Chapter 1

Rough Set Applied to Air Pollution: A New Approach to Manage Pollutions in High Risk Rate Industrial Areas

Agata Matarazzo

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75630

Abstract

This study presents a rough set application, using together the ideas of classical rough set approach, based on the indiscernibility relation and the dominance-based rough set approach (DRSA), to air micro-pollution management in an industrial site with a high environmental risk rate, such as the industrial area of Syracuse, located in the South of Italy (Sicily). This new data analysis tool has been applied to different decision problems in various fields with considerable success, since it is able to deal both with quantitative and with qualitative data and the results are expressed in terms of decision rules understandable by the decision-maker. In this chapter, some issue related to multi-attribute sorting (i.e. preference-ordered classification) of air pollution risk is presented, considering some meteorological variables, both qualitative and quantitative as attributes, and criteria describing the different objects (pollution occurrences) to be classified, that is, different levels of sulfur oxides (SOx), nitrogen oxides (NOx), and methane (CH_4) as pollution indicators. The most significant results obtained from this particular application are presented and discussed: examples of 'if, … then' decision rules, attribute relevance as output of the data analysis also in terms of exchangeable or indispensable attributes/criteria, of qualitative substitution effect and interaction between them.

Keywords: industrial areas, air pollution, meteorological attributes, rough set approach

1. Introduction

Air pollution in a region depends mainly on the emission of pollutants and on local meteorological conditions. The probability of air pollution occurrences may be estimated by simple atmospheric dispersion models with proper meteorological data and predefined typical air pollution sources [1, 2].

A lot of studies, however, do not give enough information about the possible relationships between sampling and meteorological parameters, as well as their optimal correspondence formal tools in order to enable modeling and determination of patterns which are characteristic of the investigated area. Proposing conceptual models enables decision-makers at many levels to assess and manage air quality as a whole, rather than on a pollutant-by-pollutant concentration. By developing a holistic approach to air quality, it is possible to evaluate its extensive benefits of more effective developments in existing air-quality features, thereby avoiding the growth of air-quality ceilings, and to consider air quality within its wider meteorological context [3, 4]. By establishing why and when air pollution occasions may occur across a region, strategies should be designed and implemented so as to deal with such episodes. The possibility of forecasting pollutant concentration near the ground with high spatial detail offers the opportunity of constantly monitoring and managing the territory. Air-quality modeling procedures can forecast the behavior and the effects of the substances emitted from identified sources, particularly using data from meteorological instruments. These models can supply the distribution of pollutant concentrations on the ground, and are used for thermoelectric power plant management, being very useful in the case of exceptional events, such as when a highly dangerous pollutant escapes [5].

This study analyses the main relationships between air micro-pollution and meteorological conditions of the area surrounding Siracusa, a city located in Sicily. This was done by measuring air samples from a receiving station near a small town called Melilli, a Sicilian industrial area with a high environmental risk rate [6, 7].

This station has been chosen because it allows the production of a complete picture with respect to the amount of micro-pollution data and meteorological variables descriptions [8]. Then the most reliable parameters for the phenomena of the dispersion of micro-pollutants were identified and also the various critical scenarios were checked, so that all available air pollution sources were considered [9–11]. In particular, a specially designed model, with forecasting abilities of air pollution, has been developed, working independently from the knowledge of the local sources [12]. This monitoring model uses temperature and wind vertical profiles, measured by Radar Analysis Support System (RASS, a radar manufacturer-independent system for evaluating the different elements of a radar by connecting to signals) and SOnic Detection And Ranging (SODAR, a meteorological instrument used as a wind profiler to measure the scattering of sound waves by atmospheric turbulence) and concentration data from ground stations. The local values are correlated with the characteristics of the thermal profile and the direction and intensity of the wind at a selected altitude. On the basis of stored and statistically analyzed data, the model is able to forecast the pollution in the area surrounding the ground station [13] and to give useful information about the management of its main sources.

From the methodological point of view, the proposed approach is in the framework of multicriteria decision analysis, where a lot of different points of views, often conflicting one other, are explicitly considered together to support effective decisions. The utility or, better, the necessity of a multicriteria evaluation in public policies has been recently underlined by Munda [14].

The novel method of data analysis applied to the study of air micro-pollution management, the rough set approach (RSA), considers objects described by a lot of both qualitative and/or quantitative attributes and criteria (that is their 'profile'). In this context, inconsistencies between descriptions and risk classes assignments need not be removed prior to the analysis, therefore giving useful information about the quality of the inferred decision rules; moreover, the RSA also allows for highlighting the attributes which most contribute to air pollution among those taken into account for the assessment, giving too some useful information about the management of pollution.

Furthermore, this method is able to identify redundant attributes. This concerns the elimination of superfluous data from the data table, without deteriorating the quality of the results, that is, obtaining the same information of that inferred from the original table, therefore permitting enormous savings in data collection. Additionally, the rough set theory also shows a posteriori the relative importance of the considered attributes and criteria, without requiring a priori any elicitation or assessment of technical parameters (such as importance weights, trade-off, etc.), which are often very difficult to provide and never easily understandable by decision-makers.

The results hereby obtained are just an example of the RSA application, in order to understand how and why it is possible to apply this approach to environmental problems.

