TRANSPHORM submitted deliverable

Description and development of the uncertainty methodology

Applicability of models and uncertainty analysis

D4.2.1

Partners: AUTH, UH, TNO

Prepared by: Denis Sarigiannis, Periklis Kontoroupis, Spyros Karakitsios (AUTH)

Nature: Report

Dissemination level: Public

Delivery date from ANNEX I: Month 24 (31 December 2011)

Actual delivery date: 25 April 2013 for internal review, 28 May 2013 revised

Notes:

TABLE OF CONTENTS

1 ABSTRACT **3**

2 INTRODUCTION **4**

2.1 DATA QUALITY 4

2.2 UNCERTAINTY IMPLEMENTATION TO THE TRANSPHORM PROJECT 4

3 A STEP-WISE APPROACH TO THE UNCERTAINTY ASSESSMENT **6**

3.1 QUALITATIVE UNCERTAINTY ASSESSMENT 6

3.1.1 Sources of Uncertainty 7

3.1.2 Characterising uncertainty 9

3.1.2.1 Assessment of direction and magnitude of uncertainty 9

3.1.2.2 Assessing the magnitude of influence of the source on the result(s) 10

3.1.2.3 Assessing the knowledge base of the source 11

3.1.2.4 Assessing the subjectivity of choices of the source 12

3.2 REPORTING THE QUALITATIVE UNCERTAINTY CHARACTERIZATION 14

3.3 QUANTITATIVE UNCERTAINTY ASSESSMENT TOOLS 14

3.3.1 TIER – 1 Level 15

3.3.1.1 Local Sensitivity Analysis 15

3.3.1.2 Fuzzy methods 17

3.3.2 TIER – 2 Level 18

3.3.2.1 Error Propagation 18

3.3.2.2 Monte Carlo Simulations (1D - 2D) 19

3.3.2.3 Fuzzy Monte Carlo Simulation (FMCA) 21

3.3.2.4 Stochastic Response Surface Methods 21

3.3.2.5 Global Sensitivity 22

3.4 METRICS OF UNCERTAINTY 23

3.5 TOOLS / SOFTWARE 24

4 CONCLUSIONS-RECOMMENDATIONS **26**

5 APPENDIX **30**

6 REFERENCES **32**

1 ABSTRACT

In principle, uncertainty analysis is quite a simple process. Characterisation of uncertainty can include either qualitative or quantitative evaluations, or a combination of both. The approach can also be tiered, that is, the analysis can begin with a simple qualitative uncertainty characterisation and subsequently progress to semi-quantitative and finally a complex quantitative assessment. The latter could follow when a lower tier analysis indicates a high degree of uncertainty for certain identified sources, the sources are highly influential to final result(s) and sufficient information and resources are available to conduct quantitative uncertainty assessment. This is not to suggest that quantitative uncertainty analyses should always be performed in all environmental assessments. The decision regarding the type of uncertainty characterisation performed depends on the scope and purpose of the assessment, on whether the selected analysis will provide additional information to the overall health impact assessment, whether sufficient data are available to conduct a complex quantitative analysis, and if time and resources are available for higher tier characterisations. Uncertainty assessment tools are presented in a qualitative and a quantitative form, categorized into four tiers (levels). The qualitative part summarizes all sources in a matrix, annotating direction, level of uncertainty and appraisal of the knowledge-base robustness. The quantitative tools reviewed include, sensitivity methods, error propagation techniques via Taylor expansion, Monte Carlo Modelling and Fuzzy methods.

In this report, both qualitative and quantitative tools are used with scope to identify and minimize uncertainty in the Transphorm project. All methods presented are 'tailored' according to the Tier level considered in a coherent and transparent manner. Effectively, for the full chain assessment, areas of interest include the transport and emission sources, the air quality modelling, the exposure and the health impact assessment of transport related PM. The proposed methodology is in accord to the WHO guidelines (2008) on characterizing and communicating uncertainty in exposure assessment studies and the full chain assessment is sussessfully implemented for the uncertainty needs of HEIMTSA and INTARESE projects. It is noted that different tiers of the uncertainty tools presented herein are selected based on the data availability and the complexity of the processes modelled.

2 INTRODUCTION

In this review, uncertainty assessment is investigated with regard to the objectives of each assessment level (Tiers 0, 1 and 2), as presented below:

On the Tier 0 level, a qualitative evaluation of the uncertainties and an identification of the sources of uncertainty are investigated.

On the Tier 1, interest is on the gross (overall) result addressing uncertainty from the different routes, sources and pathways. Uncertainty should be quantified in relation to the 'direction' of the overall outcome. It is noted that such assessment should not evaluate in detail the conservatism of the models used, as this is a known assumption, but it should address those elements that will provide additional uncertainty for the current assessment and to determine the future impact on the outcome.

On the Tier 2 level uncertainty is investigated at a refined level, keeping the required information (as in Tier 1) qualitatively the same. The uncertainty present in this level should be accurately quantified and minimized based on elaborate statistical techniques. It is recognized that the contribution of routes, sources and pathways are used to delineate the more detailed assessment and focus on the most important routes, sources and pathways.

2.1 Data Quality

Data in the assessment (concentrations in the contact media, individual exposure patterns etc) are uncertain to some extent with varying noise levels, according to the limitations of the measurements made. It is hence important to address data quality with respect to: the degree of relevance (to the overall assessment), the accuracy, the integrity and transparency. These characteristics are in accordance to the WHO (2008) guidelines on investigating the overall robustness of the assessment. In addition, attention should be given to the Tier considered: on a Tier 1 assessment incomplete information could be adequate, where as for a Tier 2, the most accurate and complete data should be used to produce robust answers.

2.2 Uncertainty Implementation to the Transphorm project

A pilot fully chain assessment is under progress in AUTH in collaboration with the University of Stuttgart, focusing on the road transport only, University of Stuttgart provided the gridded PM_{10} and $PM_{2.5}$ emissions together with their associated standard deviation and are further classified by the vehicle category, network, vehicle type and technology and fuel type. It is noted, uncertainty to activity data is omitted after expert elicitation.

This pilot full chain component is comprised by the assumed prior distributions, seen next by the numbered steps and by implementing sampling techniques for the Monte Carlo Simulation to follow.

The steps already taken are described next :

Firstly, based on the emission data, prior normal distributions per grid cell per country are generated and their shares (per vehicle category, network for each country) were generated following non-parametric distributions.

Secondly a statistical model inference (linear regression) is developed relating PM_{10} emission to CMAQ gridded concentration estimates under the assumptions of uniform, time-invariant emission profiles.

Thirdly, spatially distributed lognormal distributions to the estimated CMAQ, SILAM and LOTUS –EUROS models are currently under investigations, constrained by the computed emission estimates and their associated spatial link.

Fourthly, log-normal distributions are fitted to the concentration response functions relating mean and their associated confidence intervals, for a number of health end points to PM_{10} and $PM_{2.5}$.

3 A STEP-WISE APPROACH TO THE UNCERTAINTY ASSESSMENT

Uncertainty assessment is in accordance to a step-wise approach based on the work from WHO (2008) with the objective to address uncertainty from the aggregate exposure to chemical substances in the environment. This was achieved by dividing the level of uncertainty into four tiers:

- Tier 0: Qualitative evaluation of the uncertainties – identification of the sources of uncertainty.
- Tier 1: Preliminary quantitative analysis of the range of uncertainty in key parts of the assessment
- Tier 2: Monte Carlo analysis (or error propagation) to produce distribution of output.
- Tier 3: More elaborate statistical model in terms of Bayesian hierarchical modelling attempting to estimate the posterior likelihood of the final result.

The first two tiers are essentially related to a qualitative assessment of uncertainty having as a major goal the identification and qualitative evaluation of the key sources of uncertainty. The higher tiers (tier 2 and tier 3) entail a full quantitative assessment of uncertainty needed to evaluate the propagation of uncertainty and to formulate proposals to reduce its impact. Since this step-wise approach implies a gradual increase in the tier of sophistication and consequently a higher tier of knowledge and availability of quantitative information regarding the sources of uncertainty, the choice of the Tier to be applied will mainly depend on the availability of the necessary information. It is noted, that in this report tier 3 methods are not further elaborated.

3.1 Qualitative uncertainty assessment

The primary purpose of the qualitative uncertainty characterisation is to identify and compare the relative impacts that important sources of uncertainty may have on the final result(s) of the environmental health impact assessment. This approach is justified knowing that qualitative evaluation is the common denominator for all sources of uncertainty and given the limited data available to inform quantitative analyses, and time and resource constraints. Qualitative uncertainty characterisation also requires analysts to perform several quantitative sensitivity analyses to iteratively inform both model development and the qualitative uncertainty characterisation, where possible.

