MQO for percentiles: suggestions for change

Jan Horálek (CHMI, ETC/ACM)

CZECH HYDROMETEOROLOGICAL INSTITUTE

1. Motivation

2. MQO for annual averages vs. for percentiles

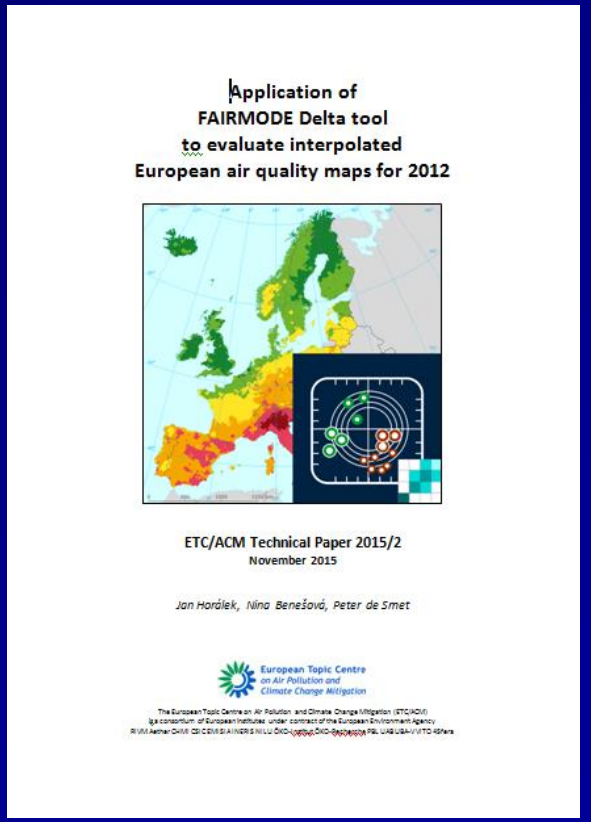
3. Observation uncertainty in MQO

4. Suggestions concerning MQO for percentiles

Motivation

Evaluation of annual data-fused maps by Delta 5.0

- Annual European-wide air quality maps for 2012, created by the Regression-Interpolation-Merging Method under ETC/ACM
- Routinely evaluated by cross-validation
- Under ETC/ACM Technical Paper 2015/2 evaluated by Delta 5.0


Application of
FAIRMODE Delta tool
to evaluate interpolated
European air quality maps for 2012

ETC/ACM Technical Paper 2015/2
November 2015

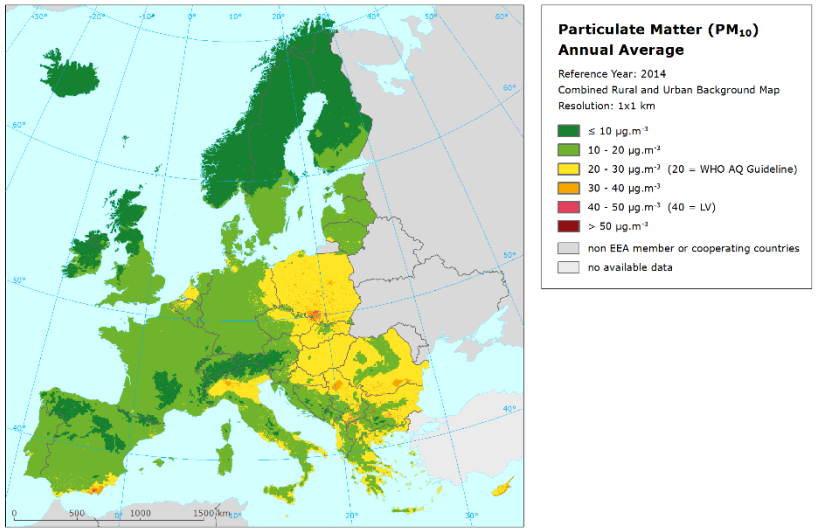
Jan Horálek, Nina Benešová, Peter de Smet

