Uncertainties of monitoring and modelling

Cross-Cutting Uncertainty

Carlos Borrego, Ana Isabel Miranda, Ana Margarida Costa, Alexandra Monteiro, Joana Ferreira, Helena Martins, Ana Cristina Carvalho, Oxana Tchepel

Air4EU

Expert Meeting

Brussels 30 June 2006

GEMAC

universidade de aveiro

MOTIVATION

The uncertainty concept is one of the crucial points of Quality Assurance/Quality Control (QA/QC) procedures that should provide information about the monitoring and modelling precison, identifying the uncertainty sources and their potencial reduction.

In real atmosphere, there are spatial and temporal variability that naturally occurs in the observed concentrations fields → direct consequence of the turbulent nature of the atmosphere.

QA/QC of air quality is a system of procedures to ensure that measurements are of known accuracy and precision, are comparable and representative of ambient condiditons.

Because of the effects of uncertainty and its inherent randomness, it is not possible for an air quality model to ever be "perfect"…

overview

Uncertainties related to monitoring methods

Uncertainties related to modelling methods

Total model uncertainty estimation

Quantitative analysis

Statistical analysis

Uncertainties according to EU Directives

Quantification of model uncertainties from other components

Model uncertainty

Input data uncertainty

Uncertainties related to monitoring methods

Data Quality Objectives of some monitoring programs.

Monitoring programme/ Monitoring objective	Compounds	Accuracy	Precision	Data time coverage
EUROAIRNET Mapping, comparability		≤10%	≤2 ppb	≥90%
EU Regulatory Monitoring [1)] Detect non-compliance with directives	SO_2, NO_2 PM, Pb	15% [2)] 25% [2)]		90% annual -
EMEP Provide basis for control of models		15-25% [3)]		90% annual
WMO-GAW Detect trends over short term (5 years)	Examples: O_3 NO_2 $PM_{2.5}$	15% or 3 ppb 20% or 50 ppt 0.05+5% M	10% or 1 ppb 10% or 25 ppt 10%	80% monthly - 90% monthly

The **EU Directive DQOs** are set with a view to the practical measurement accuracy achievable in the field with typical present-day procedures. **EUROAIRNET DQOs** are stricter than those defined by the EU Directive.

The main basis for the **EMEP DQOs** is to provide measurements to control modelling results, and an accuracy of 15-25% is considered sufficient for that purpose. Its data completeness DQO is 90%.

Accuracy and precision defined for **WMO/GAW DQOs** vary between 10% and 20% for individual gases and, in absolute terms, is much higher in the WMO network than in urban networks.

Uncertainties related to modelling methods

Total model uncertainty estimation

TOTAL UNCERTAINTY = **MODEL** UNCERTAINTY + **VARIABILITY** + **DATA** UNCERTAINTY

with...

- Statistical analysis
- EPA Quality Indicators
- Data Quality Objectives of the Air Quality Framework Directive

Statistical analysis

QUALITY INDICATORS	Formula
Correlation coefficient	$r = \left[\frac{\sum_{i=1}^{N} (Co_i - \overline{C_o})(Cp_i - \overline{C_p})}{\sigma_0 \sigma_p} \right]$
Fractional bias	$Fb = \frac{\overline{Co} - \overline{Cp}}{0.5(\overline{Co} + \overline{Cp})}$
Root mean squared error	$RMS = \sqrt{\sum_{i=1}^{N} (Co_i - Cp_i)^2}$
Normalized standard deviation	$NSD = \frac{\sigma Cp}{\sigma Co}$

QUALITY INDICATORS	Formula
Normalized mean square error	$NMSE = \frac{\overline{(Co - Cp)^2}}{\overline{Co}\,\overline{Cp}}$
Average normalized absolute bias	$ANB = \overline{\left(\frac{\lvert Co - Cp \rvert}{Co} \right)}$
Geometric mean bias	$MG = \exp(\overline{\ln Co} - \overline{\ln Cp})$
Geometric variance	$VG = \exp\left[\overline{(\ln Co - \ln Cp)^2} \right]$
Fraction of predictions within a factor of 2 of observations	$0.5 \leq \frac{Cp}{Co} \leq 2.0$

Statistical analysis

1. two regional models were applied to Continental Portugal with a resolution of 10 km, for a 48 hrs simulation
2. results were compared with measured data from 5 monitoring sites
3. statistical parameters were calculated…

