Proceedings of the Third International Symposium on Air Quality Management at Urban, Regional and Global Scales. pp. 210-219, 26-30 September 2005, Istanbul – Turkey

MODELING UNCERTAINTY ESTIMATION PROCEDURES FOR AIR QUALITY ASSESSMENT

C. Borrego*, A. Monteiro*, J. Ferreira*, A. I. Miranda*, A. M. Costa* and M. Sousa*

* Departamento de Ambiente e Ordenamento, Universidade de Aveiro, 3810-193 Aveiro, Portugal
borrego@ua.pt

ABSTRACT

This work has a two-fold objective: a review of the current existent methodologies to estimate modelling uncertainty; the preparation of some guidelines for this modelling uncertainty estimation, which can be used by local and regional authorities responsible for air quality management. Examples of modelling uncertainty estimation, using statistical analysis and the European Directives quality indicators, are presented and discussed.

Key Words: Air Quality, Modelling, Uncertainty, Quality Indicators

1. INTRODUCTION

Air quality models are powerful tools to predict the fate of pollutant gases or aerosols upon their release into the atmosphere. Dispersion is primarily controlled by turbulence, which is random by nature, thus cannot be precisely described or predicted by means of basic statistical properties. As a result, there is spatial and temporal variability that naturally occurs in the observed concentration field. On the other hand, uncertainty in the model results can also be due to factors such as errors in the input data and model formulation. Because of the effects of uncertainty and its inherent randomness, it is not possible for an air quality model to ever be "perfect", and there is always a base amount of scatter that cannot be removed (Chang and Hanna, 2004). Nevertheless, air quality models need to be properly evaluated before their predictions can be used with confidence, since model results often influence decisions that have large public to models health and economic consequences. Therefore information about uncertainties associated application is as important as their resulted data, and should be correctly estimated and interpreted. The uncertainty concept is one of the crucial points of Quality Assurance/Quality Control (QA/QC) procedures that should provide quantitative information about the modelling precision, identifying the uncertainty sources and their potential reduction. The present European legislation defines the requirements of QA/QC procedures for air quality modelling, including the definition of Quality Objectives as an acceptability measure, to guarantee they indicate a good model performance and reliable modelling results for decision makers. However, a practical application of these requirements and interpretation of the uncertainty analysis results based on the recommended methodology is difficult, and in some cases incomprehensible for non-expert users. The development of a consistent procedure for the uncertainty evaluation is still a challenge for the scientific community.

2. MODEL UNCERTAINITY ESTIMATION METHODOLOGIES

Uncertainty analysis is defined by Morgan and Henrion (1990) as the computation of the total uncertainty induced in the output by quantified uncertainty in the inputs and model, and the attributes of the relative importance of the input uncertainties in terms of their contributions. Thus, total model uncertainty can be defined by the sum of the model uncertainty, variability and uncertainty on input data. Uncertainties associated with model formulation may be due to erroneous or incomplete representation of the dynamic and chemistry of the atmosphere, incommensurability, numerical solution techniques, and choice of modelling domain and grid structure. Uncertainties in input data are described in terms of emissions, observational data, meteorology, chemistry and model resolution. Variability refers to stochastic atmospheric and anthropogenic processes. It contributes to uncertainties discussed previously, like those associated with emissions estimation and representations of chemistry and meteorology.

The total model uncertainty can be determined by comparison between observations and model predictions through the application of data Quality Indicators, which reflect the ability of a model to simulate real world phenomena. Besides being difficult to define quantitative Quality Indicators for model evaluation, applications of such indicators help to understand model limitations and provide a support for model intercomparison.

There can be three components to the evaluation of air quality models: scientific, statistical and operational (Chang and Hanna, 2004). In a scientific evaluation, the model algorithms, physics, assumptions and codes are examined in detail for their accuracy, efficiency and sensitivity. This exercise usually requires in-depth knowledge of the model. For statistical evaluation, model predictions are examined to see how well they match observations. The operational evaluation component mainly considers issues related to the user-friendliness of the model (user's guide, user interface, etc). This work will focus mainly on scientific and statistical model evaluation and on the estimation of the total uncertainty.

2.1 Sensitivity analysis

Uncertainties in modelling systems are often studied using sensitivity analysis procedures. It is assumed that a model consists of a set of equations with *m* dependent or "output" variables and *n* independent variables plus input parameters. The sensitivity coefficient can be defined as the ratio of the fractional change in an output variable to the corresponding fractional change in an input variable. The combined effects of variations in multiple input parameters can be estimated by assuming that there are no correlations among variables and there are no nonlinear effects, giving the result that the total fractional uncertainty in a given dependent or "output" variable is the square root of the sum of the squares of each individual sensitivity coefficient.