This chapter contains other five sections; Section 2 explains the basic principles of rough set theory and its main methodological features; Section 3 shows air micro-pollution analyzed data; Section 4 presents the main decision rules obtained; Section 5 discusses the interpretations of the results from the methodological and operational points of view; and lastly, Section 6 concludes this chapter.

2. The rough set theory

The rough set theory (RST), introduced by Pawlak [15–17], has proved to be an excellent tool for data analysis, even in the presence of inconsistencies and ambiguities. The main idea of the RSA is that every object in the universe U (data to be analyzed) is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for their description (e.g. if the objects are air pollution observed by monitoring stations, attributes may be air temperature, the relative humidity index, direction and wind speed, quantities of some micro-pollutants, etc.). Objects having the same description [18] in terms of these attributes are called indiscernible (similar); the indiscernibility relation thus generated induces a partition of the universe U into blocks of indiscernible objects, called elementary sets or granules of knowledge, which therefore result in information granulation. If set U is divided in some classes, objects indiscernible should belong to the same class to be consistent with the indiscernibility principle.

From the universe U, any subset X can be expressed either precisely (as a union of elementary sets) or approximately. In the latter case, the subset X may be characterized by two ordinary sets, called the lower and upper approximations. The lower approximation of X is composed of all the elementary sets included in X (whose elements, therefore, certainly belong to X), while the upper approximation of X consists of all the elementary sets which have a non-empty intersection with X (whose elements, therefore, may belong to X). A rough set is defined by means of these two approximations, which coincide in the case of an ordinary set. The difference between the two approximations represents the boundary region, whose elements cannot be characterized with certainty as belonging or not to X. The information about objects from the boundary region is, therefore, inconsistent or ambiguous.

The original RSA based on the indiscernibility relation (usually called classical rough set approach) is not, however, able to deal with preference ordered attribute domains (so-called criteria) and preference ordered decision classes (sorting problem), very often crucial for application to real problems in the field of multicriteria decision analysis.

To be able to deal with criteria and ordered decision classes, Greco et al. [19–25] have proposed an extension of the original rough set theory, called dominance-based rough set approach (DRSA), mainly based on the substitution of the indiscernibility relation by a dominance relation: object *a* dominates object *b*, if and only if *a* is at least as good as *b* with respect to all considered conditional criteria. In a similar way, the decision attribute *d* makes a partition of U into a finite number of preference ordered classes, Cl = {Clt, t = 1,…, *n*}, each $x \in$U belonging to one and only one class, Clt∈Cl. We can therefore state a basic consistency principle with respect to the dominance relation: if object *a* dominates object *b* with respect to a set of criteria and *b* belong to class Clt, *a* should belong at least to class Clt (upward union of Clt). Otherwise, there is an inconsistency with respect to the dominance principle. Therefore, *x* belongs to the lower approximation of any subset X of U if all objects dominating *x* belong to at least the same class of *x*, that is, *x* belongs to Clt or better without any ambiguity; x belongs to the upper approximation of X if among the objects dominated by x there is at least an object y belonging to Clt or better. In a similar way, it is possible to define lower and upper approximation of downward union of classes. Also in DRSA, the difference between the two approximations represents the boundary region.

The objects from U can be split into some decision classes by decisional criterion d, obtaining a decision table (DT), where each object x is described using some independent variables, called conditional attributes/criteria, and each object is assigned to a class of this partition, considered as a dependent variable. The quality of classification expresses the ratio between the objects which have been correctly classified and the total number of the elements of the DT, it lies between 0 (any object is not correctly classified) and 1 (all the objects of the universe are correctly classified), and therefore it can measure the goodness of the classification.

Besides, the classification quality may be unaltered if certain conditioned attributes are eliminated because they are superfluous. The minimal sets of the attributes which maintain the same classification quality of the entire table are called reducts. The intersection among all the reducts generates the core (the set of the most important attributes, which consequently cannot be eliminated without deteriorating the quality of the classification). Therefore, the attributes belonging to the core are indispensable, while the attributes belonging to the reducts are exchangeable with one another; the others are actually superfluous.

The relations existing among conditional attributes/criteria and decisional classes in the multicriteria sorting problem are expressed by decision rules. These are logical statements of the type 'if…, then…', where the antecedent (condition part) specifies values assumed by one or more condition attributes/criteria and the consequence specifies an assignment to one or more decision classes. If there is only one possible consequence, then the rule is said to be certain, otherwise, it is said to be approximate or ambiguous. An object $x \in U$ supports decision rule r if its description is matching both the condition part and the decision part of the rule; certain rules are supported only by objects from the lower approximation of the corresponding decision class; approximate rules are supported only by objects from the boundaries of the corresponding decision classes.

Procedures for the generation of decision rules from a decision are complex tasks, and a number of procedures have been proposed to solve it [18, 25–27].

The existing induction algorithms use one of the following strategies:

- the generation of a minimal set of rules covering all objects from a decision table;
- the generation of an exhaustive set of rules consisting of all possible rules for a decision table; and
- the generation of a set of 'strong' decision rules, even partly discriminant, covering relatively many objects from the decision table (but not necessarily all of them).

In this chapter to infer the rules, the jMAF software has been used, that is available for free in the Internet: RSES – Rough Set Exploration System, http://logic.mimuw.edu.pl/~rses , ROSE – ROugh Set data Explorer http://idss.cs.put.poznan.pl/site/rose.html, –jMAF, java Multi-criteria and Multi-attribute Analysis Framework http://www.cs.put.poznan.pl/jblaszczynski/Site/jRS.html, and jRank – ranking generator using Dominance-based Rough Set Approach http://www.cs.put.poznan.pl/mszelag/Software/jRank/jRank.html [28].