European Topic Centre on Air Pollution and Climate Change Mitigation

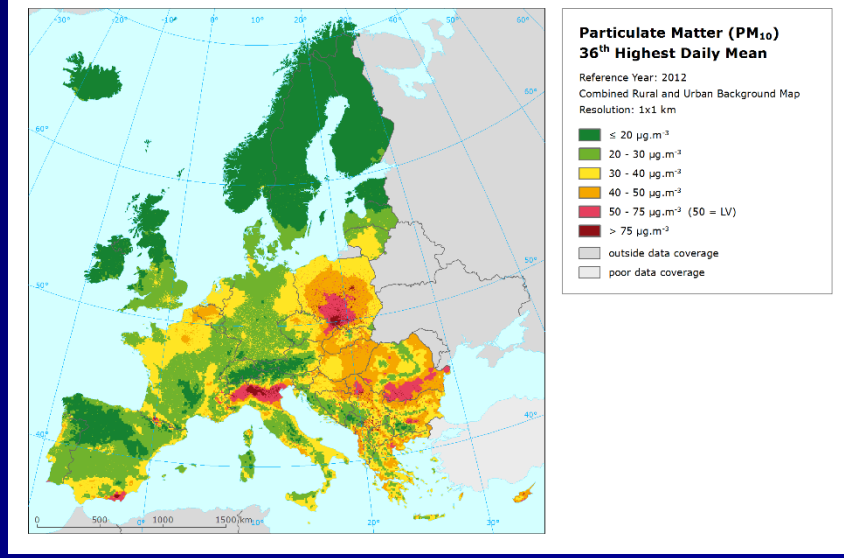
- Among others: PM_{10} annual average and 36th highest daily mean (cc. 90.4 percentile of daily means) maps

PM_{10} annual indicator maps for 2012

Annual average



36^{th} highest daily mean



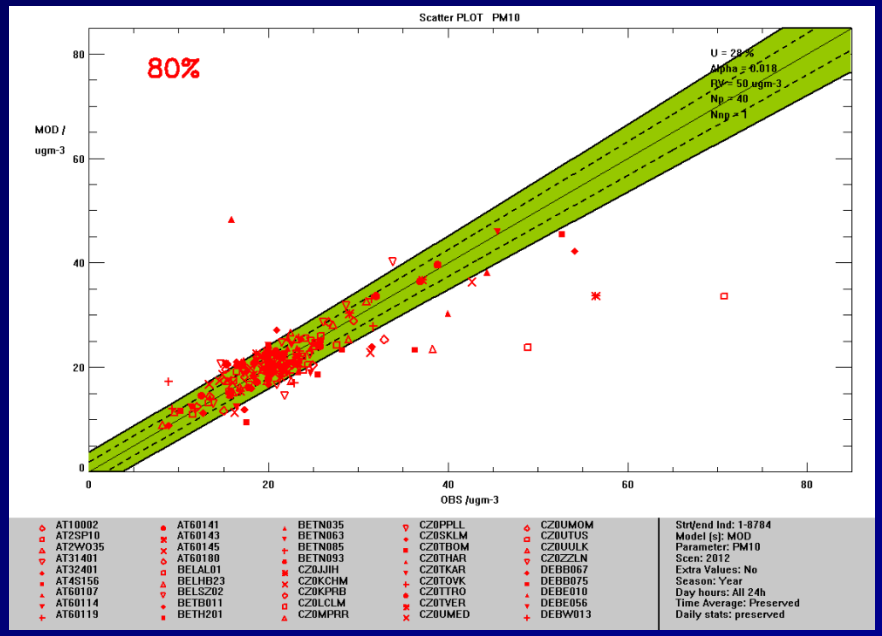
Evaluation by ***cross-validation***: similar level of relative uncertainty (for annual average slightly lower)

