2.280	-0.547
	Gravelines	LOCAL-D5	60 (5.80)	0.174	0.189	1.385	1.508
	St Pol sur Mer	LOCAL-D6	83 (8.03)	0.024	0.359	0.228	3.406
	Grande Synthe	LOCAL-D7	52 (5.03)	-0.365	0.224	-2.699	1.660
	Petite-Synthe	LOCAL-D8	35 (3.38)	-0.518	0.007	-3.114	0.044
	Dunkerque	LOCAL-D9	30 (2.90)	0.030	0.106	0.166	0.590
	Malo/Rosendael	LOCAL-D10	88 (8.51)	0.166	-0.041	1.629	-0.403
Vitoria	Laranjeiras	LOCAL-V1	51 (4.93)	0.424	-0.219	3.102	-1.606
	Ibes	LOCAL-V2	82 (7.93)	0.104	-0.154	0.986	-1.450
	Carapina	LOCAL-V3	37 (3.58)	-0.074	0.108	-0.456	0.669
	Jardim Camburi	LOCAL-V4	67 (6.48)	-0.489	0.116	-4.138	0.980
	Centro de Vitoria	LOCAL-V5	51 (4.93)	-0.160	-0.127	-1.171	-0.928
	Enseada do Sua´	LOCAL-V6	52 (5.34)	-0.651	0.430	-4.812	3.177
	Cariacica	LOCAL-V7	85 (8.22)	-0.023	-0.274	-0.220	-2.639
	Centro de Vila Velha	LOCAL-V8	90 (8.71)	-0.096	-0.315	-0.958	-3.130

Table 6. MCA results of the supplementary variables grouped under the "Sociodemographic" factor group

City	Supplementary variables	Category	Code	Freq. (%)	Coordinates		Test value	
					Axis F1	Axis F2	Axis F1	Axis F2
Dunkirk	Gender	Male	SEX- D1	211 (20.426)	0.232	0.090	2.337	0.818
		Female	SEX-D2	307 (29.719)	0.145	0.219	1.881	2.535
	Age	16-24	AGE-D1	77 (7.454)	0.066	-0.040	0.598	-0.368
		25-34	AGE-D2	83 (8.035)	0.017	0.100	0.162	0.947
		35-54	AGE-D3	191 (18.490)	0.046	-0.036	0.699	-0.547
		>55	AGE-D4	167 (16.167)	0.256	0.296	3.610	4.183
	Occupation	Employed	OCCUP-D1	247 (23.911)	-0.035	0.028	-0.635	0.502
		Unemployed	OCCUP-D2	107 (10.358)	0.215	0.290	2.344	3.164
		Retired	OCCUP-D3	108 (10.455)	0.142	0.029	1.558	0.322
		Student	OCCUP-D4	54 (5.227)	0.520	0.069	3.925	0.523
		NK	OCCUP-D99	2 (0.194)	0.128	1.531	0.181	2.167
	Education level	No formal educ.	EDUC-D1	48 (4.647)	0.213	0.417	1.510	2.957
		Primary school	EDUC-D2	92 (8.906)	0.064	0.164	0.641	1.643
		High school	EDUC-D3	213 (20.620)	0.107	0.087	1.759	1.431
		University	EDUC-D4	164(15.876)	0.123	-0.041	1.723	-0.568
		NK	EDUC-D99	2 (0.194)	-0.784	-0.021	-1.109	-0.030
Vitoria	Gender	Male	SEX-V1	205 (19.845)	0.153	-0.180	1.517	-1.597
		Female	SEX-V2	310 (30.010)	-0.404	-0.159	-5.251	-1.858
	Age	16-24	AGE-V1	140 (13.553)	0.263	-0.362	3.344	-4.608
		25-34	AGE-V2	71 (6.873)	0.040	-0.204	0.350	-1.782
		35-54	AGE-V3	145 (14.037)	-0.308	0.045	-3.996	0.583
		>55	AGE-V4	156 (15.102)	-0.346	0.082	-4.683	1.106
		NK	AGE-V9	3(0.290)	0.325	-0.632	0.563	-1.096
	Occupation	Employed	OCCUP-V1	226 (21.878)	-0.080	-0.140	-1.362	-2.378
		Unemployed	OCCUP-V2	113 (10.939)	-0.184	-0.008	-2.077	-0.089
		Retired	OCCUP-V3	100 (9.681)	-0.327	0.109	-3.440	1.147
		Student	OCCUP-V4	72 (6.970)	0.205	-0.326	1.804	-2.870
		NK	OCCUP-V9	4 (0.387)	-0.258	-0.690	-0.517	-1.382
	Education level	No formal education	EDUC-V1	7 (0.678)	0.037	-0.064	0.097	-0.171
		Primary school	EDUC-V2	131 (12.682)	0.109	-0.025	1.335	-0.307
		High school	EDUC-V3	238 (23.040)	-0.079	-0.207	-1.386	-3.633
		University	EDUC-V4	138 (13.359)	-0.387	0.043	-4.883	0.546

Table 7. MCA results of the supplementary variables grouped under the "Health" factor group