The rules inferred by DRSA can use also the 'at least' and 'at most' terms in their conditional and decisional parts. All these rules are expressed in a natural language, simple to understand the studied phenomenon and for decision support [22]. This means that the proposed approach actually is also able to explain the reasons of a particular pollution situation, moreover showing the real examples of these (traceability of decisions), and is able to support the management in preventing pollution damages, presenting them the situations where some critical events are most probable. Moreover, parameters like the support (the number of the objects which satisfy both the conditional part and the decisional part of the rule) and the confidence (the ratio between support and the number of the objects which satisfy only the conditional part of the rule, expressed in percentage) help the decision-maker in their choice of the most relevant rules.

We can summarize the main characteristics of the rough set approach as follows. With respect to input information (object description), both quantitative and qualitative data can be considered, even if they present some inconsistencies. With reference to output, information about the relevance of attributes and the quality of approximation can be acquired, and the final results are expressed in the form of 'if…, then…' decision rules, which are sentences that decision-makers find easier to understand [29–31] and using only the most relevant attributes/criteria (i.e. some reduct).

In the case of air pollution problem at hand, for example, we can consider some different decision classes of pollution according to an increasing level of some micro-pollutants (SOx, NOx,…). Since some meteorological variables (conditional attributes/criteria) present a monotonic relationship with the degree of pollution (e.g. the air temperature, the degree of humidity) and other no (e.g. wind direction, etc.), it is very important from both the operational and methodological points of view to take into consideration and to exploit in the appropriate way in the description of the objects and in the rule induction attributes and criteria distinctly. Therefore, we have to consider the indiscernibility relation with respect to the former, the dominance relation with respect to the latter, and the assignment to ordered classes with respect to the decision.

Greco et al. [26] proposed an approach for this kind of real-life multicriteria problems. This can be easily modeled by introducing some appropriate thresholds to discretize the conditional attributes and to characterize different levels of air pollution, for the decision classes. No discretization is required with respect to criteria, using the DRSA.

Consequently, the rough sets could be very efficiently applied in the case of uncertainty derived from the granularity of information. Actually, granules of condition attributes/criteria (objects having the same descriptions or respectively belonging to the same dominating/dominated sets) are used to approximate granules of decision (assignment to some decision classes).

The RSA is therefore very different with respect to the fuzzy sets, where the linguistic imprecision due to the use of natural language is mainly considered, and the membership function aims at indicating in what degree each object belongs to a particular class. Of course, the two approaches are not mutually exclusive, but they can actually be used in a complementary way [32–34]. Using a terminology from image representation, we could say that rough sets are related to the number of pixels of an image (its resolution), while the fuzzy sets represent the number of gray levels between black and white. At an operational level, the implementation of fuzzy sets always requires the definition and specification of particular membership functions, one for each attribute, not easy to specify analytically. Therefore, both classical rough set approach and fuzzy sets are sensitive to the specification of these values and both interesting and useful sensitivity and robustness analysis are actually useful and recommended by moving the level of the thresholds and other parameters [30, 35, 36]. It is not the case of DRSA, where actually no parameter should be elicited, but only some example of decision (from the past experiences of from expert knowledge) is needed to model the preference of the decision-maker.

3. Data description

Air micro-pollution-analyzed data come from an air monitoring network, working since 1975, covering an industrial area of 500 km^2, including the towns of Priolo–Melilli–Siracusa, situated in the province of Siracusa, in the region of Sicily. This industrial area was declared 'a high environmental risk rate place' by the Law 349/86 and covers six surrounding towns (Augusta, Priolo, Melilli, Siracusa, Floridia, Solarino); the landscape is very varied and is formed by sandy hills, mountains, and plains near the coast [37, 38].

In this territory, a lot of chemical plants, energy production industries, and oil refineries are found, as well as members of a private organization, the industrial trust for environmental safety (CIPA, Consorzio Industriale Protezione Ambiente–Environmental Protection Industrial Consortium). In its operative center, CIPA assembles and works out different micro-pollution parameters and various meteorological variables, measured by 12 different monitoring stations. Data collection and processing is useful in statistical analysis and in upgrading air pollution management in order to avoid the air-tested-exceeding threshold qualities, previously established [38–42].

This chapter studies monitoring station in Melilli only, because in this place, data concerning air micro-pollution quantities and weather conditions, present at the moment of pollution sample construction, are thoroughly collected. In fact, in Melilli, monitoring station hourly quantities of some micro-pollutants, such as sulfur oxide, nitric oxide, non-methanic hydrocarbon, ozone, sulfonyl hydrogen, and different meteorological conditions present at the moment of their observations, such as air temperature, relative humidity index, wind direction, and speed are observed and stored. Some previous studies show the evident correlation between these environmental variables and the quantity of air micro-pollution found in the samples. Because of the complete data present in the samples studied, levels of four micro-pollutants (SOx, CH_4, NMHC, NOx) in correlation with the meteorological variables previously mentioned [42, 43] are analyzed in this chapter.

