18th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 October 2017, Bologna, Italy

EVALUATING THE ROBUSTNESS OF THE SHERPA AIR QUALITY MODEL THROUGH THE APPLICATION OF GLOBAL SENSITIVITY ANALYSIS TECHNIQUES

E. Pisoni[1], D. Albrecht[2], R. Rosati[2], S. Tarantola[3], P. Thunis[1]

[1]European Commission, Joint Research Centre, Directorate for Energy Transport and Climate, Air and Climate Unit
[2]European Commission, Joint Research Centre, Directorate for Competences, Modelling Indicators and Impact Evaluation Unit
[3] European Commission, Joint Research Centre, Directorate Energy Transport and Climate, Energy Security Distribution and Markets Unit

Abstract: Air quality has improved in Europe in the last decades, but exceedances of limits values are still observed, mainly in hot-spots at regional and city level. A dimension shift (from "Europe-wide" to "local" exceedances) calls for novel approaches to regional air quality management, to complement existing EU-wide existing ones. The SHERPA modelling tool has been developed for this reason, to support regional/local decision makers in designing air quality plans, and to support (more in general) the evaluation of the impact on air quality of locally-tailored policies. The model implements surrogate modelling techniques to mimic, in a quick way, the behaviour of fully-fledged physically-based models. Given that SHERPA is used in the policy arena, it of utmost importance to evaluate its robustness, against all kinds of uncertainty. To do so, sensitivity analysis (SA) techniques can be applied. Sensitivity analysis are important ingredients for the quality assurance of models used in evidence-based policy. They reveal links between assumptions and predictions, help in model simplification and may show unexpected relationships between input and output. One of the SHERPA modules, predicting air quality improvement linked to emission reduction scenarios, is evaluated in this paper. As SHERPA is a model with spatially-varying coefficients and inputs, the SA has been performed on a number of selected European cities. The results confirm the robustness of the SHERPA model, and help identify where prioritise further model improvement.
.

Key words: *Integrated assessment modeling; uncertainty and sensitivity analysis; Air quality modeling; Surrogate models; Source-receptor relationships*

INTRODUCTION

Air quality has strongly improved in Europe in the last decades (EEA, 2016), but exceedances of the legislation limit values still persist, mainly for pollutants as ozone (O3), nitrogen dioxide (NO2) and particulate matter (PM10 and PM25). While in the past years these exceedances were wide-spread across Europe, they now tend to concentrate in specific regions or cities (Kiesewetter et al., 2015). This changed situation calls for novel approaches tailored to local air quality management, to complement EU-wide existing policies.

Recently, the SHERPA (Screening for High Emission Reduction Potentials on Air quality) modeling tool has been developed (Clappier et al., 2015, Thunis et al., 2016, Pisoni et al., 2017), as a tool to support regional/local decision makers to design air quality plans. It is distributed with default data covering the whole Europe and allows decision makers to work on his own regional domain. It can be used without the need to perform complex scientific/technical tasks beforehand.

As the tool will be used in the policy arena, it is of utmost importance to evaluate its robustness/uncertainty. In this paper we apply Sensitivity Analysis (SA) on the SHERPA "scenario analysis" module (Thunis et al., 2016). This module allows to estimate how concentrations change for a given emission reduction scenario. This module is used as a basis for all SHERPA modules and is therefore the key element to test.

As SHERPA is a model characterized by spatially-varying coefficients and inputs, this SA has been performed on a selected number of cities (London, Milan, Utrecht, Konstanz, and Helsinki; even due to lack of space only London is here shown), representative of different meteorological and emission inventory conditions. In particular, global sensitivity analysis (GSA) has been carried out using the popular variance based methods described in (Saltelli et al., 2010). The results provide information on the level of robustness of the model output, and allow for evaluating where future modelling effort should focus, to improve the SHERPA model in an effective way.

METHODOLOGY

The main goal of this work is to evaluate the robustness of the SHERPA model, using indicators to evaluate its sensitivity. The two main instruments applied in this study, the "SHERPA model" itself and the "sensitivity analysis indicators" are presented in the next subsections.