Supplementary Variables	Category	Cod.	Freq. %		coordinates		Test value	
					Axis F1	Axis F2	Axis F1	Axis F2
Health problem	No	HEFE-1	304	29.429	0.249	0.048	5.175	0.999
	Yes	HEFE-2	694	67.183	-0.111	-0.013	-5.113	-0.581
	NR/NK	HEFE-99	26	3.388	0.039	-0.167	0.234	-1.007
Health effects	Lung/respiratory	HPROB-1	454	43.950	-0.027	0.057	-0.759	1.630
	Cancer	HPROB-2	95	9.197	0.041	0.086	0.420	0.880
	Allergies	HPROB-3	192	18.587	-0.186	-0.082	-2.856	-1.262
	Bronchitis	HPROB-4	35	3.388	-0.003	-0.258	-0.017	-1.550
	Asthma attacks	HPROB-5	5	0.484	-0.162	0.228	-0.362	0.510
	Eye irritation	HPROB-6	46	4.453	0.027	0.152	0.188	1.057
	Stress	HPROB-7	5	0.484	-0.846	0.179	-1.895	0.402
	Others	HPROB-8	11	1.065	0.648	0.453	2.159	1.510
	NK	HPROB-99	190	18.393	0.214	-0.123	3.267	-1.880

Table 8. MCA results of the active variables grouped under the "Cause" factor group

Active variables	Categories	Cod.	Freq. (%)		Coordinates		Contributions		Squared cosines	
					AxisF1	Axis F2	Axis F1	Axis F2	Axis F1	Axis F2
Source	Vehicle	SOURCE-1	190	18.393	-0.111	-0.614	0.001	0.053	0.003	0.085
	Industry	SOURCE-2	688	66.602	0.054	0.474	0.001	0.115	0.006	0.448
	Soil suspension	SOURCE-3	22	2.130	0.062	-1.481	0.000	0.036	0.000	0.048
	Building	SOURCE-4	85	8.228	-0.006	-1.551	0.000	0.152	0.000	0.216
	Sea breeze /wind	SOURCE-5	21	2.033	-0.410	-0.968	0.002	0.015	0.003	0.019
	Others	SOURCE-6	12	1.162	-0.479	-1.071	0.001	0.010	0.003	0.013
	NK	SOURCE-99	15	1.452	-0.182	-0.798	0.000	0.007	0.000	0.009
Meteo	Yes	METEO-1	711	68.829	-0.651	0.074	0.150	0.003	0.935	0.012
	No	METEO-2	251	24.298	1.419	-0.336	0.252	0.021	0.646	0.036
	NK	METEO-99	71	6.873	1.501	0.449	0.080	0.011	0.166	0.015
Season	Summer	SEASON-1	392	37.948	-0.702	0.323	0.096	0.030	0.301	0.064
	Autumn	SEASON-2	47	4.550	-0.703	-0.001	0.012	0.000	0.024	0.000
	Winter	SEASON-3	189	18.296	-0.709	-0.888	0.047	0.111	0.112	0.177
	Spring	SEASON-4	49	4.743	-0.558	1.615	0.008	0.095	0.016	0.130
	Depend industries	SEASON-5	6	0.581	-0.472	2.529	0.001	0.029	0.001	0.037
	NK	SEASON-99	350	33.882	1.349	-0.151	0.318	0.006	0.933	0.012
Day/Night	Yes	DN-1	427	41.336	-0.255	-0.556	0.014	0.098	0.046	0.218
	No	DN-2	407	39.400	0.078	0.003	0.001	0.000	0.004	0.000
	NK	DN-99	199	19.264	0.388	1.187	0.015	0.209	0.036	0.336

Table 9. MCA results of the supplementary variables grouped under the "Local" factor group

Supplementary Variables	Category	Cod.	Freq. (%)	Coordinates		Test value	
				Axis F1	Axis F2	Axis F1	Axis F2
Dunkirk	Bourbourg	LOCAL-D1	24 (2.32)	-0.057	0.752	-0.284	3.723
	BrayDunes/Leffrinckoucke	LOCAL-D2	33 (3.19)	-0.001	0.673	-0.007	3.925
	Téteghem/ Coudekerque Village	LOCAL-D3	31 (3.00)	0.170	0.446	0.960	2.521
	Coudekerque Branche	LOCAL-D4	82 (7.93)	-0.087	0.763	-0.825	7.198
	Gravelines	LOCAL-D5	60 (5.80)	-0.069	0.513	-0.551	4.089
	St Pol sur Mer	LOCAL-D6	83 (8.03)	0.351	0.495	3.330	4.697
	Grande Synthe	LOCAL-D7	52 (5.03)	0.240	0.559	1.772	4.132
	Petite-Synthe	LOCAL-D8	35 (3.38)	0.184	0.259	1.106	1.560
	Dunkirk centre	LOCAL-D9	30 (2.90)	0.165	0.075	0.919	0.414
	Malo/Rosendael	LOCAL-D10	88 (8.51)	0.094	0.559	0.926	5.485
Vitoria	Laranjeiras	LOCAL-V1	51 (4.93)	0.021	-0.491	0.152	-3.596
	Ibes	LOCAL-V2	82 (7.93)	-0.303	-0.556	-2.860	-5.240
	Carapina	LOCAL-V3	37 (3.58)	-0.205	-0.352	-1.268	-2.177
	Jardim Camburi	LOCAL-V4	67 (6.48)	-0.035	0.003	-0.293	0.023
	Vitoria-centro	LOCAL-V5	51 (4.93)	-0.209	-1.001	-1.531	-7.330
	Enseada do Sua	LOCAL-V6	52 (5.34)	0.049	-0.347	0.360	-2.564
	Cariacica	LOCAL-V7	85 (8.22)	-0.305	-0.893	-2.933	-8.590
	Vila Velha-centro	LOCAL-V8	90 (8.71)	0.155	-0.551	1.537	-5.471