Data recovered from the Melilli monitoring station during 2 weeks, more precisely 1 week in January and 1 in August 2010, have been studied, in order to observe differences of analysis results also on the basis of the different seasons of the year. Daily available recorded 'objects' described both by meteorological variables (condition attributes) and by micro-pollution quantity (decision attribute) have been considered. More than 1000 data records have been analyzed, as an example to which the RSA could be applied.

The selected condition attributes/criteria (descriptors) considered in this analysis are the hour of observation (attribute), wind speed (criterion) and wind direction (attribute), air temperature (criterion), and the relative humidity index (criterion), whereas the levels of the aforementioned micro-pollutants are the decision classes. The descriptors have been chosen because in previous studies [36, 42] they looked like some very important factors, at a local level, influencing air micro-pollution quantity.

4. Results

In spite of the fact that data samples used are restricted to a relatively short period of time (each one only 2 weeks), their analysis allowed us to obtain some interesting results, both from methodological and from operational points of view, which give an idea of the knowledge extraction (in terms of decision rules) from available data using the considered approach and the possibility to use this new method to improve air pollution management. As mentioned before, the final results are expressed in the form of 'if…, then…' decision rules, using at any time a particular (relevant) subset of attributes (reducts), according to the season and the micro-pollutant considered at each time.

In the following sections of this chapter, just some examples of decision rules obtained in our study are presented, useful for understanding and describing concisely pollution effects caused by particular combinations of conditional attribute/criteria values. Such rules, as mentioned before, are very useful in explaining the main reasons of some particular pollution events and can be also used in forecast analysis and for decision support too. The rules were chosen as the most representative among those with the highest degree of confidence, indicating the relative frequency of antecedent ('if' part of the rule) also matching the consequent ('then' part of the same rule) of the considered rule. Apart from the analysis of SOx, CH_4, and NMHC with respect to the January observations (**Tables 1–4**), the considered rules have a confidence equal to one, that means that all the objects match both the antecedent and the consequent in each rule [37, 41, 43].

These decision rules are presented in the form of tables which are very easy to read, showing in the first column the number of the rule and in the other columns the values of the conditional attributes/criteria characterizing that rule. These values are expressed as intervals with respect to attributes (corresponding to the partition of their domain) and as real numbers ('vertices' of dominance cones) with respect to criteria. The last columns of **Tables 1–3** display the confidence of each rule; in the other tables, the confidence of the rules is one. In particular, **Tables 1–4** show results from Melilli Monitoring Station during the month of January, and **Tables 5–7** show the results in Melilli Monitoring Station during the month of August. Each table concerns a different micro-pollutant.

The threshold interval values for the conditional attribute were chosen as following: hour: 0, 1, … 23 and wind direction: N (North), S (South), E (East), W (West), as main direction ±45°. On the contrary, the criteria values were automatically determined by the method applied for

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction	Confidence
1		≥11.55					0.54
2					E		0.54
3						81.9–137.7	0.46
4	13–16				E		0.58
5			≥52.3		E		0.58
6			≥52.3			81.9–117.9	0.58
7					E	81.9–117.9	0.58

Table 1. Melilli monitoring station, January 2010 SOx; threshold = 70 μg/Nm^3.

http://dx.doi.org/10.5772/intechopen.75630

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction	Confidence
8	13–17	≥9.5					0.21
9	14–17				N		0.21
10	13–17					134.2–190.2	0.21

Table 2. Melilli monitoring station, January 2010 CH_4; threshold = 950 μg/Nm³.

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction	Confidence
11		≥11					0.36
12					E		0.40
13						107.3–149.7	0.37
14					E	107.3–149.7	0.42

Table 3. Melilli monitoring station, January 2010 NMHC, threshold = 90 μg/Nm³.

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction
15					E	
16						66.4–111.5
17	11–14	≥ 10.5				
18	10–21		≥ 75.5			
19		≥ 9.7				18.8–117.9
20			≥ 49.6		E	
21			≥ 53.8			67.3–111.5
22				≤ 3.4	E	

Table 4. Melilli monitoring station, January 2010 NOx, threshold = 20 μg/Nm³, confidence = 1.

them (DRSA). Since the dominance approach is also used for the decisional attribute, pollution is reached 'by definition' depending whether or not the observed value of micro-pollutant is at least equal to the threshold value defined by law and indicated in each table. They are usually some threshold values that it is not allowed to exceed more than three times a year.

In each table, the conditional attributes/criteria are the following: Attribute 1: hour of observation; Criterion 1: air temperature (°C); Criterion 2: air relative humidity index (%); Criterion 3: wind speed (m/s); Attribute 2: wind direction; Attribute 6: wind direction (degrees) with respect to wind speed (measured by SODAR).

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction	Risk Class
23			≥72.8				B–C
24				≤4.8	E		A
25		≥31.3	≥54	≤3.9			A
26			≥37.2	≥3.3			B–C
27			≥33	≥3.7			B–C
28			≥37		N		C
29			≥31.9		N		C

Table 5. Melilli monitoring station, august 2010 SOx, confidence = 1.

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction
30	14–20		≥37.2			
31	6–14					132.5–269.1
32		≥26.3	≥37.1			
33		≥25.9	≥37.6, ≤42.9			

Table 6. Melilli monitoring station, August 2010 CH_4 threshold =845 μg/Nm^3, confidence = 1.