Parameter	Average for all stations		Average for background stations	
	MODEL1	MODEL2	MODEL1	MODEL2
r	0.52	0.62	0.56	0.70
Fb	[illegible]	[illegible]	[illegible]	[illegible]
RMS	[illegible]	[illegible]	[illegible]	[illegible]
NSD	[illegible]	[illegible]	[illegible]	[illegible]
NMSE	[illegible]	[illegible]	[illegible]	[illegible]
ANB	[illegible]	[illegible]	[illegible]	[illegible]
MG	[illegible]	[illegible]	[illegible]	[illegible]

correlation coefficient is one of the most important parameters, as it reflects the ability of the models to simulate measured data

RMSE does not ignore the range of a variable, thus, a normalized

every statistical parameter plays a role in evaluation of model performance and uncertainties estimation, but some could be considered more important, useful and required for such an analysis, namely *r*, *Fb* and *NMSE*

Uncertainties according to EU Directives

EU Directives establish requirements for air quality modeling as a measure of modeling results acceptability:

"**maximum deviation** of the measured and calculated values, over the period for calculating the appropriate threshold, without taking into account the timing of the events"

Pollutant	Quality Indicator	Quality Objective	Directive
SO_2, NO_2, NO_x	Hourly mean	50-60%	1999/30/EC [1)]
	Daily mean	50%	
	Annual mean	30%	
PM, Pb	Annual mean	50%	
CO	8-hour mean	50%	2000/69/EC [2)]
Benzene	Annual mean	50%	
Ozone	8-hour daily mean	50%	2002/03/EC [2)]
	1-hour average	50%	

1) The term "Accuracy" is used in the Directive.
2) The term"Uncertainty" is used in the Directive.

Uncertainties according to EU Directives

EU directives **model quality measure** is interpreted as the relative maximum error without timing - largest concentration difference of all percentile differences normalized by the respective measured value

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

1) 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

2) 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

3) model quality objectives for the allowed uncertainty are given as a relative uncertainty, **without clear guidance** on how to calculate this relative uncertainty

Uncertainties according to EU Directives

an alternative model error measure is…

concentration relative error at the percentile corresponding to the allowed number of exceedances of the limit value normalized by the observation

$$RPE = \frac{|Co_p - Cp_p|}{Co_p}, p$$

(Relative Percentile Error)

p - percentile

1) more **robust** (uses the percentile instead of the maximum value)
2) **evaluates model** performance in the high concentration ranges
3) direct **link to the EU Directives**, since the model uncertainty is examined in the concentration range of the limit values

Uncertainties according to EU Directives

Practical example

EU model quality measure

SUGGESTED model quality measure

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

> 50%

Calculated at the highest measured value
Model accuracy depends on the model performance in a concentration range having an extremely small probability

Quite total compliance with legislation accuracy requirement for all the pollutants indicators

only background monitoring stations

Quantification of model uncertainties from other components

Model uncertainty

Sensitivity analysis and/or model intercomparison exercises are useful to evaluate the performance of different model modules.

4 different chemical mechanisms included in 2 photochemical mesoscale modelling systems (**MODEL 2 and MODEL 3**), were applied to a particular area of Portugal and tested/intercompared, where a field campaign was carried out during a summer period of 2001 year

MODEL 2 includes the Carbon Bond IV and the revised radical termination reactions for the same mechanism (CB IV extensions)

MODEL 3 includes the KOREM and EMEP mechanisms

Quantification of model uncertainties from other components

Model uncertainty

48 hours simulation was performed for ozone using a 200 x 140 km^2 domain and a vertical structure with 23 unequally spaced levels, with first level height at 22 m.

To estimate the **level of uncertainty associated with the chemical mechanism**, it was performed a statistical analysis according to the relative difference obtained using the different chemical schemes.

Relative difference (%) associated with the application for the different chemical mechanisms integrated in the same numerical model (MODEL 2 and MODEL 3)

Mechanisms	Monitoring site 1 Aveiro	Monitoring site 2 Sangalhos	Monitoring site 3 Covelo	Average
MODEL 2 (CBIV and CBIV extensions)	No significant	No significant	No significant	-
MODEL 3 (KOREM and EMEP)	18.51	19.30	21.90	**20**

Significant (and similar) differences at all the monitoring sites.

Relative difference is in the order of 20% and is an indicator of the uncertainty level associated with the chemical scheme used in an air quality model.

