1. INTRODUCTORY ANALYSIS

Air pollution impacts are very difficult to be quantified, since they include consequences at (i) individual, like human health (from respiratory to heart and brain problems), (ii) regional, like acid rain, eutrophication, effects on wildlife, and (iii) planetary (e.g. global warming and depletion of ozone layer) level. Moreover, these impacts perform interactively, other directly (e.g., the depletion of ozone layer on human health) or indirectly via a food chain (e.g. eutrophication, due to high amounts of nitrogen present in certain air pollutants, may develop in great extent on sea’s surface and turn itself into algae and adversely affect fish, plants, and animal species). In all these cases, the main independent/explanatory variable is the pollutant concentration, which, in its turn, depends on emission sources and sinks in the area under consideration as well as on the pollutant material balance over this area. The critical pollutant concentration $P_{cr}$ can be estimated by minimizing direct environmental total cost $C(P) = C_1(P) + C_2(P)$, where $C_1$ is the indirect environmental cost (measured in terms of hospital admissions for respiratory problems, expressed in monetary units by using the national health service payment list) and $C_2$ is the economic cost for pollutants abatement at source. The first partial cost is an increasing function of $P$ with an increasing rate (i.e. $dC_1/dP > 0, d^2C_1/dP^2 > 0$), since health is aggravated disproportionally in the region of high $P$ -values. The second partial cost is a decreasing function of $P$ with an increasing algebraic or a decreasing absolute rate (i.e., $dC_2/dP < 0, d^2C_2/dP^2 > 0$ or $d|dC_2/dP|/dP < 0$), since the lower the pollutant released
Fig. 1a. Dependence of the indirect and the direct/economic partial costs $C_1$ and $C_2$, respectively and shifting of the critical pollution concentration from $P_{cr}$ to $P'_{cr}$ in case of deep/lasting economic recession crisis.

Fig. 1b. Dependence of the indirect and the direct/economic partial costs $C_1$ and $C_2$, respectively and shifting of the critical pollution concentration from $P_{cr}$ to $P''_{cr}$ when the cumulative effect of air pollution on human health is taken into account.

to atmosphere (i.e. the higher the abatement of releases) the higher the cost, again disproportionally due to the validity of the well established in Economics Law of marginal or differential Returns (LDR). Evidently, the critical value $P_{cr}$ is the equilibrium point of the tradeoff between the partial costs $C_1$ and $C_2$, where the marginal costs $MC_1 = dC_1/dP$ and $MC_2 = |dC_2/dP|$ are equal.

Although this optimization procedure is right from a conceptual point of view, it may be complicated in practice, because of the intervention of exogenous factors causing change of the partial cost $C_1$ and consequent shifting of the initial $P_{cr}$ value. E.g., in case of economic crisis, hospital admissions may decrease giving underestimated values for $C_1$ for the same pollutant concentrations, since people prefer treatment while staying at home in order to avoid additional expenses; as a result, the corresponding curve $C_1$ moves downwards to $C'_1$, becoming more flat, since the respective influence is more expressed in the region of higher $P$-values, as a result, $P_{cr}$ is shifting to $P'_{cr}$, where $P'_{cr} > P_{cr}$, as shown in Fig. 1a. On the contrary, the $C_2$-curve is moving upwards to $C''_1$, becoming steeper, if the cumulative effect of air pollution on human health is taken into account, leading to overestimation of the current cost (in order to counterbalance respective underestimation in the previous time periods [1,2]); as a result, the initial $P_{cr}$-value is shifting to $P''_{cr}$, where $P''_{cr} < P_{cr}$ (see Fig. 1b). Since the vectors $(P'_{cr} - P_{cr})$ and $(P''_{cr} - P_{cr})$ have opposite direction, the final position of the $P_{cr}$, in comparison with its initial value, depends on the form of the partial cost functions and their estimated parameter values.

2. METHODOLOGY

For taking into account the exogenous factors mentioned in the previous paragraph, as well as hidden assumptions and inadequacies/pitfalls, we have designed/developed a methodological framework under the form of an algorithm procedure, including the following 23 activity stages and 7 decision nodes (for their interconnection, see Fig.2).

1. Description of the spatiotemporal domain under consideration.
2. Registration of population growth per administrative district and industrial development per zone.
3. Collection of meteorological, air pollution, and
urban activities/traffic data.