Rule	Hour	Air temperature	Humidity	Wind speed	Wind direction	Wind speed and direction
34					E	
35						88.9–103.8
36				≥3.7	E	
37					E	90.5–103

Table 7. Melilli monitoring station, August 2010 NOx, threshold = 20 μg/Nm^3, confidence = 1.

With respect to the decision, in **Table 5**, we consider three decisional classes, coded as following according to the law: A (Emergency), B (Alarm), C (Alert). The corresponding decision rules are expressed in terms of 'at least (≥)' or 'at most (≤)'; therefore, for example, B–C means 'at most Alarm' and A means '(at least) Emergency'. In all the other tables, the decision rules indicate if the considered threshold values are overtaken or not.

The rules in all the above tables represent only a few of several rules obtained by applying the RSA and they are presented here just as examples of easily understandable samples of the results of this analysis. All these rules are of the type 'At least' with respect to the decision, in the sense that if the antecedence is verified, the level of the corresponding micro-pollution is greater than the threshold value. We observe that the selected rules involve in the conditional part only few attributes/criteria each time.

In the following lines, some examples of how reading the decisional rules are presented from **Tables 1–7**. Rule 18 (**Table 4**): between hours 10.00 and 21.00, if the air relative humidity index is at least 75.5%, then the NOx is at least 20 μg/Nm^3, with a confidence of 1. Rule 33 (**Table 6**): if the air temperature is at least 25.9°C and air relative humidity index lies in the interval 37.6–42.9%, then the CH_4 level is at least 845 μg/Nm^3 with a confidence of 1.

5. Discussions

The decision rules concerning CH_4 and NMHC of January 2010 (**Tables 2** and **3**) have a very low confidence level; this means that the considered attributes are not sufficient to explain the phenomenon. Perhaps some attributes are missing and therefore, in order to improve this result, it would be useful to consider further attributes. This is another important methodological feature of the rough set approach, underlining that sometimes more information is needed to better describe some object in order to be able to arrive at well-founded conclusions, that is with a high degree of confidence. On the other hand, the same analysis regarding the observations of August gives very interesting results; we can observe the particular relevance of the degree of humidity in the CH_4 level (**Table 6**) and the crucial role of the wind direction and speed in NOx analysis (**Table 7**). We can also observe that sometimes (e.g. rules 12, 15, 23) it is possible to explain a result using only one attribute, that is with very short and simple decision rules. It should be remembered that a general property of the rough set approach is one that uses all conditional attributes; instead of only attributes from a reduct, we can obtain more concise rules, that is with a greater variety a fewer number of attributes in the conditional part of the rules.

With respect to SOx, we used a greater value for the threshold in January than in August both in order to present the relative pollution level more clearly and to obtain an acceptable confidence degree for same decision rules. Both the analyses of NOx (**Tables 4** and **7**) give excellent results in terms of confidence with respect to all the decision rules obtained, where we can observe a greater relevance of the attributes hour of observation, air temperature, and humidity degree in the analysis of the January data, while a crucial role in the August results is played by winds.

Actually, a first idea about the relevance of the conditional attributes can be directly revealed by the presence frequency of each conditional attribute in the decisional rules, as shown in the **Tables 1–7** (see, e.g. how important the conditional attribute for air relative humidity is in **Table 5**). Some more sophisticatedly important indices can also be computed, for example, according to the Shapley value in the cooperative games in the framework of game theory; the main idea is to compute the contribution to the quality of results by adding another attribute/criterion in the conditional part of the rules, in other words, a degree of the involvement of each attribute in all coalitions of attributes, measuring therefore also the interaction (synergy or redundancy) between the considered attributes [20] (Greco, S. et al., 2001 b). It should be observed that this kind of importance is therefore an output of the analysis within the rough set approach, and not an input information, as usually happens when we use other approaches, as, for example, weighted sum or outranking methods for the comparative evaluation of some objects.

Moreover, the results also show interesting interpretations in terms of a particular kind of trade-off. From the analysis of the couples of decision rules (26,27) (32,33) we can easily observe some cases of trade-off between the values of the couple of attributes in the classical meaning of 'compensation'. From rules (32,33), there is a relationship between air temperature and degree of humidity, in the sense that the different values of air temperature can be compensated by the different degrees of humidity obtaining the same results in terms of pollution. A similar relationship can be observed in the pair of decision rules (26,27) with respect to the degree of humidity and wind speed. This means that a certain capacity of compensation is allowed (trade-off) between the performances of a couple of attributes: a better value on one attribute is able to compensate the worst value on the other and vice-versa.

By observing the following couples of rules (5,6) (9,10) (20,21), we can see that it is possible to obtain the same results in terms of level of pollution considering the combination of one (fixed) attribute/criterion and step by step another one associated with it (example of exchangeable attributes/criteria). In other words, the same decision could be described and explained by different rules, where at each time are present different combinations (in this case, couple) of attributes/criteria that, therefore, are able to describe the same phenomenon independently one another. So, for example, from the couple of rules (5,6), it can be observed that the same result in terms of level of pollution, with the same degree of confidence, is the consequence of the degree of humidity ≥52.3 and wind direction E (rule 5) or the consequence of the same degree of humidity associated with the wind speed and wind direction between 81.9 and 117.9 (rule 6).