Nevertheless, this sensitivity coefficient approach has problems for large scale photochemical grid models with input parameters exhibiting relatively large uncertainties for systems that are nonlinear and input variables strongly correlated. Advanced mathematical procedures and software systems have been developed allowing sensitivity coefficients to be evaluated for more complex modelling systems (e.g. Carmichael et al., 1997; Saltelli et al., 2000).

2.2 Statistical analysis

Uncertainties can be characterized and air quality model evaluation can be determined by statistical analysis, where model predictions are examined to see how well they match observations. Many scientists have carried out discussion on the evaluation of air quality models and on the development of general evaluation methods; however, standard evaluation procedures and also performance standards still do not exist. Traditionally, model predictions are directly compared to observations, but this direct comparison method may cause misleading results because uncertainties in observations and model predictions arise from different sources (Chang and Hanna, 2004).

As already mentioned, the uncertainty in observations may be due to random turbulence in the atmosphere and measurements errors, whereas the uncertainty in model predictions may be due to input data errors and model physics. Hanna et al. (1993) recommended a set of quantitative statistical performance measures for evaluating models, which have been widely used in many studies (e.g. Nappo and Essa, 2001; Ichikawa and Sada, 2002) and have been adopted as a common European model evaluation framework (Olesen, 2001). Table 1 presents the main statistical parameters used as quality indicators in these studies.

Table 1. Quality indicators for air quality model performance evaluation

Quality indicators	Formula	Observations	Ideal value
Correlation coefficient	$r = \left[\frac{\sum_{i=1}^{N} (Co_i - \overline{C_o})(Cp_i - \overline{C_p})}{\sigma_0 \sigma_p} \right]$	Co and Cp are the concentration observed and predicted $\overline{Co}$ and $\overline{Cp}$ are the averaged concentration observed and predicted	1.0
Fractional bias	$Fb = \frac{\overline{Co} - \overline{Cp}}{0.5(\overline{Co} + \overline{Cp})}$		0.0
Root mean squared error	$RMS = \sqrt{\sum_{i=1}^{N} (Co_i - Cp_i)^2}$	σ_O and σ_P are the standard deviations of observations and predictions Co_i and Cp_i are the observed and predicted concentration in monitoring station ''i''; n the total number of monitoring stations.	0.0
Normalized standard deviation	$NSD = \frac{\sigma_{Cp}}{\sigma_{Co}}$		1.0
Normalized mean square error	$NMSE = \frac{\overline{(Co - Cp)^2}}{\overline{Co}\,\overline{Cp}}$		0.0
Average normalized absolute bias	$ANB = \overline{\left(\frac{\lvert Co - Cp \rvert}{Co} \right)}$		
Geometric mean bias	$MG = \exp(\overline{\ln Co} - \overline{\ln Cp})$		1.0
Geometric variance	$VG = \exp\left[(\overline{\ln Co} - \overline{\ln Cp})^2 \right]$		1.0
Fraction of predictions within a factor of 2 of observations	$0.5 \leq \frac{Cp}{Co} \leq 2.0$		1.0

Since the distribution is close to log-normal for most atmospheric pollutant concentrations, the linear measures FB and NMSE may be overlay influenced by infrequently occurring high observed and/or predicted concentrations, whereas the logarithmic measures MG and VG may provide a more balanced treatment of extreme high and low values. Nevertheless, MG and VG may be overly influenced by extremely low values, near the instrument thresholds and are undefined for zero values. FAC2 is the most robust measure, because it is not overly influenced by outliers. FB and MG are measures of mean relative bias and indicate only systematic errors, whereas NMSE and VG are measures of mean relative scatter and reflect both systematic and unsystematic (random) errors. The correlation coefficient (r) reflects the linear relationship between two variables and is thus insensitive to either an additive or a multiplicative factor. Some authors recommend this parameter when large-scale models with gridded fields are involved (McNally and Tesche, 1993). Elbir (2003) proposed a statistical analysis that included the index of agreement (d), which determines the degree to which magnitudes and signs of the observed value about mean observed value are related to the predicted deviation about mean predicted value, and allows for sensitivity toward difference in observed and predicted values as well as proportionality changes. Studies conducted by Sivacoumar and Thanasekaran (1999) and Karppinen et al. (2000) also established the usefulness of the index of agreement and other above-mentioned statistical parameters for evaluating model performance. The referred index of agreement (d) is defined as follows:

$$d = 1 - \frac{\sum_{i=1}^{N} (Cp_i - Co_i)^2}{\sum_{i=1}^{N} \left(\left| Cp_i - \overline{Co} \right| + \left| Co_i - \overline{Co} \right| \right)^2} \tag{1}$$

considering the same definitions of variables of Table 1. The index of agreement varies from 0.0 (theoretical minimum) to 1.0 (perfect agreement between observed and predicted values) and gives the degree to which model predictions are error free. To show how this statistical analysis methodology could be applied to a test case, an air quality model performance evaluation was conducted to an application of two air quality models to an ozone episode occurred in Portugal, in 2001. A 48 hours simulation was performed for Continental Portugal, over a gridded domain with 10 km horizontal resolution, using CHIMERE and CAMx models, aiming to estimate hourly ozone concentrations (Ferreira et al., 2004). Models performance was evaluated for 5 air quality monitoring stations, three of them considered as background stations and two located in industrial areas. Results of the referred analysis are summarised in Table 2. It must be stressed that the correlation coefficient is one of the most important parameters, as it reflects the ability of the models to simulate measured data, and, based on the results presented in Table 2, its values reveal a little better behaviour of CAMx model. The values obtained for the bias, reported by FB, ANB and MG, reflect the differences between average observed and simulated results. Considering that these three parameters contribute with the same kind of information, only one of them is, in fact, required for a statistical analysis of modelling results. RMSE and NMSE give information about the errors obtained within the observed-predicted pairs of results, but RMSE does not ignore the range of the variable in cause, ozone concentration, which in some cases could lead to incorrect interpretations of the results of this parameter. Thus, a normalized form of the parameter, NMSE, could be in such cases more adequate. FB and MG obtained values for CHIMERE are closer to the ideal values, 0 and 1 respectively, than the correspondent results for CAMx, meaning that systematic errors are higher in CAMx simulation results.

Table 2. Statistical analysis results for CHIMERE and CAMx simulation.

Parameter	Average for all stations		Average for background stations	
	CHIMERE	CAMx	CHIMERE	CAMx
r	0.52	0.62	0.56	0.70
Fb	-0.16	0.28	-0.19	0.20
RMS	49.10	51.15	43.14	42.66
NSD	1.12	0.77	1.13	0.84
NMSE	0.27	0.45	0.25	0.34
ANB	1.06	0.52	1.26	0.53
MG	0.76	1.44	0.73	1.48
VG	1.16	1.16	1.21	1.19
FAC2	1.19	0.77	1.24	0.84
d	0.71	0.69	0.78	0.79

However, both modelling applications have systematic and random errors as indicated by VG results. The parameters MG and VG are, in some way, useless, since they are sensitive to very low concentrations, which occur at night during the application example. Therefore, these parameters should be carefully used in such an evaluation. Regarding FAC2 and d, similar results were obtained for both models. Based on the ideas pointed out, it can be concluded that every statistical parameter plays a role in the evaluation of model performance and uncertainties estimation, but some of them could be considered more important, useful and required for such an analysis, namely, the correlation coefficient (r), the fractional bias (FB), the root mean square error (RMSE) (without forgetting its accounting on the magnitude of the studied variable), and the normalised mean square error (NMSE). For EPA regulatory applications, the primary objective is to evaluate how well an air quality model simulates the maximum one-hour averaged concentration anywhere on the sampling network. USEPA (1996) presents a compilation of a series of photochemical model validation exercises focused on the model's ability to predict the domain-wide peak ozone concentration and the concentrations at all locations with observed ozone data above 60 ppb. These quality indicators are described in Table 3, including the ideal values, which are merely indicative, once they were defined based on tests performed. Table 4 presents the EPA quality indicators that were also computed for the application described above.