4. Mapping on the respective layers of a Geographical Information System (GIS), properly designed to give/receive/store (in retrievable format) information on a real time basis [3-5].

5. Addition of a separate layer with hospital sites and other health services providers.

6. Design of environmental economic functions aiming at minimizing total cost.

7. Identification/replacement (or monitoring/fixing) of the inadequate AQMSs.

8. Localization of proper candidate alternative places for installation of new equipment.

9. Multicriteria ranking of place(s) and installation/trial of the new equipment.

10. Preliminary overall synthesis for quantifying the air pollution impacts.

11. Fault Tree Synthesis (FTS) by setting ‘inadequacies/pitfalls’ as a top event for diagnosing its possible/probable causes [6,7].

12. Assignment of possibility grades of dependence/importance [8] between events (in fuzzy version to count for uncertainty) by experts, and Fault Tree Analysis (FTA), according to [9,10].

13. Localization/Identification of ‘possible’ (or ‘probable’, in case that probabilities are used instead of fuzzy sets) ultimate causes or intermediate ones, if relevant data are available and reliable.


15. Remedial proposals, including external quality assessment schemes in monitoring and improving the categorization/standardization process.

16. Recognition of certain categories of diseases (especially commutable ones, like influenza), which should be excluded, and final registration of hospital admissions, recognized as possibly related to air pollution.

17. Time series analysis and elimination of pitfalls.

18. Truncation or elongation, accordingly.


20. Sensitivity/Robustness analysis of the total optimum by simulating real conditions.

21. Limitation of the spatiotemporal range of validity and/or change of the respective significance level.

22. Development/enrichment/updating of the internal Knowledge Base (KB).

23. Design/development of an Intelligent Agent (IA) searching within external KBs, according to [11].

A. Is the AQMSs network providing adequate data for quantifying the independent/exploratory variables at the required information granularity level?

B. Are all the AQMSs of the same satisfactory technology in order to provide reliable/comparable measurements?
C. Is the network of AQMSs covering completely and equally well the area described in stage 1?
D. Is the probability of occurring the top event quite low (i.e. within a range of pre-set acceptable values)?
E. Are the hospitals (and the rest health services/institutions) capable to provide properly categorized/standardized data on emergency and programmed admission of people with asthmatic-type problems?
F. Are the results of this analysis satisfactory?
G. Is the series too long or too short?
H. Is it robust?

The fuzzy sets used as possibility grades of dependence/importance for performing FTA, according to stage 12, enter the respective calculations as crisp number after defuzzification in order to simplify the algorithmic procedure, especially as regards sensitivity/robustness analysis of the total optimum under applicable conditions (see stage 20). In Fig. 3, an example is presented in case that three experts assign three different triangular LR fuzzy numbers on a certain dependence index; the crisp value $m_n$ results after determining the mean fuzzy set and defuzzifying this set by means of the centroid method.

![Fig. 3. Transformation of three fuzzy input estimators into a mean crisp number (obtained after defuzzification with the centroid method) for the final cause 1.3.3 (uneven coverage of the domain under surveillance with the AQMSs network) by adopting the assumption consensus in input.](image)

3. IMPLEMENTATION

The tree synthesized, for the implementation case example, i.e., “Increased Uncertainty of Pollution Impact Quantification” is shown in the Appendix, whereas the description of the events is presented below:

1.1 Improper spatiotemporal data acquisition.
1.1.1 Spatial Heterogeneity.
1.1.1.1 Permanent heterogeneity.
1.1.1.2 Accidental heterogeneity.
1.1.1.3 Periodical heterogeneity.
1.1.1.4 Extraordinary limited or extended area under examination, according to stage 1 in the flowchart depicted in fig.2.
1.1.2 Temporal heterogeneity.
1.1.2.1 Two or more predictor variables in the multiple regression model are highly correlated (i.e., the value of at least one independent variable can be estimated from at least one other independent variable with significant accuracy degree), weakening the predictive power of at least one independent variable [12].
1.1.2.2 The regression residuals/errors exhibit a time series structure, violating the basic assumption of error independency, influencing model explainability/predictability.
1.1.2.3 Time series too long, in which case the ceteris paribus clause loses progressively its validity, in the time course.
1.1.2.4 Time series too short in which case statistical analysis becomes progressively unreliable with the data (sample size) decrease.
1.2 Lack of data representativeness
1.2.1 Economic recession
1.2.1.1 Under the form of a monotonic function (demand side analysis).
1.2.1.2 As a phase of economic cycle in the medium run (demand side analysis).
1.2.1.3 Specifically as reduction of the budget for hospitals and the rest health services operability (supply side analysis).
1.2.2 Incomplete categorization of people admissions into hospitals and relevant health centers, mainly as regards the real event caused their decision to ask for medical help/support/treatment.
1.2.3 Inadequate sensitization of people concerning the asthmatic-type diseases and the relation to air pollution.
1.2.3.1 Lacking of diffusion of information into the public at proper granularity level.
1.2.3.2 Low public awareness in general, not necessarily/exclusively in relation with understanding the situation under consideration.
1.2.4 Measures taken on individual basis to avoid human exposure to pollutants.
1.2.4.1 Measures taken on systematic way to avoid human exposure to pollutants.
1.2.4.2 In case of emergency conditions.
1.2.5 Measures taken on societal basis to avoid human exposure to pollutants.
1.2.5.1 Measures taken on systematic way to avoid human exposure to pollutants.