Parameter	PM_{10} annual average				PM_{10}, 36^{th} highest daily mean			
	Full set		Assimil. subset		Full set		Assimil. subset	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
RMSE [µg.m^{-3}]	3.8	6.1	4.4	7.4	7.7	11.9	8.7	14.9
Relative RMSE	21.4%	22.1%	25.0%	27.1%	24.5%	24.5%	27.5%	30.6%
Bias [µg.m^{-3}]	0.1	0.0	0.7	0.2	0.1	-0.1	1.0	-1.1

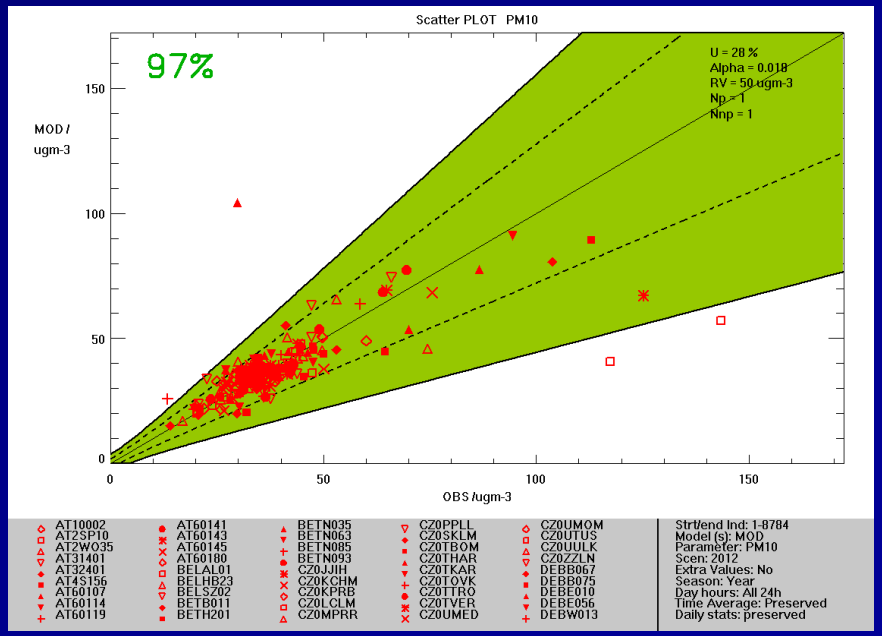
Evaluation of PM_{10} annual maps for 2012 using Delta 5.0

Map created using assimilation subset of stations evaluated against validation subset of stations (of all background types).