Table 3. EPA's quality indicators for air quality model performance evaluation

Quality indicators	EPA	Ideal values
Normalized accuracy of the maximum 1h concentration unpaired in space and time	$\mathrm{A_u} = 100\left(\frac{\mathrm{Co}_{\max} - Cp_{\max}}{Cp_{\max}}\right)$	± 15-20%
Normalized bias test	$D = \frac{1}{N}\sum_{i=1}^{N}\frac{Co_i - Cp_i}{Co_i}$	± 0.05-0.15
Gross error of all pairs Co > 60 ppb	$E = \frac{1}{N}\sum_{i=1}^{N}\frac{Co_i - Cp_i}{Co_i}$	± 30-35%

Table 4. EPA quality indicators obtained for CHIMERE and CAMx simulation

Parameter	Average for all stations		Average for background stations	
	CHIMERE	CAMx	CHIMERE	CAMx
Au	18.0	46.6	10.1	26.5
D	-0.8	0.1	-1.1	0.1
E	0.0	0.1	0.0	0.1

This group of parameters complements the previous analysis, since it evaluates the model capability to simulate peaks, which is particularly important for the evaluation of atmospheric pollutants episodes, like the example exposed.

Another way to evaluate model performance in a stochastic framework is to assume that the observed concentration is simply a random sample taken from the probability density function (PDF) of the predicted concentration, which can be estimated by such techniques as higher-order turbulence closure schemes and Monte Carlo analysis (Chang and Hanna, 2004). Several authors already applied the Monte Carlo method for different urban/regional scale models (e.g. Moore and Londergan, 2001; Beekmann and Derognat, 2003). This method is also the most commonly used to estimate uncertainties in model input variables, since it has a quite simple principle and it can applied to a complete set of more than 100 input parameters and it allows use of standards nonparametric statistical tests concerning confidence intervals (Ang and Tang, 1984; Borrego et al., 2004). In summary, multiple performance measures and methods should be applied and considered in any model evaluation exercise, as each measure/method has advantages and disadvantages and there is not a single method that is universally applicable to all conditions.

2.3 Model uncertainty according to the EU Directives

The Framework Directive (FWD) and Daughter Directives establish requirements for air quality modelling, including the definition of the Modelling Quality Objectives, as a measure of modelling results acceptability. In this context, the uncertainty for modelling and objective estimation is defined as the maximum deviation of the measured and calculated concentration levels, over the period for calculating the appropriate threshold, without taking into account the timing of the events. The quality objectives defined for each quality indicator are listed in Table 5.

Table 5. Modelling Quality objectives established by EU Directives

Pollutant	Quality Indicator	Quality Objective	Directive
SO_2, NO_2, NO_x	Hourly mean Daily mean Annual mean	50-60% 50% 30%	1999/30/EC
PM, Pb	Annual mean	50%	1999/30/EC
CO Benzene	8-hour mean Annual mean	50% 50%	2000/69/EC
Ozone	8-hour daily mean 1-hour average	50% 50%	2002/03/EC

Model quality measures described in the EU Directives are interpreted as the relative maximum error without timing (RME), which is the largest concentration difference of all percentile (p) differences normalized by the respective measured value.

$$RME = \frac{\max(|Co_p - Cp_p|)}{Co_{p_{\max(Co_p - Cp_p)}}} \qquad (2)$$

The question of timing is relevant for those hourly and daily limits, or target values, which are defined as a number of allowed exceedances of a given threshold concentration. Besides that, the model quality objectives for the allowed uncertainty are given as a relative uncertainty, without clear guidance on how to calculate this relative uncertainty. It could be assumed that the respective measured value shall be used to normalize the absolute difference between the maximum deviation of the measured and calculated concentration levels. Another possibility would be to take the maximum relative deviation, but this approach could shift the emphasis to the very low measured concentration ranges, where usually the largest relative deviations between observations and calculations occur, which could be the main reason for non-compliance of annual mean values accuracy requirements. Besides that, other problems of the interpretation of the model accuracy requirements, according to the EU Directives could occur since there are no differences between a short-term and long-term model application accuracy analysis, being the first one in advantage due to the number of paired-in-time results. An alternative model error measure was already proposed by Stern and Flemming (2004), defining the quality indicator as the concentration difference at the percentile corresponding to the allowed number of exceedances of the limit value normalized by the observation (RPE).

$$RPE = \frac{|Co_p - Cp_p|}{Co_p}, p \qquad (3)$$

This measure is more robust than the error defined in the EU Directive and also evaluates the model performance in the high concentration ranges, but without the sensitivity to outliers. Since the model accuracy is examined in the concentration range of the limit values, there is also a direct link to the EU Directives. In order to test and illustrate these model accuracy measures, a one-year simulation of the chemistry-transport model CHIMERE was used. CHIMERE was applied in the regional scale mode, covering Portugal with a resolution of 10 km for the entire year 2001 (Borrego et al., 2005). The model results were compared with measured data from 23 sites of the national air quality monitoring network according to the EU directives thresholds. In Table 6 is presented an average of the relative maximum error (RME) and the relative error at the percentile that correspond to the allowed number of exceedings of the limit value threshold (RPE) for the background and all the monitoring sites, for each pollutant indicator defined by the EU Directives. Concerning the hourly and daily averages indicators, the analysis of the relative maximum error (RME) defined by the EU directives reveals that it is calculated at the highest measured value.