1.2.5.2 In case of emergency conditions.

1.3 Inadequacy/error in the network of AQMSs

1.3.1 Insufficient definition/determination of metrological validity.

1.3.2 Inadequate nodes of the network (in the measuring or/and the communication subsystems).

1.3.2.1 Equipment out of order.

1.3.2.2 Unreliable calibration (e.g., due to lack of maintenance).

1.3.3 Uneven coverage of the respective GIS domain.

1.3.4 Heterogeneity in methods/technology applied in measurements.

1.3.4.1 Usage of equipment of varying technology, even of different ‘generation’.

1.3.4.2 Application of multimodal methodology to assign pollutant concentration values within a geographical grid in the respective GIS layer.

For the assignment of a respective grade on the event 1.2.3.2, we used the chi-square test by questioning 117 people in Athens during an episode of high dust concentration. The two main questions were: (I) According to your empirical knowledge/opinion, is your respiratory system a sensitive one, especially when breathing dust, even at low/medium concentration level in open air? (II) Supposing you were at workplace with moderate dust concentration and without formal obligation to wear a mask, with half of your colleagues usually taking such a precautionary measure, would you undertake the initiative to wear a similar mask?

In the sample of 117 people, 31 answered ‘yes’ in the question and the rest ‘no’ forming two groups, named A and B, respectively. In group A, 3 were wearing a sanitary mask, while in group B only 1 was wearing such a mask. The results of the chi-squared test are 5.0031 for the original and 2.7567 for the corrected (after Yates) test. Since the critical value is 3.84 for one degree of freedom at significance level 0.050, we conclude that the null hypothesis (stating that the criterion divisionis of what people think about their sensitivity against air pollution is not adequate to make them wear a sanitary mask, so the groups A and B behave similarly in this respect) is rejected only if the uncorrected $\chi^2$ value is taken into account. Since the Yates' correction gives the more reliable value $2.7567 < 3.84$ (leading to no-rejection conclusion), we considered the significance level 0.025, where the critical value is 5.02 and consequently the null hypothesis is not rejected independently of correction. Therefore, In relation with the event 1.2.3.2, the conclusion is that public awareness is not adequate and further measures should be taken for sensitization of people (event 1.2.3).

In an attempt to identify the causes for this low awareness, we asked the reviewees about it and the answers might be categorized as follows: ‘because others are not wearing a sanitary mask’ and ‘providing that I wore one, others would think that I have a disease, potentially contagious, and they would avoid me’. In an additional attempt to further investigate quantitatively these qualitative answers, we analyzed the numerical results obtained through the question (II): from the groups A and B, 30 and 75 people gave positive answers, respectively, resulting to corresponding chi-square uncorrected and corrected values 2.2648 and 1.3449; since both are < 3.84, the ‘imitation effect’ is confirmed in its positive version while its negative version was previously confirmed, according to the analysis presented in the paragraph above.

The categorization of people admissions into hospitals and relevant health centers, mainly as regards the real event caused their decision to ask for medical help/support/treatment, according to the event description given in 1.2.2 has been investigated by means of correlating the concentrations of two pollutants as measured by the same AQMS. In Fig. 4, the concentration co-variation of the separately measured air pollutants, Nitrogen Oxides (NOX) and Nitrogen Dioxide (NO2) by the AQMS ‘Liosia’ in Athens, Greece, during the photochemical episode of 26th July 2001 (taken as a reference episode) is shown.