Annual average



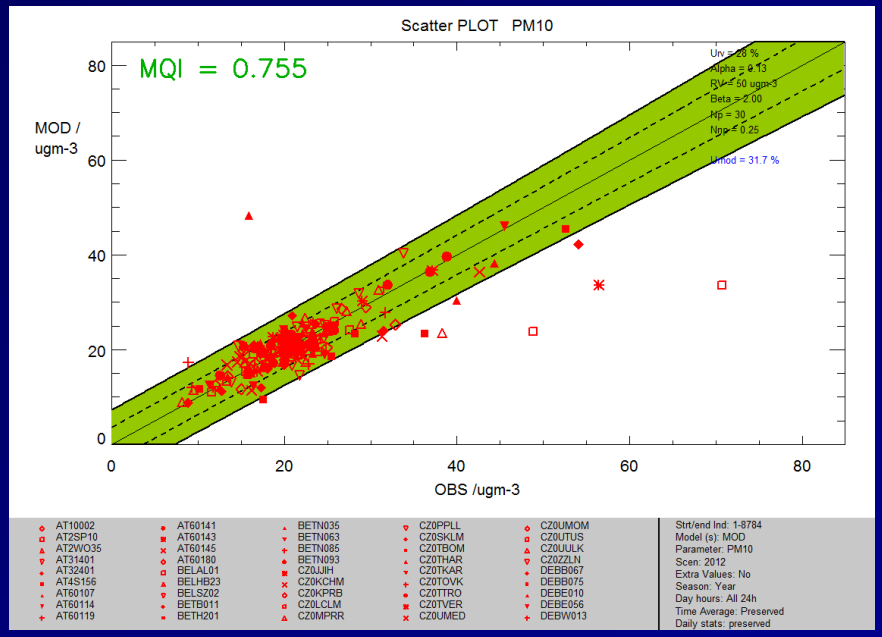
36th highest daily mean



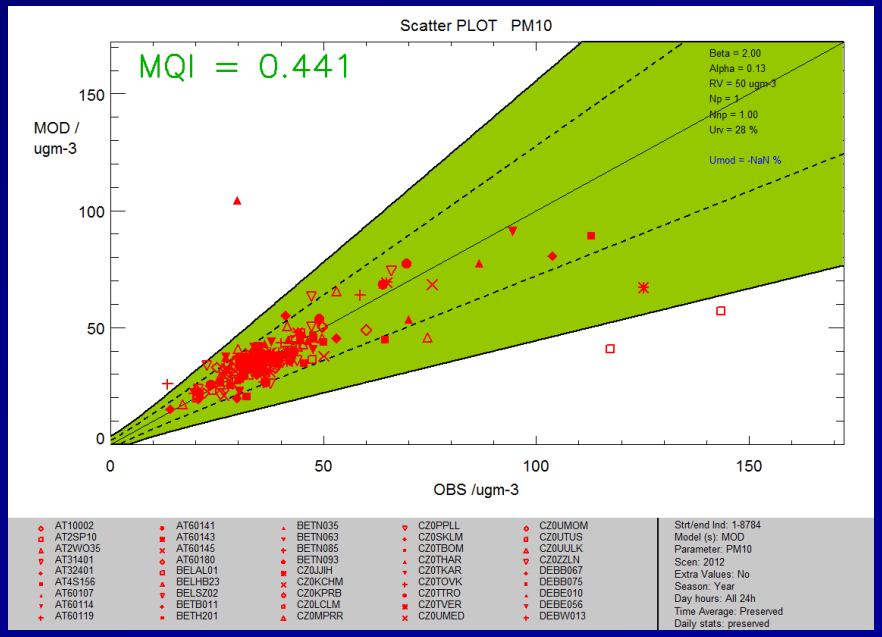
Evaluation of PM_{10} annual maps for 2012 using Delta 5.3

Map created using assimilation subset of stations evaluated against validation subset of stations (of all background types)