Another similar observation can be made comparing rules 6 and 7, where again the phenomenon of exchangeable attributes can be observed that in this case are the air humidity and the wind direction. This means that using the RSA the same effect in the pollution class assignment can be obtained as a result of a combination of an attribute/criterion value each time with other different attributes/criteria, as a particular very interesting 'qualitative substitution effect' between different attributes/criteria. The exchangeable role played independently by some conditional attributes/criteria in combination with a given level of another conditional attribute/criterion (in the previous example, the degree of humidity or the wind direction and speed) results therefore in the assignment of an object to the same decision class of pollution.

With respect to the operational aspect of this approach, it is important to emphasize how obtained results can be used to capably support the decision-maker to manage the pollution risk. Actually, the information given by decision rules can help to understand the main reasons of a pollution event, giving us the explanation of this (its 'traceability') but also for preventing or forecasting dangerous situations, very probable when meteorological conditions similar to those described by the obtained decision rules are approaching (air temperature, humidity degree, wind direction, …).

Another very interesting result using this approach concerns the information we can receive by so-called non-activated rules in improving or in deteriorating the results of a decision. See, for example, rules from **Table 5** and at levels of air relative humidity index. It can be observed that if this value is smaller than 31.9, the SOx will never be at a level higher than the threshold of $10/gr/Nm^3$. These rules, therefore, are able to give us useful information about 'critical values' of the conditional criterion air of relative humidity.

More generally, we can say that using this approach we are able to detect some threshold values of one or more condition criteria that can be considered as boundary values to be reached or to be avoided and the combination of two or more attributes/criteria that can be really dangerous for the air pollution. Of course, the meteorological variables cannot be changed by decision-makers. But the rules inferred using the rough set approach can be actually used as guidelines for forecasting in some areas particular cases of pollution events (e.g. emergency, alarm, alert), consequently giving people useful information and suggestions concerning the probable danger of air pollution.

6. Conclusions

The aim of this chapter is to give a first idea of the possibilities offered by rough sets data analysis in the field of air pollution management. In the following, we summarize its main methodological and operational contributions of this exemplary application.

From the methodological point of view, the RSA allows us to take into consideration quantitative and qualitative data, without being in need of their arbitrary transformations.

The relevance of each subset of attributes/criteria is an output of the analysis, and not an input, and therefore does not require elicitation of a priori subjective weights.

It is possible to underline the role of each attribute/criterion in terms of reducts and core; the attributes belonging to the core are indispensable, while the attributes belonging to the reducts are exchangeable with one another; the others are actually superfluous or redundant.

The significance of the results can be measured by peculiar indicators (quality, strength, support, confidence, etc.).

The results are presented in the form of 'if… then…' logical statements, decision rules expressed in simple language and very understandable for the decision-makers.

It is not necessary to remove a priori some inconsistency in the data to be analyzed, but—on the contrary—also these inconsistencies are an important piece of information about the degree of uncertainty of the decision rules inferred (certain or approximate rules, degree of confidence, etc.).

From the operational point of view, the decision rules inferred can be used immediately for managerial purposes as guidelines for preventing or warning people about the risk of air pollution (emergency, alarm, or alert situations), when the weather conditions match or are similar to those shown on the tables and to other rules not included in this chapter.

The decision rules are also able to explain the reasons of particular pollution occurrences, describing the consequences of different meteorological scenarios and their giving a traceability of possible decisions.

Moreover, the results obtained point out other relevant profiles of the phenomenon considered. They clearly show, for example, the more or less important role played by each meteorological variable in the assignment of the actions to different pollution classes, the fundamental relationships between the antecedent (attributes and conditional criteria), and the consequent (ordered decision class). Furthermore, they provide interesting information about the semantic importance of quantitative and qualitative trade-offs between attribute/criteria, that is the role of combination of different levels and/or different pollutants considered together, showing therefore also the main interaction among some meteorological factors. Finally, using the RSA it is also possible to detect some particularly interesting threshold values, of one or more condition criteria, that can be considered as boundary values to be reached or to be avoided.

The decision rules, in fact, could be the basis for the development of air-quality management strategies under the impacts of climate change, that is fundamentally a risk valuation and risk management process involving priority assessment of the impacts of climate change and associated uncertainties, including determination of air-quality targets, the selection of potential management options, and identification of effective air-quality management strategies through decision-support models.

The simple application of the method presented in this chapter shows how it can effectively help decision-makers in making appropriate responses to climate change, since it provides an integrated approach for climate risk assessment and management when developing air-quality management strategies. The risk-based decision-making framework can also be applied to develop climate-responsive management strategies for the other environmental dimensions and appraise costs and benefits of future environmental management policies.

Like any study, this could be improved and a more in-depth study can be carried out. For example, the original database could be enlarged, both in time limits and with reference to the variables considered. If we take into consideration data concerning different years or places, and analyze them by using the same methodology, we can, for instance, eliminate the peculiar effects related occasionally to atypical weather conditions. Moreover, if we extend the analysis to other meteorological variables we could obtain decisional rules which are sometimes easier, more intuitive, and more precise than those obtained by using a smaller number of descriptors.