Qualitative uncertainty assessment is organised in three distinct steps, as seen in figure 1:

1. Identification of all uncertainty sources;
2. Qualitative characterisation of uncertainty comprising three dimensions:
    a. Assessing the magnitude of the influence of the uncertainty source on the result(s)
    b. Assessing the knowledge base of the uncertainty source;
    c. Assessing the subjectivity of the choice of uncertainty sources
3. Qualitative uncertainty reporting

Uncertainty analysis is iterative in the sense that only the "Assessing the magnitude of influence of the source on the result(s)" part of the second step can be performed during first iteration, followed by the third step. A subsequent iteration could consist of "Assessing the knowledge base of the source" part of the second step, followed by the third step. The added value of the second iteration lies in greater focus on the uncertainty sources which resulted in high scores during the first iteration. Similarly, the third iteration would consist of "Assessing the subjectivity of choices of the source" part of the second step followed by the third step. The same added value can also be applied between the second and third iterations.

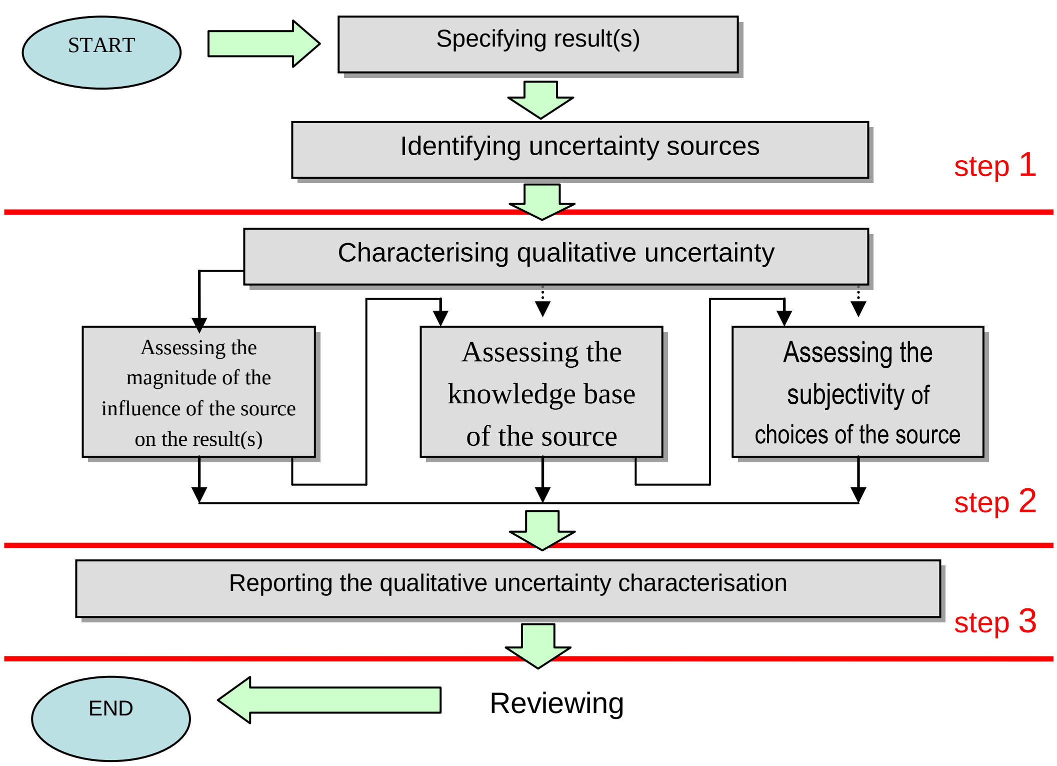

Figure 1. Three steps methodology for characterising and reporting qualitative uncertainty

3.1.1 Sources of Uncertainty

There are three main sources of uncertainty described in this document:

i) **Scenario uncertainty** refers to the description of the context (scenario setting) as a prerequisite for either modelling or measuring experimental data. It includes descriptive errors, aggregation errors, errors in selection of the assessment tier and errors due to incomplete analysis. It often includes the purpose of the environmental health impact assessment and consistency between the scenario definition and the scope and purpose of the assessment.

ii) **Model uncertainty** reflects the limited ability of mathematical models to represent the real world accurately and may also reflect lack of sufficient knowledge. It is principally associated to model boundaries, extrapolation limits, modelling errors and correlation (dependency) errors. It also includes errors due to the implementation of tools and software.

iii) **Parameter uncertainty** refers to data values that are not known with precision due to measurement error or limited observations (sampling error). Sometimes it consists of variability as an inherent property of the heterogeneity or diversity in the parameter, such as parameters expressed as a function of the entire population. Usually, variability cannot be reducible through further investigation. It is also possible for the uncertainty and variability of parameters to be combined.

Classification using the three categories defined above is not as strict as it may seem, and uncertainties may arise in overlapping areas. Thus it is sometimes difficult to reach a clear decision as to whether the uncertainty is related to the scenario, model or parameters because there may be overlaps and even expert opinions may differ. Identification of the sources of uncertainty is a matter of interpretation and depends on the clarity with which the scope and purpose of the assessment are given.

A comprehensive list of relevant sources of uncertainty composing the environmental health impact assessment should be assembled. At this stage, it is not necessary to be concerned about the quantification of individual components; the aim is to be completely clear about what should be considered. In practical terms, all potential sources of uncertainty should be identified at a first instance; judging whether a particular source is important enough to be included in the assessment can be done after evaluating the associated uncertainty at a later stage.

Table 1. Overall qualitative assessment matrix on level of uncertainty – here only the sources of uncertainty are filled in

Sources of uncertainty	***Dimensions of uncertainty***		
Policy scenario			
Conceptual model			
Mathematical model			
Parameters			

3.1.2 Characterising uncertainty

3.1.2.1 Assessment of direction and magnitude of uncertainty

Following WHO (2008) recommendations, this guidance proposes a concrete compromise of three dimensions characterisation of qualitative uncertainty. In other words, each source of uncertainty should be scored according to three different dimensions of qualitative uncertainty characterisation as shown in figure 2. Each dimension gives a complementary view on the uncertainty scoring relative to an uncertainty source. This methodology is systematically applicable for any module of the integrated environmental health impact assessment chain as well as for the interaction between modules.

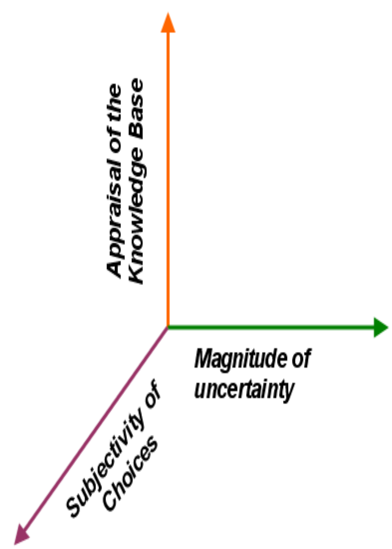

Figure 2. Three dimensions of qualitative uncertainty characterisation

The three dimensions of uncertainty are defined as follows:

1) Magnitude of uncertainty: based on three mutually exclusive criteria of direction, magnitude and data sets;

2) Appraisal of the knowledge base: based on five mutually inclusive criteria of accuracy, reliability, plausibility, scientific consistency and robustness;

3) Subjectivity of choices: based on five mutually inclusive criteria of choice space, intersubjectivity among peers and among stakeholders, influence of situational limitations on choices, sensitivity of choices to the analysts' interests and influence of choices on results.

Distinguishing between the three dimensions of uncertainty could lead to the identification of areas where improvement in the environmental health impact assessment is needed. The first dimension (*i.e.*, magnitude of uncertainty) induces the analysts to gain a better knowledge of the sources, their relationships, their correlations and thus their influence on the final result(s). It can often require a sensitivity analysis when possible. The second dimension (*i.e.*, appraisal of the knowledge base) highlights the lack of knowledge of the team of analysts and of the lack of available knowledge. The third dimension (*i.e.*, subjectivity of choices) explicitly assesses the context and the limitation of the assessment. This step consists of ranking uncertainty as *high*, *medium*, *low* or *not applicable* according to the source, to the dimension and to the audience involved. In the following three sub-chapters, the scoring according to each dimension is detailed.

3.1.2.2 Assessing the magnitude of influence of the source on the result(s)

The first dimension (*i.e.,* magnitude of uncertainty) aims to characterise how the direction and magnitude of each identified source of uncertainty influences the assessment results. Three criteria are considered for scoring the first dimension: direction (see table 1), magnitude, and data sets. The last two criteria are additional (*i.e.*, 'and' condition): outcomes and data sets (*e.g.* ranges and probability distributions).

The direction of influence indicates how the source of uncertainty is judged to affect estimated results; the estimated results is either considered to be *over*- or *under-estimated* (see table 1). When the component of uncertainty can affect the assessment results in either direction, the influence is judged as *both*. The direction of influence is characterised as *unknown* when there is no evidence available to judge the directional nature of uncertainty associated with the particular source.