![Fig. 4. Concentration co-variation of the separately measured air pollutants, Nitrogen Oxides (NOX) and Nitrogen Dioxide (NO2) at the AQMS ‘Liosia’ in Athens, Greece, during the photochemical episode under consideration; the coefficients](image-url)
of determination $R^2$ for the linear and the parabolic model are also quoted.

Fig. 5. Concentration co-variation of the air pollutants, Nitrogen Oxides (NO$X$) and Ozone (O$_3$), during the 24h photochemical episode under consideration (the six hour time-lag is clearly shown).

Referring to two distinct pollutants (in contrast with the nitrogen oxides which are chemically correlated, although separately measured), we have examined the concentration co-variation of the air pollutants nitrogen oxides (NO$X$) and ozone (O$_3$), during the same 24h photochemical episode (the six hour time-lag is clearly shown in Fig. 5). Consequently, any categorization attempt performed by the hospital personnel would be incomplete, because it is not possible to discriminate the real cause of the respiratory problem (or the weighted contribution of each pollutant to the creation of such a problem).

As regards the examination of the coverage of the geographical domain under consideration, according to 1.3.3, we have constructed a Pearson correlation matrix of ozone measurements for the existing main seven AQMSs of the Greater Athens Area (GAA), shown in Table 1. The numerical results presented in this Table indicate that the AQMS ‘Peristeri’ might be considered as redundant; meaning that the gathered data from this AQMS might be predicted satisfactorily at a network level by taking into account the rest measurements;

Table 1: Pairwise Pearson correlation matrix of ozone measurements for the existing main seven AQMSs of the GAA (photochemical episode on 30th July 2013).

<table>
<thead>
<tr>
<th></th>
<th>PER</th>
<th>LYK</th>
<th>GEO</th>
<th>LIO</th>
<th>AGP</th>
<th>N.SM</th>
<th>ELE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PER</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LYK</td>
<td>0.9366</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEO</td>
<td>0.9593</td>
<td>0.8761</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIO</td>
<td>0.9253</td>
<td>0.9048</td>
<td>0.8801</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGP</td>
<td>0.5996</td>
<td>0.5053</td>
<td>0.6378</td>
<td>0.4390</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.SM</td>
<td>0.8432</td>
<td>0.8075</td>
<td>0.9034</td>
<td>0.8655</td>
<td>0.5516</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>ELE</td>
<td>0.7545</td>
<td>0.7686</td>
<td>0.7892</td>
<td>0.8477</td>
<td>0.3881</td>
<td>0.9009</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

as a result, the equipment of this station might be relocated to another site, in order to ensure the uniformity/evenness coverage from a metrological point of view.

Fig. 6. Concentration co-variation of the air pollutant Ozone (O$_3$), during the 24-hour photochemical episode under consideration, as registered in certain AQMSs of the Athens Metropolitan Network.

For further supporting this view we introduced the dimension of time in following the concentration of the same pollutant during the episode under consideration, as measured by the most AQMSs; we observed that the fluctuations shown (see Fig. 6) behave in a similar way, so that prediction is feasible at a significant level. Surprisingly, the same phenomenon is observed even
when usual/normal conditions prevail, as regards the same pollutant (see Fig. 7).

4. DISCUSSION

The metrological validity, mentioned in the event 1.3.1 of the fault tree described in the Implementation Section and presented in the Appendix, depends on the precision (repeatability and reproducibility) and the accuracy of the measurement of the air pollutant. In the case of representing the measurement as a triangular fuzzy number we might define accuracy as the distance between the reference or ‘true’ value and the crisp number representing the fuzzy number after defuzzification, as shown in Fig. 8.

![Fig. 8. Accuracy as the proximity of measurement results (presented as a fuzzy set, giving the corresponding crisp number by defuzzification with the centroid method) to the reference or ‘true’ value in relation with respective precision range, determined via a repeatability/reproducibility Standard.](image)