Author details

Agata Matarazzo

Address all correspondence to: amatara@unict.it

Department of Economics and Business, University of Catania, Catania, Italy

References

[1] Chen KS, Ho YT, Lai CH, Chou Y-M. Photochemical modeling and analysis of meteorological parameters during ozone episodes in Kaohsiung, Taiwan. Atmospheric Environment. 2003;**37**:1811-1823

[2] Vukovich FM, Sherwell J. An examination of the relationship between certain meteorological parameters and surface ozone variations in the Baltimore-Washington corridor. Atmospheric Environment. 2003;**37**:971-981

[3] Bojkov RD. Surface ozone during the second half of the nineteenth century. Journal of Applied Meteorology and Climatology. 1986;**25**:343-352

[4] Premec K. Ultraviolet solar radiation in the northern Adriatic. Croatian Meteorological Journal. 2002;**3**:79-88

[5] Xue M, Droegemeir V, Wong V. The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. Meteorology and Atmospheric Physics. 2000;**75**(3/4):161-193

[6] ARPA (Agenzia Regionale per la Protezione dell'Ambiente) (2002). Siracusa, Environmental Report 2002. Siracusa, 2002

[7] Decreto Assessoriale (Territorio e Ambiente) 14 giugno 2006. Approvazione del piano di azione con gli interventi di prevenzione dell'inquinamento atmosferico dell'area ad elevato rischio di crisi ambientale della provincia di Siracusa. Gazzetta Ufficiale della Regione 2006.25.2041

[8] Armendia A. A study of correlation between atmospheric stability and atmospheric pollution in Salamanca (Spain). In: Proc. 8th World Clean Air Congress "Man and His Ecosystem", The Hague, Holland, 11-15 September 1989; 1989

[9] Di Lorenzo A. New technologies in air pollution prevention and management: A case study. In: Proc 12th World Clean Air & Environment Congress and Exibition "Greening the New Millennium", vol. 2; 26-31 August 2001; Seoul Korea. 2001

[10] Hsin-Chung L. The influence of meteorological conditions to the statistical characters of air pollutants.In: Proc. 12th World Clean Air & Environment Congress and Exibition, "Greening the New Millennium", vol. 1; 26-31 August 2001; Seoul, Korea. 2001

[11] Seinfeld JH. Atmosperic Chemistry and Physics of Air Pollution. New York: Wiley; 1986

[12] Biamino W, Gambadoro A, Trivero P, Zerbo G. An experimental approach to forecasting the atmospheric pollution in a complex industrial area. In: Proc. 12th World Clean Air & Environment Congress and Exhibition; 26-31 Aug. 2001; Seoul. 2001

[13] Tinarelli G, Anfossi D, Brusasca G, Ferrero E, Giostra U, Morselli MG, Moussafir J, Tampieri F, Trombetti F. Lagrangian particle simulation of tracer dispersion in the lee of a schematic two-dimensional hill. Journal of Applied Meteorology. 1994;**33**(6):744-756

[14] Munda G. Beyond welfare economics: Some methodological issues. Journal of Economic Methodology. 2016;**23**(2):186-202

[15] Pawlak Z, Slowinski R. Rough set approach to multi-attribute decision analysis. European Journal of Operational Research. 1994;**72**(1994):443-459

[16] Pawlak Z. Rough sets. International Journal of Information and Computer Science. 1982;**11**:341-356

[17] Pawlak Z. Rough Sets. Theoretical Aspects of Reasoning about Data. Dordrecht: Kluwer; 1991

[18] Greco S, Matarazzo B, Slowinski R. Granular computing for reasoning about ordered data: the dominance-based rough set approach. Chapter 15. In: Pedrycz W, Skowron A, Kreinovich V, editors. Handbook of Granular Computing. Chichester: Wiley; 2008. pp. 347-373

[19] Greco S, Matarazzo B, Slowinski R. Rough sets theory for multicriteria decision analysis. European Journal of Operational Research. 2001;**129**:1-47

[20] Greco S, Matarazzo B, Slowinski R. Assessment of a value of information using rough sets and fuzzy measures. In: Chocjan J, Leski J, editors. Fuzzy Sets and Their Applications. Dordrecht: Silesian University of Technology Press; 2001. pp. 185-193

[21] Greco S, Matarazzo B, Slowinski R. Rough sets methodology for sorting problems in presence of multiple attributes and criteria. European Journal of Operational Research. 2002;**138**:110-123

[22] Greco S, Matarazzo B, Slowinski R. Decision rule approach. In: Figueira J, Greco S, Ehrgott M, editors. Multiple-Criteria Decision Analysis. State of the Art Surveys. New York, NY: Springer International Series in Operations Research and Management Science; 2005. pp. 507-561

[23] Greco S, Inuiguchi M, Slowinski R. Fuzzy rough sets and multiple-premise gradual decision rules. International Journal of Approximate Reasoning. 2006;**41**:179-211

[24] Greco S, Matarazzo B, Slowinski R. Dominance-based rough set approach to decision involving multiple decision makers. In: Greco S, Hata Y, Hirano S, Inuiguchi M, Miyamoto S, Nguyen HS, Slowinski R, editors. Rough Sets and Current Trends in Computing (RSCTC 2006). Springer LNCS 4259, Berlin; 2006. pp. 306-317

[25] Greco S, Matarazzo B, Slowinski R. Dominance-based rough set approach as a proper way of handling graduality in rough set theory. Transactions on Rough Sets, Springer LNCS 4400, Berlin; 2007. pp. 36-52

[26] Greco S, Matarazzo B, Slowinski R. On joint use of indiscernibility, similarity and dominance in rough approximation of decision classes. Research Report RA-012/98. Institute of Computing Science, Poznan University of Technology, Poznan, November 1998. In: Proc. 5th International Conference of the Decision Science Institute; July 4-7, 1999; Athens, Greece. 1998. pp. 1380-1382