Table 1. Direction of uncertainty related to the magnitude of influence of the source of uncertainty on the result(s)

Score	**criteria**
	direction
Over	the source is judged to over-estimate the final result
Under	the source is judged to under-estimate the final result

The magnitude of influence scores the overall impact of the uncertainty on the final outcome of the assessment by considering the severity of the uncertainty given by the relationship between the source of the uncertainty and the results, seen in table 2. The magnitude of uncertainty is rated *low* when it is judged that large changes within the source of uncertainty would have only a small effect on the assessment results. A designation of *medium* implies that a change within the source of uncertainty is likely to have a moderate (or proportional) effect on the results. A characterisation of *high* implies that a small change in the source would have a large effect on results.

Table 2. Scale of uncertainty related to the magnitude of influence of the source on the result(s)

Score	criteria	
	magnitude	**data sets**
Low	known	known
Medium	known	unknown
High	unknown	unknown

3.1.2.3 Assessing the knowledge base of the source

The second dimension (i.e. the knowledge base) is assessed when **(i)** a detailed list of sources has been detailed and agreed upon and **(ii)** a minimum agreement has been found on the scoring of the magnitude of uncertainty. This dimension of qualitative uncertainty characterisation scales the knowledge base uncertainty associated with each identified source using a three level scale: *low* indicates significant confidence in the data used and their applicability to the assessment; *medium* implies that there are some limitations regarding consistency and completeness of the data used or scientific evidence presented; and *high* indicates that the knowledge base is extremely limited. It comprises five criteria (see table 3) which are assessed for scoring. They can be used in an <u>exclusive</u> manner (*i.e.*, 'or' condition) at the discretion of the analysts.

Table 3. Scale of uncertainty related to the knowledge base

score	criteria				
	accuracy	**reliability**	**plausibility**	**scientific consistency**	**robustness**
Low	expected results	best available practice	plausible	extended scientific backing	results quality well analysed
Medium	results partly expected	limited consensus on reliability	acceptable	independent scientific backing	some analysis of the results quality
High	unexpected results	no discernible rigor	fictive or speculative	no scientific backing	no information on the results quality

The table can be used as decision-aid tool for selecting the adequate score of the identified source of uncertainty. In other words, analysts select the right cell according to the considered criterion and the textual description in the cell. Then, the score is identified accordingly. In addition, each criterion composing the table shall be used independently as suggested in figure 3 before completing the score justification comment. Scoring according to five criteria is instructive but time consuming.

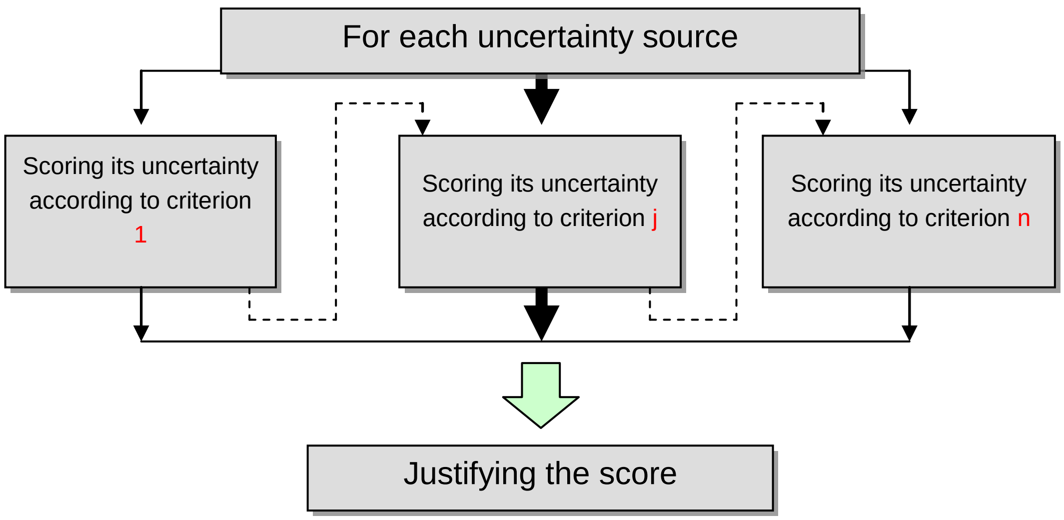

Figure 3. Characterising uncertainty according to several criteria composing a specific dimension

Therefore under time constraints, it would be possible to assess the uncertainty of the knowledge base according to a limited number of criteria. The prioritisation of the criteria to be used (*e.g.*, consistency) should be made explicit during the uncertainty assessment of the complete environmental health impact assessment.

3.1.2.4 Assessing the subjectivity of choices of the source

The third dimension (*i.e.,* subjectivity of choices) aims to characterise how the subjectivity of choices on the source influences the assessment results. While the first two dimensions of uncertainty characterisation can be undertaken by analysts, the third dimension needs to include a selected wider audience with the aim to reach consensus on the discordance. Nevertheless, as it is time consuming to organise a dissemination process, it is recommended to address these same issues via an extensive literature review at AUTH. This dimension of qualitative uncertainty characterisation scales the subjectivity of choices associated with each identified source using three levels: *low* indicates significant consensus on the source and its local influence on the assessment results; *medium* implies that some disagreement on the source is clearly identified and its influence can determine the results; and *high* indicates divergent opinions on the source with radical influence on opposing results.

In the third dimension, *i.e.* subjectivity of choices, five criteria are considered for the score (see table 4). They can be used in exclusive manner (*i.e.*, connected by an 'or' condition) at the discretion of the analysts.

Table 4. Scale of uncertainty related to the subjectivity of choices

score	**Criteria**				
	choice space	**Inter-subjectivity among peers**	**influence of situational limitations on choice**	**sensitivity of choices to the analysts' interest**	**Influence of choices on results**
Low	hardly any alternative assumptions available	many would make same assumption	choice assumption hardly influenced	choice assumption hardly sensitive	only local influence
Medium	limited choice from alternative assumptions	several would make same assumption	choice assumption moderately influenced	choice assumption moderately sensitive	greatly determines the results of link in chain
High	ample choice from alternative assumptions	few would make same assumption	totally different assumption when no limitation	choice assumption sensitive	greatly determines the results of the indicator

The table can be used as decision-aid tool for selecting the suitable score of the identified source of uncertainty. In other words, analysts select the right cell according to the considered criterion and the textual description on the criterion that best fits the situation they evaluate. Then, the score is identified accordingly. In addition, each criterion composing the table should be used independently as suggested in figure 3 before completing the score justification comment. Scoring according to five criteria is instructive but time consuming. Therefore under time constraints, it would be possible to evaluate the subjectivity of the assessment on the basis of limited criteria. Prioritisation of criteria (*e.g.*, influence of choices on results) should be made explicit during the uncertainty assessment of the complete environmental health impact assessment.

3.2 Reporting the qualitative uncertainty characterization

One option for reporting the qualitative uncertainty characterisation is to use the matrix format, accommodating the sources of uncertainty in line and the three dimensions of uncertainty in columns. Particular attention is given to the modules composing the environmental health impact assessment and to regrouping sources into the three identified categories. The scoring of each source according to a specific dimension thus becomes a cell. The scoring value can be coloured to facilitate reading: red means *high*, orange means *medium*, yellow means *low* and white means *not applicable*. For specific cases of intervals [*medium – high*] or [*low – medium*] the two composing colours are used to maintain reading consistency, seen in figure 4.

Sources of uncertainty		***Characteristics of uncertainty***		
		Level of Uncertainty	***Appraisal of Knowledge Base***	***Subjectivity of Choices***
Scenario		High		Medium
Model	*Conceptual*	Medium		
	Mathematical	High	Low	
Parameters		High		

Figure 4. Reporting the qualitative uncertainty characterizations in a matrix format.

Justification comments can be accommodated within the cells when necessary. The immediate objective of this reporting is to facilitate the decision regarding the validity of the overall impact analysis based on the uncertainty assessment. A critical decision consists in allowing further investigation in higher tiers uncertainty assessment for a specific subset of uncertainty sources, *e.g.* key parameters having high scores.

3.3 Quantitative uncertainty assessment tools

In this section essential tools for uncertainty assessment in tier 1 and tier 2 are presented, including sensitivity analysis, error propagation methods, Monte Carlo Simulations and Fuzzy methods. The most suitable statistical tool should be selected in accordance to the Tiered method considered and the corresponding uncertainty level identified. Before calculating various statistical performance measures (or metrics), it is recommended that exploratory data analysis should be firstly performed via plotting the data in different ways. Human eyes can often glean much more inherent information from these plots than pure statistics. These plots can also provide clues as to why a model performed in a certain way. Some of the commonly-used plots are, scatter plots, quantile-quantile plots, residual (box) plots and conditional scatter plots.

3.3.1 TIER – 1 Level

At tier 1 level the interaction between input parameters and outputs is evaluated and when possible the number of influencing parameters are reduced. Furthermore, using the flowchart presented in figure 5, the recommended methods depend on the type of this interaction, including the linear relationships, the non-linear and monotonic and the non-monotonic. The variation in model output caused by specific model input is quantified using the sensitivity methods presented in the following sections. It is noted that some of the methods illustrated in this flowchart are under tier 2 level only due to their complexity.