This study reflects a simple and easy way to evaluate the influence of a specific model formulation in the results and to estimate its contribution to the total model uncertainty.

Quantification of model uncertainties from other components

Input data uncertainty

Example

Input data uncertainty estimation through
Monte Carlo Experiments (Hanna *et al.*, 2001)

1. Identify the input parameters that will be considered for variations into the Monte Carlo experiment

   to get knowledge on each specific input variable uncertainty

2. To each input parameter it should be associated a distribution function (shapes and key parameters such as median and variance)

Quantification of model uncertainties from other components

Input data uncertainty

Example of the type of information compiled for the considered input variables

Variable	**Uncertainty range (includes 95 % of data) (log-normal distribution)**
Initial ozone concentration	**Factor of 3**
Initial NO_x concentration	**Factor of 5**
Top ozone concentration	**Factor of 1.5**
Side ozone concentration	**Factor of 1.5**
Wind speed	**Factor of 1.5**
Wind direction	± **40 º (normal)**
Ambient temperature	± **3 K (normal)**
Major point NO_x emissions	**Factor of 1.5**
Area biogenic NO_x emissions	**Factor of 2**
Photolysis rates	**Factor of 2**
CB-4 reactions 1-94	**Factor of 1.01 to 3.02** **Median 1.80, Mode 2.5**

Quantification of model uncertainties from other components

Input data uncertainty

Major limitations on Monte Carlo methodology

Dependence on good knowledge of input variable uncertainty

Dependence on satisfactory model

The model must be able to account properly for the major physical and chemical effects.

The methodology **does not account for deficiencies**

in model formulations

The uncertainty analysis looks only at the influence of **uncertainties in model input** parameters and variables on the uncertainties of model output variables.

Quantification of model uncertainties from other components

Input data uncertainty

Advantages of Monte Carlo methodology

The application of the method allows to:

- include a huge amount of model input variables into the uncertainty study
- determine the relative importance of initial parameters on the final modelled concentrations
- establish the uncertainty of final modelled concentration due to all considered input parameters

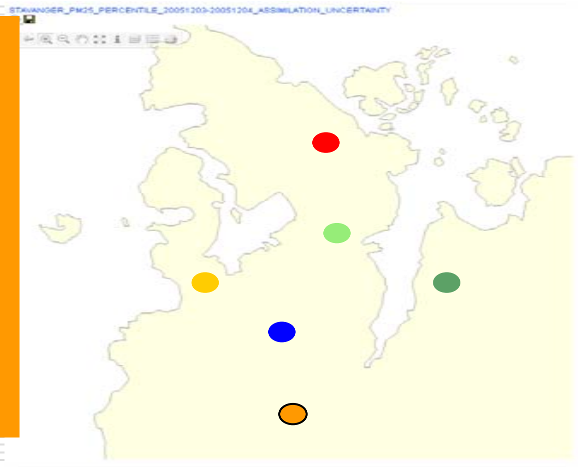
Uncertainty mapping

Estimation of the total model uncertainty parameters by comparison between modelled data and measured data, at the measuring point.

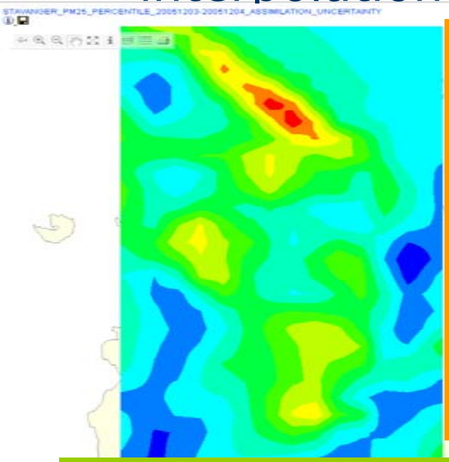
How to show the uncertainty parameters?

post maps

In case of **deficient spatial air quality network coverage** is recommendable to use **point representation** of model errors.



interpolation



Use **spatial interpolation** of the model errors (kriging or other method) to create **maps**, in case of **good spatial air quality network coverage** of the study domain

Interpolation adds uncertainty to the process

Maps of **RPE** for daily or hourly averages and **RME** for annual averages for **directive purposes**

Also maps of normalised **NMSE** (Normalised Mean Square Error) and **Fb** (Fractional Bias)