Normalizing metrological validity and representing it with Index $I$, we can estimate the optimal value $I_{opt}$ by maximizing total benefit $B(I) = B_1(I)+B_2(I)$, where the partial benefits $B_1$ and $B_2$ depend on mitigating air pollution consequences/impacts (to health, materials, and human activities) and reducing the economic expenditure, respectively. The first partial benefit is an increasing function of $I$ with a decreasing rate (i.e., $dB_1/dI<0$, $dB_1^2/dI^2<0$), because of the Law of (differential or marginal) Diminishing Returns (LDR). On the other hand, the second partial benefit is a decreasing function of $I$, with a decreasing algebraic or an increasing absolute rate (i.e., $dB_2/dI<0$, $dB_2^2/dI^2>0$) or $d|dB_2/dI|/dI>0$, because of the validity of the LDR, which is more expressed in the region of higher $I$ values. Evidently, $I_{opt}$ is the abscissa of $B_{max}$, where the marginal benefits $MB_1=dB_1/dI$ and $MB_2=dB_2/dI$ are equal since this point represents equilibrium in the trade-off between $B_1$ and $B_2$. In case of installing higher precision/accuracy and more expensive equipment, the $B_2$-curve will move downwards because of cost increase in the short/medium run, especially during the depreciation period, becoming steeper, since the deviation from the initial $B_2$-curve is expected to be higher in the region of higher $I$-values (implying higher technology equipment too); as a result, $I_{opt}$ is shifting to $I_{opt}'$, where $I_{opt}'<I_{opt}$ (see Fig.9a). In the medium/long run (after depreciation), the $B_2$-curve is expected to move upwards because of cost reduction (especially as regards the part corresponding to capital cost), becoming more flat since the deviation from the initial $B_2$-curve is expected to be higher in the region of higher $I$-values; as a result, $I_{opt}$ is shifting to $I_{opt}''$, where $I_{opt}''>I_{opt}$ (see Fig. 9b). Since the vectors $(I_{opt}''-I_{opt})$ have opposite directions, the final optimal value of $I$, as well as the final $B_{max}$ value depend on the form of the partial benefit functions and their estimated parameters. On the other hand, the pollutant monitoring methods and the respective equipment, as well as their standards, change in the time course. The changes create successive ‘generations’ of measuring methods.
Fig. 9b. Dependence of the partial benefits $B_1$ and $B_2$ (environmental and economic, respectively) on the validity index $I$ (representing precision and accuracy of the air pollution measurements) and shifting of $I_{opt}$ to $I''_{opt}$ when introducing/installing next generation equipment, implying lower capital cost in the long/medium run after the depreciation period; the thick line in the upper diagram is the locus of the $B_{max}$ points.

5. CONCLUSIONS

Fault Tree Synthesis/Analysis (FTS/A) has been successfully implemented for depicting/determining inadequacies/pitfalls in quantifying air pollution impacts. The original design of FTS was heavily based on a deductive procedure by setting as top event the “increased uncertainty of pollution impact quantification”, which is quite wide in content, in order to (i) ensure/guarantee extensibility when/as required, and (ii) advanced statistical processing of spatiotemporal data and empirical modelling. Special emphasis was put on cases where (a) the regression residuals/errors exhibit a time series structure, and (b) two or more predictor variables (in the multiple regression model), are highly correlated, weakening the predictive power of certain independent variables (increasing also uncertainty about redundancy), as it was observed in the case of data processing for Athens, Greece. The subsequent FTA was based on an inductive procedure through the assignment of probabilities or weight/possibility values by experts in a fuzzy version to count for uncertainty.

Both FTS and FTA were successfully embedded with an algorithm serving the main purpose of the present work not only by suggesting the ‘top event’ but also by interconnecting the ‘final causes’ harmonically, aiming at the complete coverage of the whole geographical area under investigation, without excluding further extension in the time course, giving a dynamic character to the whole project.

Last, the similarity of the trade-off optimization methods presented for determining critical pollutant concentration and estimating the optimal validity index (incorporating/combining precision and accuracy of the air pollution measurements), forms the necessary background for developing a consistent platform for solving relevant environmental problems.

ACKNOWLEDGMENT

The publication of this paper has been partly supported by the University of Piraeus Research Center.

References:


Fig. 10. The fault tree synthesized and subsequently analysed, according to stages 10-13 of the algorithmic procedure developed for the determination of inadequacies/pitfalls in quantifying air pollution impacts based on a fault tree analysis (FTA) approach, including remedial proposals with external quality assessment schemes in monitoring and improving the categorization process; the events 1.2.3.2, 1.2.2, 1.3.3 are further analysed with empirical evidence/quantification, using reference data and numerical results obtained for Athens, Greece, through samples representative of usual/normal and emergency conditions as regards air pollution.