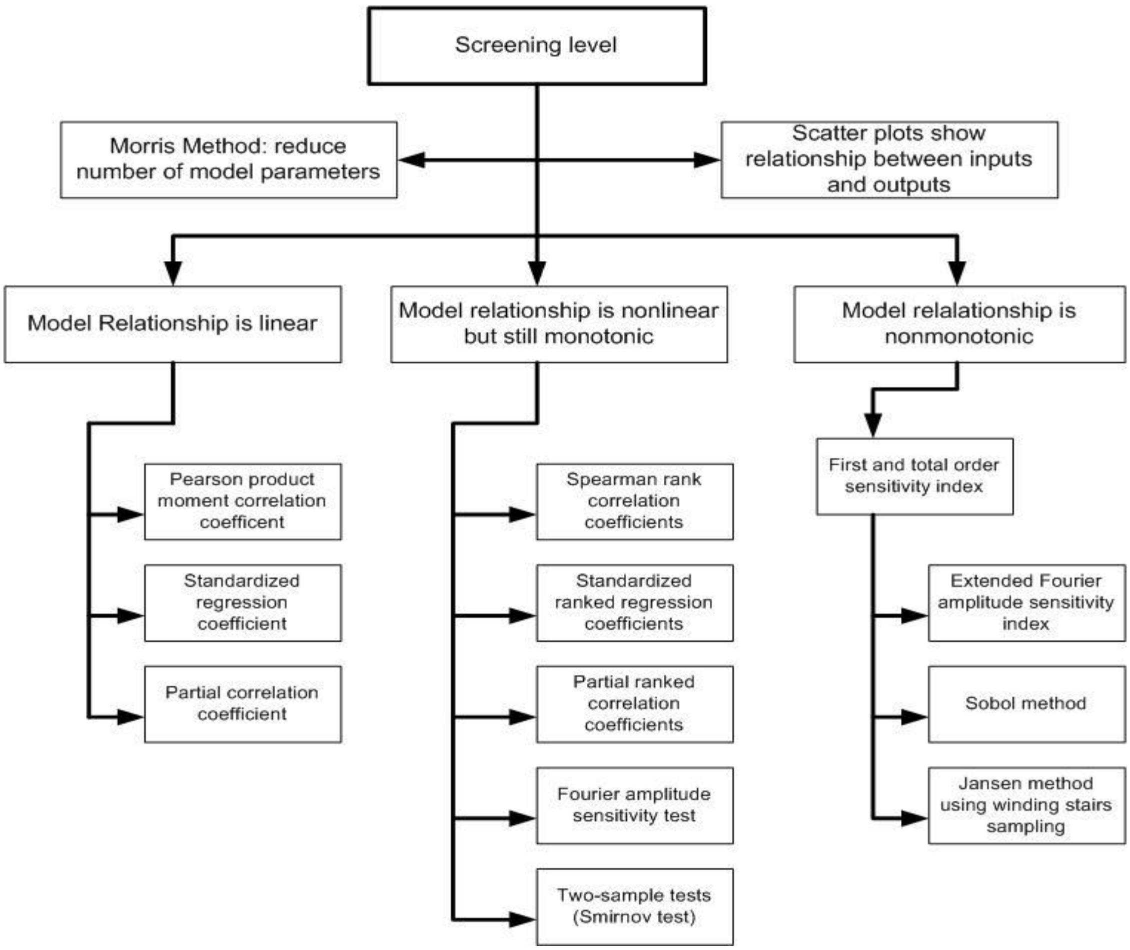

Figure **5**. Selection of methods to perform sensitivity analysis

3.3.1.1 Local Sensitivity Analysis

Sensitivity analysis methods quantify the variation in model output that is caused by specific model inputs. The most commonly used sensitivity analysis methods are often relatively simple techniques that evaluate the local linearized sensitivity of model response at a particular point in the input domain. This type of approach is typically used if the model inputs are treated as point estimates, often representing the "best guess" as to the true but unknown value of each input. The sensitivity analysis of point estimates is often done for the purpose of evaluating how much the model would respond to a unit change in the input. According to a report by WHO (2009), sensitivity analysis should be used to answer the following types of questions (Mokhtari and Frey 2005)

- What is the rank order of importance among the model inputs?
- Are there two or more inputs to which the output has similar sensitivity, or is it possible to clearly distinguish and discriminate among the inputs with respect to their importance?
- Which inputs are most responsible for the best (or worst) outcomes of the output?
- Is the model response appropriate?

For the computationally expensive models, containing large amounts of uncertain input factors, screening methods (Morris 1991) could be used to isolate the set of factors that have the strongest effect on the output variability with very few model evaluations. In this way the number of uncertain input factors under investigation could be reduced, focusing in a relatively small number of uncertain variables. Obviously the most appealing property of the screening methods is the small required number of model evaluations (computational cost). The simplest method is the "one-factor-at-a time" design (OAT). In OAT designs the input factors are varied in turn and the effect each has on the output is measured. Normally, the factors that are not varied are fixed at "nominal" values, which are the values that best estimates the factors. A maximum and minimum value is often used representing the range of likely values for each factor. Usually the nominal value is chosen to be midway of these extremes.

In addition the OAT designs can be used to compute the local impact of the input factors to the model outputs. This method is often called local sensitivity analysis. It is carried out by computing partial derivatives of the output functions with respect to the input variables or by varying the input factors in a small interval around the nominal value. The interval is usually a fixed (e.g. 5%) fraction of the nominal value and is not related to the uncertainty in the value of the factors. In general, the number of model evaluations required for an OAT design is of the order O(k) (often, 2k + 1), k being the number of factors examined.

Forward to the Morris method, the input factor x_i may be important if,

a) $f(x_i + \Delta, x_{\sim i}) - f(x)$ is nonzero, then x_i affects the output

b) $f(x_i + \Delta, x_{\sim i}) - f(x)$ varies as x_i varies, then x_i affects the output non-linearly.

c) $f(x_i + \Delta, x_{\sim i}) - f(x)$ varies as $x_{\sim i}$ varies then x_i affects the output with interactions

where Δ is the variation size.

The input factor is "discretized" and the possible input factor values are restricted inside a regular k-dimensional p-level grid, where *p* is the number of "levels" of the design. The elementary effect of a given value x_i of input factor X_i is defined as a finite difference derivative approximation for any x_i between 0 and $1-\Delta$, as seen by formula 1,

$$ee_i(x) = [f(x_1, x_2, ..., x_{i-1}, x_i + \Delta, x_{i+1}, ..., x_k) - f(x)]/\Delta \qquad (1)$$

Where $ee_i(x)$ is the elementary effect with $x \in \{0, 1/(p-1), 2/(p-1)...,1\}$ and Δ is a pre-determined multiple of $1/(p-1)$.

If all samples of the elementary effect of the i^th input parameter are zero, then x_i does not have any effect on the output y, the sample mean and standard deviation will both be zero. If all elementary effects have the same value, then y is a linear function of x_i. The standard deviation of the elementary effects will then equal zero. For more complex interactions, due to interactions between parameters and nonlinearity, if the mean of the elementary effects is relatively large and the standard deviation is relatively small, the effect of x_i on y is mildly nonlinear. Alternatively, if the mean is relatively small and the standard deviation is relatively large, then the effect is supposed to be strongly nonlinear. Thus we can summarize that,

- a high mean of elementary effects indicates a parameter with an important overall influence on the output
- a high standard deviation in elementary effects indicates that either the parameter is interacting with other parameters or the parameter has nonlinear effects on the output.

3.3.1.2 Fuzzy methods

Fuzzy methods were introduced to represent and manipulate data and information possessing no statistical uncertainties (Zadeh 1965). Fuzzy methods differ from statistical methods in that they do not conform to axioms of probability. Fuzzy methods are based upon fuzzy sets (Jablonowski 1998). An element of a fuzzy set, such as a particular number for an input to a model, has a grade of membership in the set. The grade of membership is different in concept from probability and is often referred to simply as "membership." Membership is a quantitative noncommittal measure of imperfect knowledge. Specifically, the available knowledge on a variable X can be represented by a ''fuzzy number'' which is characterized by its membership function $\mu(x)$. The membership in a fuzzy set is not a matter of acceptance or rejection (i.e., membership function for crisp set is a binary pair (2001) 1 representing membership and 0 representing non-membership) but rather a matter of degree. Thus, $\mu(x)$ is the grade of membership of x in the fuzzy set. The closer the value of $\mu(x)$ is to 1, the more belongs to the fuzzy set. For example, suppose that the height of a person is classified as "tall" or "not tall." In a fuzzy representation, the person's height might be given a partial membership of, say, 0.7 in the tall set, and therefore would have a partial membership of 0.3 in the “not tall” set. For the case of uncertainty, the theory of possibility is possible, based on fuzzy numbers (Klir 1995; Dubois 1988). The membership function, the values of which lie between zero and one, describes the ‘‘possibility’’ that variable X may take a certain value x.

$$\Pi(A) = \underset{x \in A}{Sup}(\mu(x))$$

$$N(A) = 1 - \Pi(A^c) \qquad (2)$$

Where $\Pi(A)$ is the degree of possibility of an event A, $N(A)$ is the necessity function (i.e. the dual function of a possibility function) and A^c denotes the complement of A.