Annual average

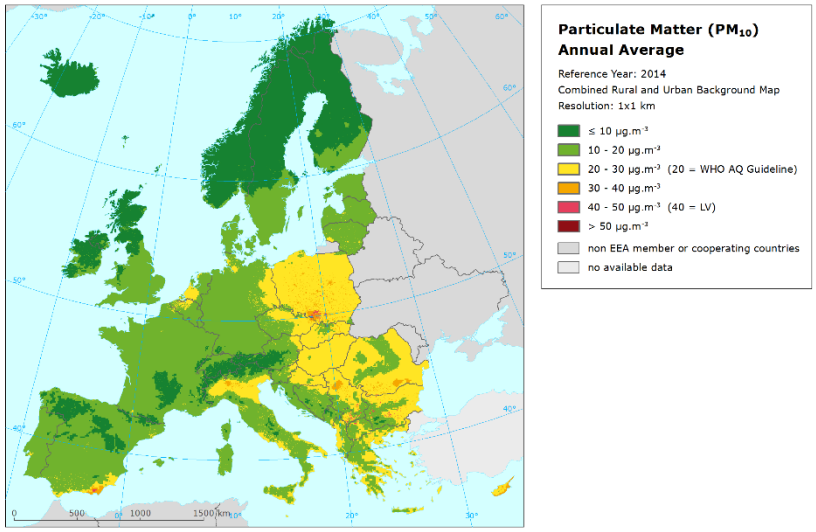


36th highest daily mean

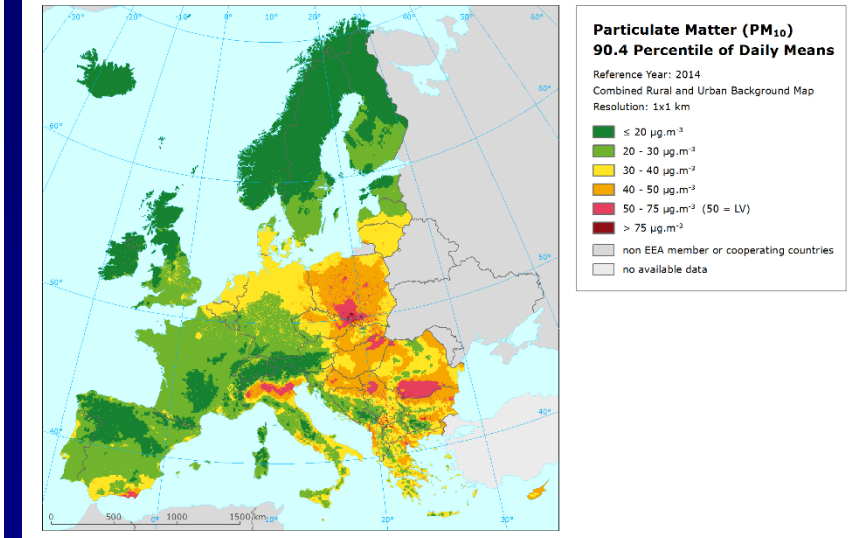


PM_{10} annual indicator maps for 2014

Annual average



90.4 percentile of d. means



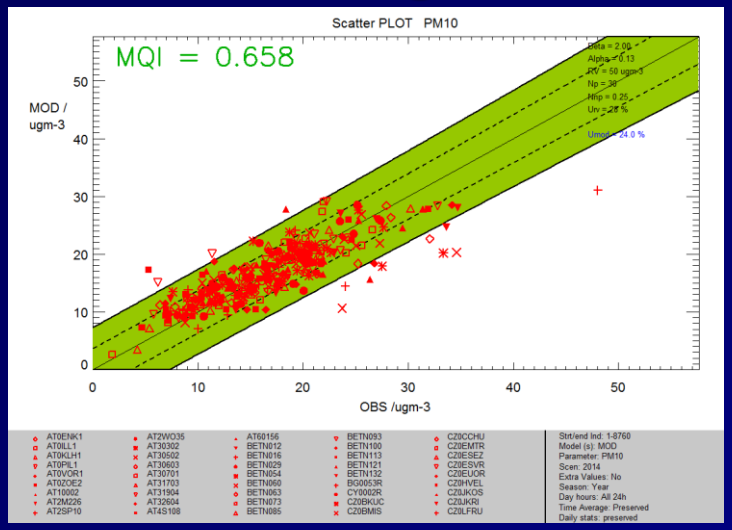
Evaluation by ***cross-validation***: similar level of relative uncertainty (for annual average slightly lower)

Parameter	Annual average		90.4 percentile of daily means	
	Rural areas	Urban areas	Rural areas	Urban areas
RMSE [µg.m^{-3}]	3.5	4.2	6.5	8.6
Relative RMSE	20.7%	17.7%	21.5%	20.4%
Bias (MPE) [µg.m^{-3}]	0.1	0.0	0.2	-0.1