The Sherpa model

SHERPA has been developed to provide a fast modeling approach to calculate concentration fields resulting from emission reduction scenarios, mimicking the behavior of a full Chemical Transport Model (CTM). The aim of SHERPA is to mimic the CTM's behavior with a more simple relation derived from a set of full CTM simulations built with various emission reduction scenarios. This set of scenarios should be sufficiently varied (in terms of concentration changes responses to emission changes) to provide SHERPA training phase with enough data variability.

In SHERPA emissions and concentration changes are computed on a cell by cell basis according to the following equation:

$$\Delta C_i = \sum_{j}^{N_{prec}} \sum_{k}^{N_{cell}} a_{i,j,k} \Delta E_{j,k}$$

where the delta concentration (ΔC) in a destination grid cell "*i*" is expressed as a linear combination of the aggregated emissions delta $\left(\Delta E_{j,k}\right)$ for each "k" source cell and precursor "j". The $a_{i,j,k}$ coefficients are estimated using the results of a statistical analysis performed on all available CTMs simulations (base-case and scenarios) over the entire modelling domain performed with the air quality model. This analysis shows that the correlation between ΔC_i (at one receptor cell "*i*") and $\Delta E_{j,k}$ (at source cell "*k*") decreases with "d_{ik}", the distance between these two cells ("*i*" and "*k*"). It has been assumed that the coefficients "*a*" in the previous equation follow a similar trend to this correlation and can therefore be approximated by the following distance-function:

$$a_{i,j,k} = \alpha_{i,j}\left(1 + d_{ik}\right)^{-\omega^{i,j}}$$

where "*i*" is a grid cell within the domain in which the concentration delta is estimated, the index "*k*" runs over all grid cells within the domain and "d_{ik}" is the distance between cells '*i*" and "*k*". The two unknowns α and ω need to be defined for each precursor and each grid cell (see Pisoni et al., 2017, for more details).

After α and ω have been computed, SHERPA can be used to evaluate concentration changes resulting from any emission reduction scenario. The current SHERPA implementation (Pisoni et al., 2017) will be used to perform the uncertainty and sensitivity analysis.

Sensitivity analysis approach

In this paper Global sensitivity analysis (GSA) is used, to overcome the drawback of the Once-at-a-time (OAT) approaches, i.e. making use of methods (such as variance based techniques, regression/correlation, graphical tools…) based on the simultaneous exploration of all uncertain inputs and thus being able to capture nonlinearities and interactions among model input. GSA allows a fully exploration of the input space, in order to exhaustively assess the output uncertainty.

Variance–based techniques are based on the decomposition of the total variance of the model output into terms of increasing dimensionality (Sobol', 1993). These methods are model independent, are capable to highlight interactions among the model inputs (including non–linear, non–additive models), and when applicable, they allow an analysis (and uncertainty decomposition) based on groups of inputs. As for all GSA techniques, the drawback of variance–based measures is their computational cost and the potential necessity of calculating higher-order terms. Also, each investigated input must be characterized by a probability density function (p.d.f.).The method of Sobol' (Sobol', 1993) allows computing the terms of the variance decomposition, in a quite intuitive way, by estimation of a multidimensional integral through Monte Carlo (MC) techniques.

$$V(Y)=\sum_{i=1}^{k}V_i+\sum_i\sum_j V_{ij}+\sum_i\sum_j\sum_k V_{ijk}\cdots+V_{1,2,\ldots,k}$$

This is the so-called ANOVA decomposition over the space Ω^{κ} where the total variance of the output is the sum of the V_i (1st order effect), the V_{ij} that measure the joint effect of the pair (X_i, X_j) on Y (2nd order effect), and the $V_{12\ldots k}$ for the higher order interactions.

Through the previous equation we can then compute all the relevant terms for Sensitivity Analysis. The conditional variance $V[E(Y|X_i)]$ is known as the first-order effect of model input X_i on the model output Y. This conditional variance, divided by the total variance V of the model output (normalization), defines the sensitivity index of X_i, also called first order index or main effect:

$$S_i=\frac{V[E(Y|X_i)]}{V(Y)}$$

S_i is by definition a number between 0 and 1. A high value of S_i denotes an important input in the sense that the uncertainty of the input X_i has an important effect on the uncertainty in the model output Y.