It is noted that fuzzy methods are suitable for approximate reasoning (EPA 2001), especially for analysis of systems where uncertainty arises due to vagueness or “fuzziness” or incomplete information rather than due to randomness alone (Evans 1986). The advantage of these methods lies in that they can characterize non-random uncertainties arising from vagueness or incomplete information and give an approximate estimate of the uncertainties. The limitations of the fuzzy methods are: (1) they cannot provide a precise estimate of uncertainty, but only an approximate estimation; (2) they might not work for situations involving uncertainty arising from random sampling error.

3.3.2 TIER – 2 Level

At tier 2 level, advanced methods are recommended (Error propagation and Monte Carlo Simulation), provided that screening and interactions between input parameters and outputs are evaluated from the tier 1.

3.3.2.1 Error Propagation

Error propagation methods are an efficient method for propagating uncertainty through a model, or set of models. They are based on the Taylor expansion where the moments of a model can be described by a set of increasingly higher order terms. Their objective is to numerically solve the first and second differential in the range where the uncertainty is being propagated. The major limitation of these methods is that they provide information only on centred moments, *e.g.* mean (μ_1), variance (μ_2) and skewness (μ_3). They are generally not applied to provide higher order moments than variance and cannot give information concerning, for example, the tails of distributions. Error propagation actually propagates the moments of a distribution, rather than the distribution itself, and so does not contain the complete information concerning the variability of distribution.

Typical implementation of the error propagation algorithm based on a Taylor series of a function about a point:

$$f(x_1, x_2) - f(x_{0,1}, x_{0,2}) = (x_1 - x_{0,1}) \frac{\partial f}{\partial x_1}\bigg|_{\substack{x_1 = x_{0,1} \\ x_2 = x_{0,2}}} + (x_2 - x_{0,2}) \frac{\partial f}{\partial x_2}\bigg|_{\substack{x_1 = x_{0,1} \\ x_2 = x_{0,2}}} \qquad (3)$$

Assuming small perturbations of x_1 and x_2 about the point $\left(x_{0,1}, x_{0,2}\right)$, changes are denoted as ε_1 and ε_2 and the change of the function under these perturbations is denoted as U_f, where:

$$U_f = \varepsilon_1 \left| \frac{\partial f}{\partial x_1} \middle| x_{0,1} = x_{0,2} \right| + \varepsilon_2 \left| \frac{\partial f}{\partial x_2} \middle| x_{0,1} = x_{0,2} \right| \tag{4}$$

Generalized in any number of variables U_f is presented next:

$$U_f = \sum_n \left\{ \varepsilon_n \left| \frac{\partial f}{\partial x_n} \middle| x_{0,1},...,.x_{0,n} \right| \right\} \tag{5}$$

Furthermore, following the implementation of Taylor expansions to the uncertainty assessment a typical formulation could be as follows:

Let a model $y = f(x_i)$

where x_i are the *i* independent (and uncorrelated) variables, can be written to the second order as:

$$E(y) = f\left(E(x_1), E(x_2), E(x_3),...\right) + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 f}{\partial x_i^{\,2}}\right) Var(x_i) + ... \tag{6}$$

The corresponding variance of the function can also be written as

$$Var(y) = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 Var(x_i) + \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)\left(\frac{\partial^2 f}{\partial x_i^{\,2}}\right)\mu_3(x_i) + ... \tag{7}$$

Given the first and second order partial differentials of *f* and the first, second and third order moments (μ) of the priori-distribution, the variance of *y, Var(y),* can be approximated. In the case where we are interested in a scenario difference we can write the scenario difference function as:

$$\Delta y = f(x_i + \Delta x_i) - f(x_i) \tag{8}$$

where Δx_i represents the change in the variable x_i which will also have some uncertainty $Var(\Delta x_i)$.

Furthermore, the error propagation equation for the second order could be modified as follows,

$$Var(\Delta y) = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 Var(x_i) + \sum_{i=1}^{n}\left(\frac{\partial f}{\partial \Delta x_i}\right)^2 Var(\Delta x_i) \tag{9}$$

3.3.2.2 Monte Carlo Simulations (1D - 2D)

A Monte Carlo simulation (MC) is a widely used method on air quality and exposure assessment (Beekmann and Derognat 2003; Moore and Londergan 2001), Involves a large number of drawings (typically hundreds of thousands) from the distribution of the input parameters in the model that are combined to obtain values for the output parameters (which will be a function of the input parameters). As many values are available for the output parameters a probability distribution can be evaluated. The outputs from each run of the model are saved and a probability distribution for the output values is generated. The output can be in the form of a probability density function or more often as a cumulative probability distribution, which is the integrated PDF. Figure 5 illustrates this process. This allows the probability of the occurrence of any particular value or range of values for the output to be calculated. Based on the distribution of the output, the desired levels of probability could be identified, including the high and low end (e.g., 95th and 5th percentile), the central tendency (e.g., mean and median), or any other level of probability.

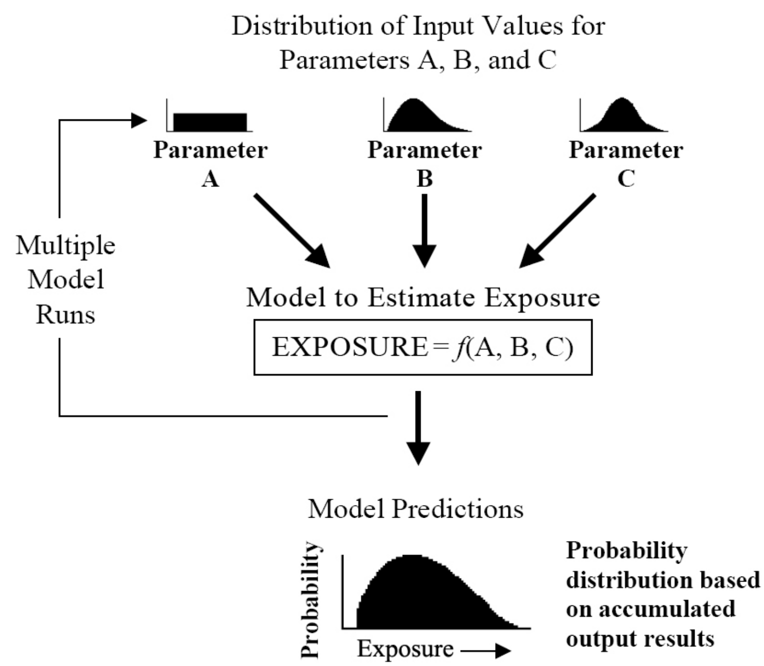

Figure 5. Diagrammatic representation of the application of Monte Carlo analysis to a model

It should be noted that Monte Carlo analysis does not require that probability distribution function are defined for all input parameters. Where there is no basis for assigning a probability distribution function to particular parameters in multiple-parameter models, it is acceptable to keep a fixed value for those parameters while assigning probability density functions to parameters where sufficient information is available. Well known probability density functions are: Normal, Triangular, Uniform and Lognormal. For discrete variables (i.e. a variable that can only assume certain isolated or fixed values), the probability mass function expresses the probability that a randomly selected discrete variable will be a specific value. A well-known probability mass function is the Poisson.

Since in some studies the output is influenced by variables that are both uncertain and variable, a quantitatively characterization of both variability and uncertainty is proposed. A two-dimensional Monte Carlo Analysis (2D MCA) can be used also, where the variables of the model equation along with the parameters of these variables (for example mean and standard deviation for a normal distribution) are described in terms of probability density functions (PDFs). A variable described in this way is called a ''second order random variable.' It is noted that complete datasets or considerable insight to uncertainty associated with these variables is necessary to develop the appropriate PDFs for these random parameters.

3.3.2.2.1 Latin Hypercube Sampling

The classic form of Monte Carlo simulation implies the use of the simple random sample technique; however it is possible to apply other sampling methods to improve the coverage and efficiency of the Monte Carlo methods. Latin Hypercube Sampling (LHS) is an example of such a method (McKay 1979). Instead of randomly generating realisations of the prior distributions this approach divides the distribution up into areas of equal probability, for example using percentile bands of interest and selecting from the median of these bands only, ('Median LHS'), or alternatively selecting randomly from within the specified probability range. The order of the sampling is usually random. LHS can be applied to reduce the number of realisations, particularly when accumulated distributions are to be assessed. For a general overview see (Cullen 1999). LHS has been used in a variety of applications where correlation between parameters exists *e.g.* (Iman 1982); (Pebesma and Heuvelink 1999). This is a particularly important point when assessing the uncertainty between two assessments where there can be a large degree of correlation.