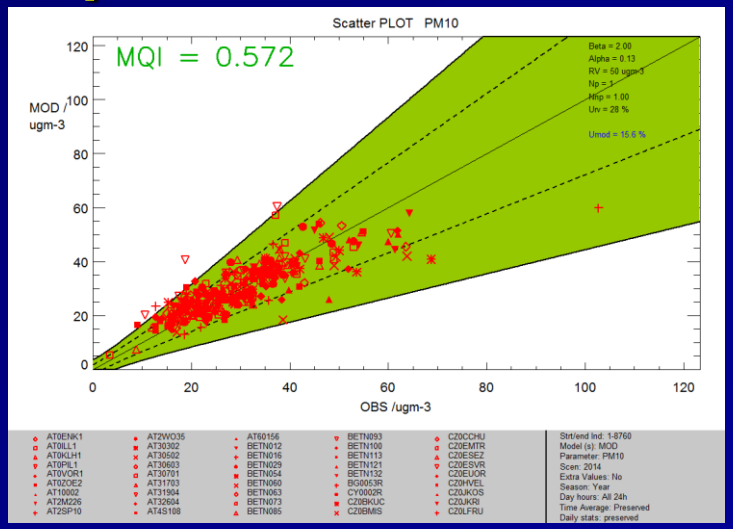
Evaluation of PM_{10} annual maps for 2014 using Delta 5.4

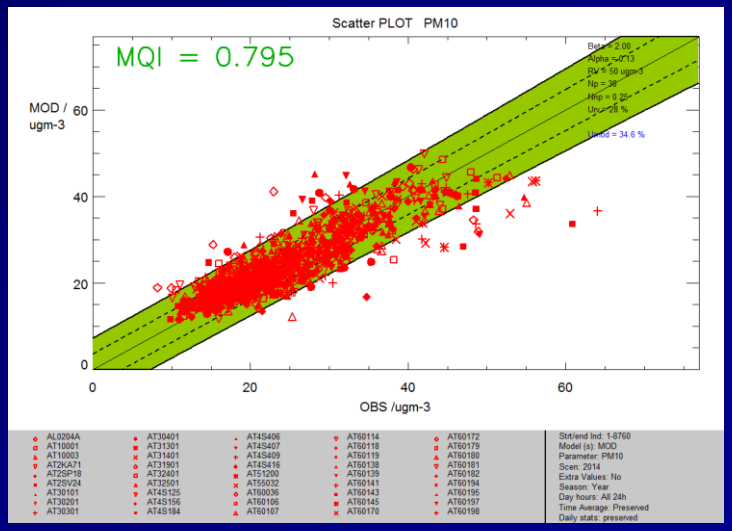
Cross-validation estimates against all stations of different types

Annual average ***90.4 percentile of d. means***

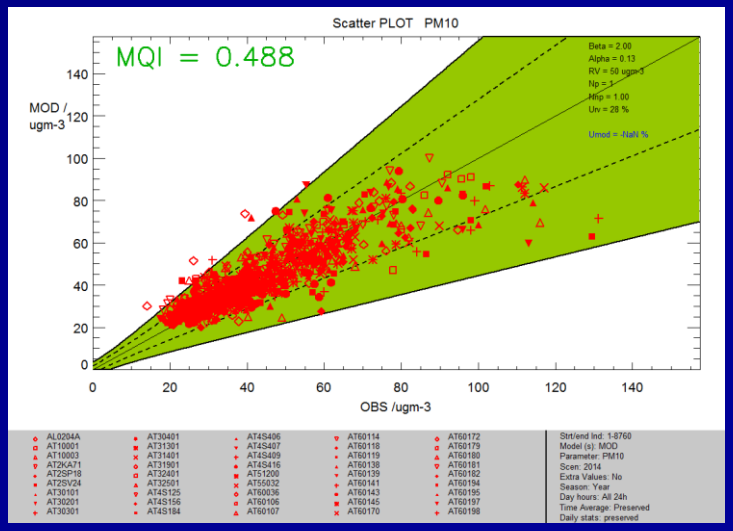


Rural areas





Urban backgr. areas



1. Motivation

2. MQO for annual averages vs. for percentiles

3. Observation uncertainty in MQO

4. Suggestions concerning MQO for percentiles

Model quality objective (MQO) formulation

MQO for hourly and daily data:

$$\text{MQO: } \text{MQI} = \frac{|O_i - M_i|}{\beta U_{95}(O_i)} \leq 1$$

where O_i ... observation value, M_i ... modelled value
$U_{95}(O_i)$ ... measurement expanded uncertainty

MQO should be fulfilled for 90% of the stations.