In addition to this, Homma and Saltelli (1996) proposed the total effect sensitivity index of a model input as the sum of all the terms of any order involving that input. Being the sum of all possible sensitivity terms equal to 1, the difference between 1 and the normalized value of $V[E(Y|X_{\sim i})]$ - which expresses all terms of any order that do not include input X_i ($\sim i$ indicates all terms but i) - represents the total effect of input X_i :

$$T_i=1-\frac{V[E(Y|X_{\sim i})]}{V(Y)}$$

For both first-order and total effect indicators we use in this paper the indicators as proposed in Saltelli et al., (2010).

THE CASE STUDY SET-UP

In this work, SHERPA implementation is based on data produced by the CHIMERE CTM model (Menut et al., 2013). In particular, the CTM has been used over the whole European territory with a spatial resolution of roughly 7 x 7 km2. The anthropogenic emissions proxies underlying the model simulations are based on the MACC-TNO emission inventory with residential sector emissions modified to account for the enhanced wood consumption at extremely low temperatures (Terrenoire et al., 2015). The meteorological input data is based on IFS (Integrated Forecasting System from ECMWF) for the year 2009. A set of CTM simulations in which emissions are reduced over the entire modelling domain are used to derive the α and ω coefficients required in the simplified SHERPA equation for each grid cell and precursor. An additional set of simulations, with reductions over specific areas, provide data for the method's validation. More details on the whole procedure can be found in Thunis et al., 2016 and Pisoni et al., 2017. In this paper, we refer to the model linking emission reduction scenarios (of nitrogen oxides-NOX, ammonia-NH3, primary PM-PPM, sulphur dioxide-SO2) to yearly average concentrations of PM25.

As previously said, a sensitivity analysis is performed on the SHERPA model. Because SHERPA is based on spatially dependent coefficients and due to the fact that we cannot analyze every grid cell, we restricted our uncertainty and sensitivity analysis to 5 specific cities: Milan, Utrecht, London, Helsinki, and Konstanz (even if here, due to lack of space, we show results only for London).

SENSITIVITY ANALYSIS RESULTS

To take into account all the quantifiable uncertainties, the analyst in collaboration with the modeler must specify a nominal or central value, and a range derived from a probability distribution, to each input and possibly each coefficient used by the model. In this particular case, it is necessary to identify all the terms affecting the modelled air quality at the city level (as said, in this specific case, PM25 yearly averages), i.e.:

- The model coefficients uncertainty: there are 4 values for α and 4 values of ω, defining the model linking emissions to concentrations (for this we use normal distributions, as estimated during the model training phase);
- The emissions (input) uncertainty, of NOX, NH3, PPM, SO2 (we will use uniform distribution, derived from literature);
- The selected policy, or level of ambition to improve air quality (in terms of emission reductions): in this paper we refer to four policies considered in the Air Quality Package Review (Amann et al., 2014).

The first analysis is made considering all together the 13 available perturbations, namely the four α, the four ω, the four emissions (NOX, NH3, PPM and SO2) and the policy option. The input "policy option" embeds the possible alternative policy options that are proposed in the analysis, each one with the same probability as the others. In all the cities, the policy options (as explained above) are related to different levels of air quality improvement. The results for London (the only case presented here,
) show that the input "policy option" is the most important factor. This means that the choice of the policy strongly influences the concentration reductions, whereas all other model-related inputs (α and ω) do not contribute to such reductions. In this case, the first action is on the policy makers who should discuss upon what is the best policy to put in place.

Let us assume that a given policy has been selected (i.e. the one at 50% air quality improvement). Now, a new sensitivity analysis has been carried out in order to see which input, among the remaining 12, are those that are mostly contributing to the uncertainty in the concentrations reduction. In the example of London (**Errore. L'origine riferimento non è stata trovata.**), we find that emission NOX and emission NH3 are the most important inputs. This means that, if we want to improve the accuracy of the concentrations reductions, and therefore the evidence for the policy, we should make efforts towards reducing the uncertainty on the emissions of NOX and NH3. One can note that, in this case, the inputs related to α and ω have negligible influence on the concentrations reductions. This means that the modeler and the analyst do not need to spend time and efforts in further reducing the uncertainty of such inputs.