3.3.2.3 Fuzzy Monte Carlo Simulation (FMCA)

An alternative method to the MCA is the 2D Fuzzy Monte Carlo Simulation proposed by Kentel and Aral (2005). Based on the fuzzy set theory, it is possible to allow utilization of incomplete information together with expert judgment. Furthermore, instead of describing the parameters of PDFs as random variables, fuzzy numbers (membership functions) are used. According to the authors, the application of the 2D MCA approach may provide with sufficient information for decision making.

3.3.2.4 Stochastic Response Surface Methods

Since the standard Monte Carlo and Latin Hypercube Sampling, for propagating uncertainty and developing probability densities of model outputs, may in fact require performing a prohibitive number of model simulations, the computationally efficient Stochastic Response Surface Methods (SRSMs) (Isukapalli et al. 1998) could be used. Stochastic Response Surface Methods (SRSMs) extend the classic response surface methodology to systems with stochastic inputs and outputs. This is accomplished by approximating both inputs and outputs of the uncertain system through stochastic series of 'well behaved' standard random variables; the series expansions of the outputs contain unknown coefficients which are calculated by a method that uses the results of a limited number of model simulations. It is shown (Isukapalli et al. 1998) that the results of the SRSMs closely agree with those of traditional Monte Carlo and Latin Hypercube Sampling methods, while significantly reducing the required number of model simulations.

3.3.2.5 Global Sensitivity

According to the (WHO 2008), the local sensitivity-analysis techniques typically suffer from two key shortcomings: (1) they do not take into account the simultaneous variation of multiple model inputs; and (2) they do not take into account any non-linearities in the model that create interactions among the inputs. Alternatively the use of global sensitivity analysis tools could aid in this direction, since they are applicable to any nonlinear function but they are computationally demanding. Specifically, Global sensitivity tools take into account all the variation ranges of the inputs and apportion the output uncertainty to the input factors. These techniques, often based, on the probabilistic framework and Monte-Carlo methods require a lot of simulations. Under global sensitivity, methods could be categorized as those computing the elementary effects (Campolongo et al. 2007), the factor mapping and metamodelling (Helton et al. 2006) and the variance based presented next.

The Fourier Amplitude Sensitivity Test (FAST) (McRae et al. 1982), associates each uncertain parameter with a specific frequency in the Fourier transform space of the system. The system sensitivities are determined by solving the system equations for discrete values of the Fourier transform variable and then computing the Fourier coefficients associated with each parameter frequency. The main advantages of this technique, is the efficiency and robustness in the estimation of the nonlinear global sensitivities model equations subjected to large parameter variations but with a high computational cost. The Random Balance Design (RBD) FAST algorithm (Tarantola et al. 2006), is shown to reduce the computational time mentioned previously, but it can only compute the first order sensitivity indices. The extended RBD-Fast Algorithm (Mara 2009), is used for any number of global sensitivity indices and is particularly efficient at small sample sizes. It’s advantage is the flexibility to include the Sobol sampling based strategy (Sobol 1993) and the ANOVA based sensitivity index (Saltelli 2002).

Below is a brief summary of the Sobol’s indices (Sobol 1993) for a model :

$$y = f(X), \tag{10}$$

where Y is the output and $X = (X_1, \quad \ldots, \quad X_p)$ are *p* independent inputs and *f* is the model function with an unknown analytical formulation.

The objective is to evaluate the contribution of the variance of an input or a group of input parameters to the output variance of *f*. These contributions are described using the following sensitivity indices:

$$S_i = \frac{Var[\mathrm{E}(Y|X_i)]}{Var(Y)},\ S_{ij} = \frac{Var[\mathrm{E}(Y|X_i X_j)]}{Var(Y)} - S_i - S_j \tag{11}$$

Following equation 9, the Sobol's indices, can be used for any complex model functions *f*. The second order index S_{ij} expresses the model sensitivity to the interaction between the variables X_i and X_j (without the first order effects of X_i and X_j). The interpretation of these indices is natural as all indices lie in [0, 1] and their sum is equal to one. It is noted that the larger an index value, the greater the importance of the variable or the group of variables related to the particular index.

3.4 Metrics of Uncertainty

In order to assure the quality of models, the COST Action 732 (Schatzmann M. 2010) guide metrics are suggested, which is widely used for assuring quality in flow and dispersion predictions in urban and industrial areas. According to the guidelines, a number of metrics could be used and implemented through each tier of the uncertainty assessment. These performance measures include, the fractional bias (FB), the geometric mean bias (MG), the normalized mean square (NMSE), the geometric variance (VG), the correlation coefficient (R), the fraction of predictions within a factor of two observations (FAC2) and the Figure of Merit in Space (FMS or threat score).

$$FB = \frac{\left(\overline{C_o} - \overline{C_p}\right)}{0.5 \cdot \left(\overline{C_o} - \overline{C_p}\right)} \tag{12}$$

$$MG = \exp\left(\overline{\ln C_o} - \overline{\ln C_p}\right) \tag{13}$$

$$NMSE = \frac{\overline{\left(C_o - C_p\right)^2}}{\overline{C_o} \cdot \overline{C_p}} \tag{14}$$

$$VG = \exp\left[\overline{\left(\ln C_o - \ln C_p\right)^2}\right] \tag{15}$$

$$R = \frac{\overline{\left(C_o - \overline{C_o}\right)\left(C_p - \overline{C_p}\right)}}{\sigma_{C_p}\sigma_{C_o}} \tag{16}$$

FAC2 = fraction of data that satisfy $0.5 \le \frac{C_p}{C_o} \le 2.0$ (17)

$$FMS = \frac{A_p \cap A_o}{A_p \cup A_o} \tag{18}$$

where C_p denotes model predictions, C_o denotes observations, overbar $\left(\overline{C}\right)$ denotes the average over the dataset, σ_c denotes the standard deviation over the dataset, $\overline{C_o} - \overline{C_p}$ is the mean bias (U.S. EPA),

A_p is the predicted contour area based on a certain threshold and A_o is the observed contour area based on the same threshold.

It is noted that a perfect model would have MG, VG, R, and FAC2 = 1.0; and FB and NMSE =0.0. Furthermore typical magnitudes of the above performance measures and estimates of model acceptance criteria as summarized by (Chang and Hanna 2004) are,

- the fraction of predictions within a factor of two of observations is about 50% or greater (i.e., FAC2 > 0.5).
- the mean bias is within ± 30% of the mean (i.e., roughly |FB| < 0.3 or 0.7 < MG < 1.3).
- the random scatter is about a factor of two to three of the mean (i.e., roughly NMSE < 1.5 or VG < 4).

However, these are not firm guidelines and it is necessary to consider all performance measures in making a decision concerning model acceptance. Since most of these criteria are based on research grade field experiments, model performance would be expected to deteriorate as the quality of the inputs decreases. The performance measures previously presented should not be used independently, but instead multiple performance measures should be used, since each measure has its advantages and disadvantages while there is not a single measure that is universally applicable to all conditions. In this manner, the relative advantages of each performance measure are partly determined by the distribution of the variable of interest.

In the case of atmospheric pollutant concentrations, linear measures FB and NMSE are strongly influenced by infrequently occurring high observed and predicted concentrations, whereas logarithmic measures MG and VG provide a more balanced treatment of extremely high and low values. Therefore, for a dataset where both predicted and observed concentrations vary by many orders of magnitude, MG and VG would probably be more appropriate. FAC2, on the other hand, is the most robust measure, because it is not overly influenced by high and low outliers.

3.5 Tools / software

There is a number of tools that could be used to implement a quantitative uncertainty assessment, as listed below:

@Risk http://www.palisade.com/risk/default.asp a user friendly software that works within Excel.

This software is a complete risk and decision analysis toolkit, including the @Risk, Precision Tree, TopRank, NeuralTools, StatTools , Evolver and RISK Optimizer. It can be used for risk assessment and perform Monte Carlo Simulation. It is easy to use for aggregated data as it is integrated within the MS-office excel. Nevertheless, it is quite complex for spatially distributed data due to the constraints of the MS-Excel. There is a license fee and a maintenance cost.

AcslXtreme http://www.acslsim.com/ is a software for modelling and simulation of dynamic systems and processes. It is excellent choice to perform exposure and risk assessment as it is fully programmable software and can meet challenging data demands. It is requires time to learn, there is a license fee and a maintenance cost.

CrystalBall http://www.oracle.com/crystalball/index.html is a spreadsheet-based software suite for predictive modeling, forecasting, Monte Carlo simulation and optimization. It can be used for risk assessment and perform Monte Carlo Simulations for aggregated only data as it is integrated within MS-office excel. Nevertheless, it is quite complex to implement on spatially distributed data due to the constraints of the MS-Excel. There is a license fee and a maintenance cost.