Table 6. Average of RME and RPE for the background and all the monitoring sites, for each pollutant indicator defined by the EU Directives

Pollutant	EU Directives indicators	RME (%)*	Percentile (P)	RPE (%)*	RPE (%)**
SO_2	Human health protection (hourly mean)	79	99.73 (25^{th} max 1h mean)	34	40
	Human health protection (daily mean)	66	99.18 (4^{th} max 24h mean)	57	69
	Vegetation protection	33	annual mean	33	46
	Vegetation protection	44	winter mean	44	58
NO_2	Human heath protection (hourly mean)	81	99.79 (19^{th} max 1h mean)	39	48
	Human heath protection	47	annual average	47	50
O_3	Human heath protection (8h running daily mean)	69	93.15 (26^{th} max 8h daily mean)	16	35
	Vegetation protection	71	AOT40	49	65

*considering only background monitoring stations; ** considering all monitoring stations

In these cases, the assessment of the model accuracy depends on the model performance in a concentration range having an extremely small probability. This also means that the model accuracy assessment could probably be based on an outlier concentration caused by an error of the monitoring unit or an extreme weather situation. In fact, and in opposite to the RME, the alternative model error measure proposed (RPE) shows a quite total compliance with the legislation accuracy requirement of 50% for all the pollutants indicators. These conclusions are in agreement with other model evaluation studies with similar or even higher complexity (Stern and Flemming, 2004; Hass et al., 2003). The analysis of Table 6 reveals also the problem of the heterogeneity of the observed concentration fields and the importance of selecting the adequate and representative monitoring sites for model resolution, since it is impossible for a grid model to simulate all stations with the required accuracy. In spite of the European Air Quality monitoring network (EUROAIRNET) considers that both spatial and temporal representativeness of monitoring stations should be addressed in uncertainty estimation procedures, in order to guarantee a more accurate comparison with air quality standards, the Daughter Directives say nothing about the monitoring stations representativeness and the selection of criteria for the number and type of stations to be used on model accuracy evaluation. Nevertheless, there is a need for pre-selecting the stations to be used for model evaluation and that should be relied on the sites classification or on the prior knowledge of the air quality regime of the measurement sites (based on daily mean and the daily variation of each pollutant). Besides the monitoring stations representativeness, there is absence of any guidance in the EU Directives about measurement inaccuracy and incomplete data coverage that should all be taken into account in the context of a model evaluation. Regarding to data coverage, the EU Directives require a minimum of 90 % data coverage of the hourly or daily values. In fact, this is another model accuracy check problem since in the past the data coverage of the Portuguese stations was mostly below 90%.

3. CONCLUSIONS

A systematic description of the modelling uncertainty analysis methodologies, based on bibliography review, was performed and discussed. Examples of air quality modelling uncertainty estimation at regional scale were presented, taking into account the review of the current existent scientific and legislated methodologies.
The statistical analysis suggested to evaluate model performance and to estimate uncertainties comprises a set of parameters, giving information about the ability of the model to predict the tendency of observed values, errors on the simulation of average and peak observed concentrations, and type of errors (systematic or unsystematic). From the application exercise, it was concluded that despite all parameters are important, it is possible to define a subset of parameters able to reproduce the general uncertainties estimation, comprising the correlation coefficient, the fractional bias and the root and normalized mean square errors. Parameters that reflect the capability to simulate peaks should be taken into consideration in air pollution episodes simulation. Concerning the quality indicators defined by EU directives, the results show that the legislated uncertainty estimation measures are ambiguous and inadequate in several aspects, mainly in what concerns the error measures for hourly and daily indicators based on the highest observed concentration. A relative error at the percentile correspondent to the allowed number of exceedances of the limit value was suggested and tested, showing that is more robust and also evaluates the model performance as required. Besides that, the EU directives do not give rules on how to deal with monitoring stations representativeness on model evaluation, an important issue to guarantee the correct information about measured or predicted exceedances of thresholds values.