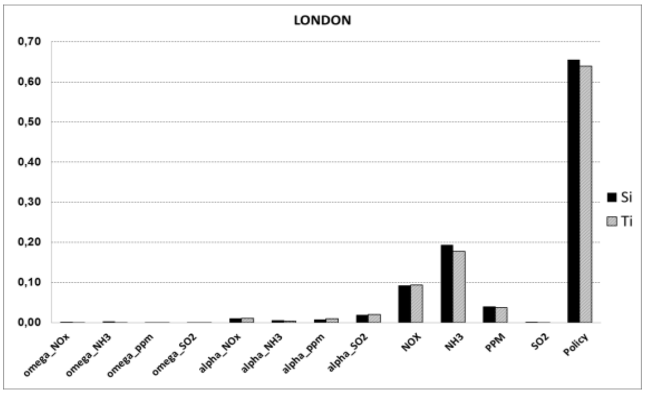

Figure 1. Results considering model coefficients, input data and policy options

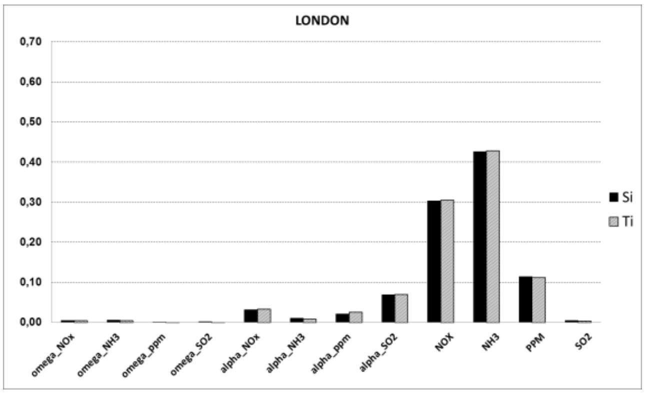

Figure 2: Results considering model coefficients and input data.

CONCLUSIONS

In Europe, we are moving to a situation in which exceedances of air quality legislation thresholds are mainly measured in specific regions or cities. This is way the focus of this paper has been on the application of sensitivity analysis to a modelling tool specifically designed for these geographical (regions/cities) scale. SA results show that the most influential inputs are by far the emissions, in particular of PPM, NOX, and NH3. The model coefficients (α and ω) are a less influential input, even if the α coefficients are more important than the ω ones. Finally, the policy selection (in this case, the level of ambition to be considered in the design of the air quality plan) remains a key aspect.

REFERENCES

Amann et al., 2011. Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, 26, 1489-1501.

Amann et al., 2014. The Final Policy Scenarios of the EU Clean Air Policy Package, TSAP Report #11, IIASA Technical Report.

Clappier A., Pisoni E., Thunis P., 2015. A new approach to design source–receptor relationships for air quality modelling, Environmental Modelling & Software, 74, 66-74.

European Environmental Agency, 2016. Air quality in Europe — 2016 report.

Homma, T., & Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. Reliability Engineering & System Safety, 52(1), 1-17.

Kiesewetter G. et al., 2015. Modelling PM2.5 impact indicators in Europe: Health effects and legal compliance, Environmental Modelling and Software, 74, 201-211.

Menut et al., 2013. CHIMERE 2013: a model for regional atmospheric composition modelling, Geosci. Model Dev., 6, 981–1028.

Pisoni E., Clappier A., Degraeuwe B., Thunis P., 2017. Adding spatial flexibility to source-receptor relationships for air quality modeling, Environmental Modelling & Software, 90, 68-77.

Saltelli, A., & Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software, 25(12), 1508-1517.

Sobol, I. M., 1993. Sensitivity estimates for nonlinear mathematical models. Mathematical Modelling and Computational Experiments, 1(4), 407-414.

Terrenoire et al., 2015. High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE, Geosci. Model Dev., 8, 21–42.

Thunis et al. 2016. On the design and assessment of regional air quality plans: The SHERPA approach, Journal of Environmental Management, 183, 952-958.