Matlab http://www.mathworks.com/ a numerical computing environment and programming language. This is a widely used programming language with many support groups in the area of statistical analysis and others. It's fully customized and it is an excellent tool to meet the challenging data demands on uncertainty assessment. This software is our first recommended choice for the of TRANSPHORM full chain uncertainty assessment. There is a license fee and a maintenance cost.

R http://www.r-project.org an environment to perform statistical analysis. This is an excellent software to implement uncertainty analysis as it is programmable hence spatially distributed data could be integrated in the uncertainty assessment. It is quite similar to Matlab and our second recommended choice for the of TRANSPHORM uncertainty assessment. It is freely distributed

WinBUGS http://www.mrc.bsu.cam.ac.uk/bugs/winbugs/contents.html a software to perform statistical analysis using MCMC (Gibbs sampling). It's of limited use and hard to implement for the needs of full chain uncertainty assessment. It is freely distributed.

MCSim http://toxi.ineris.fr/activities/toxicologie_quantitative/mcsim/mcsim.php a software to perform complicated high dimensional joint posterior distributions using Metropolis-Hastings Monte Carlo Sampling. It's of limited use and hard to implement for the needs of full chain uncertainty assessment.

SimLab http://webfarm.jrc.cec.eu.int/uasa/primer/index.asp is a software dedicated to uncertainty and sensitivity analyses using the extended Fourier amplitude sensitivity test (FAST) method. It's of limited use and hard to implement for the needs of full chain uncertainty assessment.

BOOT, http://www.harmo.org/kit/, is a software package intended to be used for evaluation of atmospheric dispersion models. It is a collection of four field data sets as well as software for model evaluation. The Kit is a practical tool intended to serve as a common frame of reference for model performance evaluation as it addresses the classic problem of dispersion from a single point source. This software is free of charge.

4 CONCLUSIONS-RECOMMENDATIONS

This report presents an overview of the areas where uncertainty ensues and the available tools to identify, quantify and minimize it. Uncertainty is firstly identified in a qualitative form, categorized into four levels (steps), where all sources are presented in a matrix, annotating direction, level and knowledge-base. Once the qualitative formulation is completed, depending on the availability and complexity of data different Tier levels are considered. Sensitivity tools are used next to examine the contribution of each model input to variation and uncertainty in the output. This helps to identify sources of uncertainty that could be manipulated or identify which inputs would benefit most from collection of additional data to reduce uncertainty. Error propagation techniques (Taylor expansions or Monte Carlo Simulations) used in accordance to quantify uncertainty propagation through the chain is adequate for the purposes of the Tiers 1 and 2.

Specifically for the needs of TRANSPHORM, different uncertainty approaches would be employed across the full chain health impact assessment methodology. These differences in the approaches depend on:

- Limitations imposed by the complexity of the models
- Scientific soundness
- Availability of data and relevance of method

Transport and emission sources: For transport and emission sources, a **Tier 2 Monte Carlo analysis** of uncertainty is recommended, since emissions estimation is based on emission factors, traffic fleet composition and activity data. For hot emission factors and fuel consumption, log-normal probability distributions could be used for fourteen different speed classes. In the absence of robust experimental data for cold start, the standard deviation over mean of the hot emission factors could be used case, also assuming log-normal probability functions (Kioutsioukis et al. 2010). With regard to the uncertainty originated from vehicles classes and age, age distribution of vehicles might be modeled by a Weibull function within given age distribution boundaries. The outcome will be a distribution of emissions for given road network segments, so as to provide the input for air quality models.

Air Quality modelling: For air quality modelling, a **Tier 0 qualitative description** is needed, based on the fact that the computationally intensive meteorological and air quality models (requiring long time to produce outcomes) will not allow for multiple iterations needed for the probabilistic analysis of a higher Tier assessment. For these models estimates, a sensitivity analysis could be a solution, accompanied by a qualitative description of uncertainty.

It is noted that in order to assess the total uncertainty and evaluate the performance of an air quality model, the uncertainty related to the different modelling components of the system has to be separately quantified. This is in accordance to the fact that usually the chemical and physical processes involved are not linear and, also, some uncertainties may compensate each other. With respect to the Air Quality (AQ) policy, the 2008/50/EC Framework Directive places more emphasis on, and encourages, the use of models in combination with monitoring in a range of regulatory applications, in comparison to previous

Directives, which have based AQ assessment and reporting almost exclusively on measurement data. However, as the directive does not provide guidelines on how to carry out model evaluation to achieve the quality requirements imposed, the development of relevant guidelines is necessary for modellers and authorities. Several attempts have been made for the establishment of uncertainty assessment guidelines within a number of projects, including AIR4EU (Denby et al. 2011) and FAIRMODE (Moussiopoulos et al. 2008). The Guidance Document that was elaborated within FAIRMODE is the current reference point for model users and regulators to ensure that their air quality model meets the quality criteria required by EU legislation. It is noted that the model uncertainty for air quality assessment could be calculated using the Relative Directive Error (RDE) (Denby et al. 2011) or via the Relative Percentile Error (RPE)(Stern J. 2004). Other factors to consider when applying AQ modelling for air quality assessment are the following:

- The "90% of stations requirement", according to which the AQ Directive states that the uncertainty will be determined from the maximum of 90% of the available monitoring stations, in order to exclude outliers from the uncertainty calculation. However, this does not apply if less than 10 monitoring stations correspond to the same scale as the model, in which case all stations have to be considered.
- In order to use model results with confidence for compliance purposes, it is important that the model has been adequately validated for the particular application and well documented and that it contains the relevant physical and chemical processes suitable for the type of application, the scale and the pollutant for which it is applied.
- Finally, the quality of required input data has to be sufficient, e.g. the relevant emission sources for the application need to be adequately represented and suitable meteorological data must be available.

<u>Exposure:</u> Poor exposure assessment is an important source of uncertainty in HIA (Martuzzi et al. 2003) and can result from errors and biases in either air quality models or in exposure models (Fuentes 2009). Exposure models are mostly probabilistic models accounting for the numerous sources of variability, including human activity data (that is often neglected, considering an immobile population/ static population distribution). The different sources of error and uncertainties in the exposure models result from variability not modelled or incorrectly modelled, inaccurate inputs, errors in coding, simplifications of physical, chemical and biological processes to form the conceptual models, and flaws in the conceptual model. The generation of the exposure estimates involves stochastic processes utilizing numerical Monte-Carlo sampling techniques to characterize the variability within an individual over time and between individuals across a population. Uncertainty in the model output is estimated by incorporating the knowledge or measurement-based uncertainty associated with the inputs through multiple iterations of the model.

Exposure and dose deposited on Human Respiratory Tract (HRT) analysis is based on several inputs (beyond concentrations of ambient air) related to parameters such as:

i) the fraction of time spent while in traffic, indoor or during other outdoor activities
ii) PM infiltration at different microenvironments
iii) inhalation rates based on the different activities and
iv) PM size-specific deposition fractions across the Human Respiratory Tract.

A mass balance equation should be used to calculate indoor PM concentrations for the different indoor locations that includes parameters for air exchange, penetration and deposition in the form of:

$$V \cdot \frac{dC}{dt} = Q \cdot \left(inf \cdot C_{out} - C_{ind} \right) - k_{dep} \cdot C_{ind} \cdot V$$

Where V is the volume of the indoor location, C_{out} the outdoor concentration, C_{ind} is the indoor concentration of the respective indoor location, Q is the indoor-outdoor air exchange rate, inf is the infiltration rate and k_{dep} is the deposition rate. Considering that neither initial concentration, nor the presence of indoor emission sources is taken into account, what is actually calculated is the contribution of outdoor air in indoor air exposure. In cases of limited data availability, presenting results from a small number of model scenarios could provide an adequate uncertainty analysis for the air quality and exposure models

Total daily PM deposited dose will be estimated using the PM exposure distribution, the activity level-specific inhalation rates based on the activities in the assigned diaries and cumulative deposition to various regions of the respiratory tract based on HRT region specific deposition fractions for the different PM size fractions. The later will be also in the form of distributions since these deposition fraction will take into account inter-individual differences related to deposition and clearance of HRT.

All these variabilities and uncertainties should be combined within a **Tier 2 Monte Carlo analysis** so as to produce distributions of exposure outputs. Similarly to above

<u>Health impact assessment of transport related PM:</u> Uncertainty for HIA originates from several sources (Martuzzi et al. 2003; Fuentes 2009), namely:

- Uncertainties related to the results of the epidemiological studies or to their generalisation (transferability, incorporation of inter-individual susceptibility, different importance of confounding factors).
- Uncertainties in estimating the impact for each health outcome (possible synergies with other factors).
- Uncertainties related to the concentration-response functions, estimated by epidemiological models.
- Uncertainties concerning the temporal scale of effects, i.e. the latency times from exposure to adverse event. This is an uncertainty mainly associated with long-term exposure studies, as acute effects follow exposure by a few days
- Uncertainties related to the exposure reference value.

Interpersonal variability and variance in exposure estimates have to be explicitly integrated in the burden of disease estimates affecting the concentration (or exposure)-response curves, as well as to incorporate the uncertainty characterizing concentration (or exposure)-response functions.