4. ACKNOWLEDGEMENTS

The authors wish to thank to the financial support of the AIREU Research Project (FP6-2002-SSP-1) and of the 3rd EU Framework Program and the Portuguese Ministério da Ciência e do Ensino Superior, for the PhD grant of A. Monteiro (SFRH/BD/10922/2002), M. Sousa (SFRH/BD/19048/2004) and A. M. Costa (SFRH/BD/11097/2002).

REFERENCES

Ang, A.H-S and Tang, W.H., 1984. Probability Concepts in Engineering Planning and Design, Vol. 2: Decision, Risk, and Reliability. Wiley, New York.
Borrego, C.; Tchepel, O.; Costa, A.M.; Amorim, J.H.; Miranda, A.I., 2003. Emission and dispersion modelling of Lisbon air quality at local scale. Atmos Environ 37, 5197-5205.
Borrego, C.; Monteiro, A.; Miranda, A.I. and Vautard, R., 2005. Air quality assessment for Portugal. Proceedings of the 5th International Conference on Urban Air Quality, Eds. Ranjeet Sokhi and Nicolas Moussiopoulos, pp. 57.
Beekmann, M.; Derognat, C., 2003. Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the Atmospheric Pollution Over Paris Area (ESQUIF) campaign, J. Geophys. Res., 108, D17, 8559.

Carmichael, G.R., Sandu, A. and Potra, F.,1997. Sensitivity analysis for atmospheric chemistry models via automatic differentiation. Atmos Environ 31, 475-489.
Chang, J. C. and Hanna, S. R., 2004. Air quality model performance evaluation. Meteorol. Atmos. Phys. 87, 167-196.
Elbir, T., 2003. Comparison of model predictions with the data of an urban air quality monitoring network in Izmir, Turkey; Atmos Environ 37 (2003) 2149–2157.
Ferreira, J., Salmim, L., Monteiro, A., Miranda, A.I., Borrego, C., 2004. Evaluation of ozone episodes in Portugal by photochemical modelling. Proceedings of the 8th Environment National Conference: Lisboa, Portugal, pp. 383-384.
Hanna, S.R., Chang, J.C, Strimaitis, D.G., 1993. Hazardous gas model evaluation with field observations. Atmos Environ 27A: 2265-2285.
Hass, H., van Loon, M. Kessler, C., Matthijsen, J., Sauter, F., Stern, R., Zlatev, R., Langner, J., Fortescu, V., Schaap, M., 2003. Aerosol modelling: results and intercomparison from European Regional-scale Modelling Systems. A contribution to the EUROTRAC-2 subproject GLOREAM. EUROTRAC report 2003.
Ichikawa, Y., Sada, K., 2002. An atmospheric dispersion model for the environment impact assessment of thermal power plants in Japan – a method for evaluating topographical effects. J. Air Waste Manage Assoc 52: 313-323.
Karppinen, A., Kukkonen, J., Elolahde, T., Konttinen, M., Koskentalo, T., 2000. A modelling system for predicting urban air pollution: comparison of model predictions with data of an urban measurement network in Helsinki. Atmos Env 34, 3735–3743.
McNally, D., Tesche, T.W., 1993. MAPS sample products. Alpine Geophysics, 16225 W. 74th Dr., Golden, CO 80403.
Moore, G.E. and Londergan, R.J., 2001. Sampled Monte Carlo uncertainty analysis for photochemical grid models. Atmos Environ 35, 4863-4876.
Morgan, M.G., Henrion, M., 1990. A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. New York: Cambridge Univ. Press.
Nappo, C.J., Essa, KSM, 2001. Modelling dispersion from near-surface tracer releases at Cape Canaveral, FL. Atmos Environ 35: 3999-4010.
Olesen, HR, 2001. Ten years of harmonization activities: past, present and future. Proceedings of the 7th Int. Conf. On Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, Belgirate, Italy.
Saltelli, A., Chan, K., Scott, E.M., 2000. Sensitivity analysis. Wiley, 475 pp.
Sivacoumar, R., Thanasekaran, K., 1999. Line source model for vehicular pollution prediction near roadways and model evaluation through statistical analysis. Environmental Pollution 104, 389–395.
Stern, J.; Flemming, R., 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, Freie Universität Berlin, Institut für Meteorologie.
US Environmental Protection Agency, 1996. Compilation of Photochemical Models' Performance Statistics for 11/94 Ozone SIP Applications. EPA-454/R-96-004.