[27] Greco S, Matarazzo B, Slowinski R. The use of rough sets and fuzzy sets in MCDM. Chapter 14. In: Gal T, Stewart T, Hanne T, editors. Advances in Multiple Criteria Decision Making. London: Kluwer; 1999. pp. 14.1-14.59

[28] Blaszcynski J, Greco S, Matarazzo B, Slowinski R, Szelag M. jMAF—Dominance-based rough set data analysis framework. In: Rough Sets and Intelligent Systems—Professor Zdzislaw Pawlak in Memoriam Intelligent Systems Reference Library. Berlin Heidelberg: Springer Verlag; 2013. pp. 185-209

[29] Clasadonte MT, Matarazzo A, Pappalardo N. Rough set analysis for urban air pollution control system. In: Book of Abstract of 22nd European Conference on Operational Research—EURO XXII; July 8-11, 2007; Prague. 2007

[30] Clasadonte MT, Matarazzo A, Zerbo A. Analisi della concentrazione di PM10 in zona ad alto rischio industriale. PM10 analysis in a High-risk Environmental Area, La termotecnica", 6, Luglio-Agosto 2004;**2004**. pp. 44-48

[31] Clasadonte MT, Matarazzo A, Pappalardo N, Collura F, Cortina G, Toscano FP, Zerbo A. Rough set analysis applied to the study of air pollutants in urban areas. In: Proc 13th World Clean Air and Environmental Protection; 22-27 August 2004; London, UK. 2004. pp. 324-330

[32] Greco S, Matarazzo B, Slowinski R. Dominance-based rough set approach to decision under uncertainty and time preference. Annals of Operations Research. 2010a;**176**:41-75

[33] Greco S, Matarazzo B, Slowinski R. Dominance-based rough set approach to interactive evolutionary multiobjective optimization. In: Greco S, Marques Pereira RA, Squillante M, Yager RR, Kacprzyk J, editors. Preferences and Decisions: Models and Applications. Studies in Fuzziness 257. Berlin: Springer; 2010. pp. 225-260

[34] Greco S, Matarazzo B, Slowinski R, Stefanowski J. An algorithm for induction of decision rules consistent with the dominance principle. In: Ziarko W, Yao Y, editors. Rough Sets and Current Trend in Computing (RSCTC'2000). Berlin Heidelberg New York: Springer Verlag; 2000. pp. 304-313

[35] Matarazzo A. Air micro-pollutants analyzed through rough set approach. Progress in Industrial Ecology. 2016;**10**:382-897

[36] Matarazzo A, Clasadonte MT, Pappalardo N. An application of dominance based rough set approach to analyse factors affecting interannual variability of ozone concentrations in urban areas. In: Book of Abstract of 22nd European Conference on Operational Research—EURO XXII; July 8-11, 2007; Prague. 2007. p. 168

[37] Matarazzo A, Clasadonte MT, Zerbo A, Pappalardo N. Innovative tecniche multicriteriali applicative all'analisi dei BTX in atmosfera. New Multicriteria Tecnhiques Applied to BTX Analysis of Air Pollutants. In: Proc I Congresso Seriale "Tematiche Ambientali – Qualità dell'area Nelle Aree Industriali Ed Urbane"; 16 Dicembre 2004; Udine. 2004

[38] Slowinski R. Intelligent Decision Support. Handbook of Applications and Advances of the Rought Sets Theory. Dordrecht: Kluwer Academic Publishers; 1992

[39] Matarazzo A, Cavallaro F, Zerbo A, Ciraolo L. La gestione della qualità dell'aria in zona industriale ad alto rischio ambientale: una applicazione dei rough set all'analisi degli inquinanti. Air Quality Management in a High-risk Environmental Area: An Application of Rough Set Analysis to Air Pollutians. In: Proc XIX Congresso Nazionale di Merceologia, Vol.II; 27-29 settembre 2000; Sassari-Alghero. 2000

[40] Matarazzo A, Clasadonte MT, Zerbo A. Ozone analysis with dominance based rough set approach in a high-risk environmental area. In: The 14th International Conference "Air

Quality-Assessment and Policy at Local, Regional and Global Scales"; 6-10 October 2003; Dubrovnik, Croatia. 2003. pp. 287-295

[41] Matarazzo A, Clasadonte MT, Zerbo A. Rough set analysis to manage urban air pollution control system. In: Proc of the NACA IUAPPA Conference "An International Perspective on Air Quality Management Issues in the Developing World with Emphasis on Africa"; 1-3 October 2008; Nelspruit (South Africa). 2008. pp. 1-8

[42] Matarazzo A, Pappalardo N, Clasadonte MT. La Rough Set Analysis per l'individuazione di Indicatori Sintetici di Performance Ambientale Nel Comparto Aria Dei Centri Urbani, the Rough Set Analysis to Achive Air Quality Performance Synthetic Indicators in Urban Centers. In: Proc Primo Incontro Sui Modelli Matematici Applicati Alle Gestione Ambientale; 26-27 Febbraio 2007; Milazzo. 2007

[43] Matarazzo A, Clasadonte MT, Ingrao C. The (dominance based) rough set approach applied to air pollution in a high risk rate industrial area. Environmental Engineering and Management Journal. 2018;**17**(3):1-8