As a recommendation, the confidence intervals derived in concentration-response functions (CRFs) (together with the mean value) should be provided as a probability distribution, rather than a single value. In this way the probability distribution of exposure will be propagated to the distribution derived for the respective CRFs, as well as to the probability distribution that describes uncertainties related to the population size, so as to derive the probability distribution of health impacts for the given population. Similarly to above, **Tier 2 Monte Carlo analysis** will be employed.

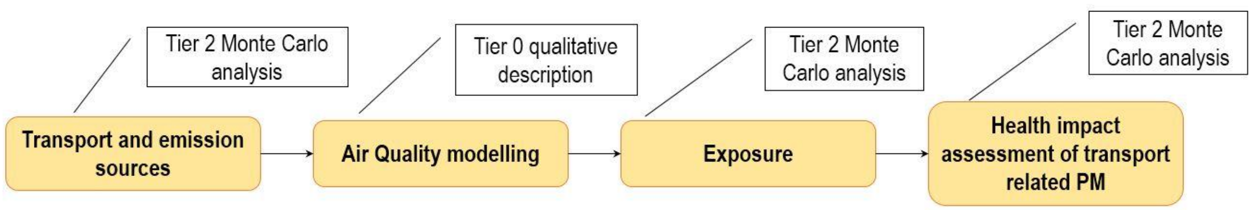

Figure 6. **Transphorm full chain health impact assessment and selected uncertainty management options**

5 APPENDIX

Error propagation algorithm for a simple linear problem

Given a simple analytical example of a function of just two variables A and B multiplied together then we define y to be

$$y = f(x_i) = A * B \tag{A-1}$$

Through error propagation the uncertainty in y becomes

$$Var(y) = B^2 Var(A) + A^2 Var(B) \tag{A-2}$$

The scenario difference function, can be written as

$$\Delta y = f(x_i + \Delta x_i) - f(x_i) = (A + \Delta A).(B + \Delta B) \tag{A-3}$$

Taking ∆A=0 (*i.e.* no change in A for the scenario) for this example then the uncertainty in the scenario difference becomes

$$Var(\Delta y) = \Delta B^2 Var(A) + A^2 Var(\Delta B) \tag{A-4}$$

Clearly the uncertainty in equation A-4 can be significantly less than that in equation 18, at least when ∆B<<B. In this simple linear case the scenario uncertainty becomes independent of the original uncertainty in B.

In this section we demonstrate with a worked example how uncertainties could be identified, characterized and reported using the matrix presented in table 5.

Table 5. Example of qualitatively identifying the dimensions of the uncertainties in the TRANSPHORM project

			Dimensions of Uncertainties			
		Uncertainties	*Source of Uncertainty*	*Direction of uncertainty*	*Level of Uncertainty*	*Appraisal of knowledge base*
Emission	Road	Activity data	parameter	Underestimation	Low	Low
		Emission factors	model	Underestimation	Low	Low
	Sea	Activity data	parameter	Underestimation	Low	Low
		Emission factors	model	Underestimation	Low	Low
	Air	Activity data	parameter	Underestimation	Low	Low
		Emission factors	model	Underestimation	Low	Low
Concentration models		Meteorological Conditions	parameter	Underestimation	Low	Low
		Model resolution	parameter	Underestimation	Low	Low
		Geographical location	parameter	Underestimation	Low	Low
Exposure/ Health impact	Inhalation	Activity Data	parameter	Underestimation	Low	Low
		Particle size	parameter	Underestimation	Low	Low
		Intake rate	model	Underestimation	Low	Low
		Uncertainty associated with the CRFs	parameter	Underestimation	Low	Low
		Population data resolution	parameter	Overestimation	Low	Low
		Background health status	parameter	Underestimation	Low	Low

6 REFERENCES

WHO. 2008. Guidance document on Characterizing and Communicating Uncertainty in Exposure Assessment. (IPCS Harmonization Project Document). ISBN 978 92 4 156376 5.

Mokhtari A, Frey HC. 2005. Recommended practice regarding selection of sensitivity analysis methods applied to microbial food safety process risk models. Human and Ecological Risk Assessment 11(3): 591-605.

Morris DM. 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33(2): 161-174.

Zadeh LA. 1965. Fuzzy sets. Information and Control 8(3): 338-353.

Jablonowski. 1998. Fuzzy Risk Analysis in Civil Engineering.

[Anonymous]. 2001. Exposure factors sourcebook for European populations (with focus on UK data). Exposure Factors Sourcebook for European Populations (with Focus on UK Data).

Klir GJ, Yuan, B. 1995. Fuzzy Sets and Fuzzy Logic: Theory and Applications. New Jersey: Prentice Hall PTR.

Dubois D, Prade, H. 1988. Possibility Theory. New York.

EPA US. 2001. Computational Methods for Sensitivity and Uncertainty Analysis for Environmental and Biological Models.

Evans GWK, W. and Wilhelm, M.R. 1986. An Introduction to Fuzzy Set Methodologies for Industrial and Systems Engineering. New York.

Beekmann M, Derognat C. 2003. Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign. Journal of Geophysical Research D: Atmospheres 108(17): ESQ 2-1 - ESQ 2-17.

Moore GE, Londergan RJ. 2001. Sampled Monte Carlo uncertainty analysis for photochemical grid models. Atmospheric Environment 35(28): 4829-4841.

McKay MD, Conover, W.J., and Beckman, R.J. . 1979. A comparison of three methods for selecting values of input variable in the analysis of output from a computer code. Technometrics 21(2): 239-245.

Cullen ACaF, H.C. . 1999. The Use of Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. New York: Plenum.

Iman RL, and Conover, W.J. . 1982. A Distribution Free Approach to Inducing Rank Correlation Among Input Variables. Communications in Statistics B11(3): 311-334.

Pebesma EJ, Heuvelink GBM. 1999. Latin hypercube sampling of Gaussian random fields. Technometrics 41(4): 303-312.

Kentel E, Aral MM. 2005. 2D Monte Carlo versus 2D Fuzzy Monte Carlo health risk assessment. Stochastic Environmental Research and Risk Assessment 19(1): 86-96.

Isukapalli SS, Roy A, Georgopoulos PG. 1998. Stochastic response surface methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems. Risk Analysis 18(3): 351-363.

Campolongo F, Cariboni J, Saltelli A. 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling and Software 22(10): 1509-1518.

Helton JC, Johnson JD, Sallaberry CJ, Storlie CB. 2006. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering and System Safety 91(10-11): 1175-1209.

McRae GJ, Tilden JW, Seinfeld JH. 1982. Global sensitivity analysis-a computational implementation of the Fourier Amplitude Sensitivity Test (FAST). Computers and Chemical Engineering 6(1): 15-25.

Tarantola S, Gatelli D, Mara TA. 2006. Random balance designs for the estimation of first order global sensitivity indices. Reliability Engineering & System Safety 91(6): 717-727.

Mara TA. 2009. Extension of the RBD-FAST method to the computation of global sensitivity indices. Reliability Engineering and System Safety 94(8): 1274-1281.

Sobol IM. 1993. Sensitivity estimates for nonlinear mathematical models. Math Modell Comput Exp 1(4): 407–414.

Saltelli A. 2002. Making best use of model evaluations to compute sensitivity indices. Computer Physics Communications 145(2): 280-297.

Schatzmann M. OH, Franke J. 2010. COST 732 Model Evaluation Case Studies: Approach and Results.

Chang JC, Hanna SR. 2004. Air quality model performance evaluation. Meteorology and Atmospheric Physics 87(1-3): 167-196.

Kioutsioukis I, Kouridis C, Gkatzoflias D, Dilara P, Ntziachristos L. Uncertainty and sensitivity analysis of national road transport inventories compiled with COPERT 4. 2010, Vol. 2. 7690-7691.

Denby B, Larssen S, Builtjes P, Keuken M, Sokhi R, Moussiopoulos N, et al. 2011. Recommendations for the spatial assessment of air quality resulting from the FP6 EU project Air4EU. International Journal of Environment and Pollution 44(1-4): 128-138.

Moussiopoulos N, Dilara P, Lükewille A, Denby B, Douros J, Fragkou E, et al. 2008. Fairmode: A forum for air quality modelling in Europe. Hrvatski Meteoroloski Casopis 43 PART 1: 187-191.

Stern J. FR. 2004. Formulation of criteria to be used for the determination of the accuracy of model calculations according to the requirements of the EU Directives for air quality – Examples using the chemical transport model REM-CALGRID. Berlin:Institut für Meteorologie.

Martuzzi M, Krzyzanowski M, Bertollini R. 2003. Health impact assessment of air pollution: Providing further evidence for public health action. European Respiratory Journal, Supplement 21(40): 86S-91S.

Fuentes M. 2009. Statistical issues in health impact assessment at the state and local levels. Air Quality, Atmosphere and Health 2(1): 47-55.