Measurement uncertainty is a ***key input to MQO***.

MQO for the ***annual average*** data:

$$MQI = \frac{|\bar{O} - \bar{M}|}{\beta U_{95}(\bar{O})} \leq 1$$

Annual average uncerainty – ***reduced*** compared to $U_{95}(O_i)$.

MQO (or MPC) for the ***percentile*** data:

$$MPI_{perc} = \frac{|M_{perc} - O_{perc}|}{\beta U_{95}(O_{perc})} \leq 1$$

Percentile uncerainty – ***not reduced*** compared to $U_{95}(O_i)$.

Measurement uncertainty for percentiles

Motivation for non reduction of percentile uncertainty:
The percentile – considered as a single observation, percentile uncertainty – considered as the uncertainty of the observation corresponding to the relevant percentile.

However, percentile is an ***annual indicator***.

Percentile – based on ranked data

$$\ldots < O_{|i-1|} < O_{|i|} = O_{perc} < O_{|i+1|} < \ldots$$

When the uncertainties are taken into account, the rank can be changed:

$$\ldots <? O_{|i-1|} \pm U(O_{|i-1|}) <? O_{|i|} \pm U(O_{|i|}) <? O_{|i+1|} \pm U(O_{|i+1|}) <? \ldots$$

Thus: measurement uncertainty reduction should be applied.

Which percentiles?

Based on EU legislation (Directive 2008/50/EC).

Particularly two percentiles used in EEA Air Quality Reports:

90.4 percentile of daily means in one year for ***PM_{10}***
(corresponds to 36th highest daily mean)

93.2 percentile of maximum daily 8-hour means in one year for ***ozone***
(corresponds to 26th highest maximum daily 8-hour mean)

Further, only these two pollutants considered, with the emphasis on PM_{10}.

1. Motivation

2. MQO for annual averages vs. for percentiles

3. Observation uncertainty in MQO

4. Suggestions concerning MQO for percentiles

Measurement uncertainty expression

Uncertainty expression (proposed in: Thunis et al. 2013, Pernigotti et al., 2013) – based on the ***assumption*** that the uncertainty of each measurement is composed of a component proportional to the concentration level and a non-proportional component, ***as in***

$$u_c^2(O_i) = u_p^2(O_i) + u_{np}^2(O_i) = (1-\alpha^2)\,(u_r^{RV} O_i)^2 + \alpha^2 (u_r^{RV} RV)^2$$

where *RV* is reference level.

Based on this – ***expanded uncertainty*** (as used in MQI):

$$U_{95}(O_i) = k u_c^{RV} \sqrt{(1-\alpha^2)O_i^2 + \alpha^2 . RV^2} = U_{95,r}^{RV} \sqrt{(1-\alpha^2)O_i^2 + \alpha^2 . RV^2}$$

where *k* is so-called coverage factor.

Measurement uncertainty expression - continuation

Uncertainty for annual average – expected to be ***reduced*** compared to the uncertainties associated to the raw measurements. To cover this – parameters parameters N_p and N_{np} are introduced:

$$U_{95}(\bar{O}) = U_{95,r}^{RV}\sqrt{\frac{(1-\alpha^2)}{N_p}\bar{O}^2 + \frac{\alpha^2 . RV^2}{N_{np}}}$$

In the Delta tool, all the parameters used in Equations 1 – 3 have been already ***predefined***.

Their values – estimated in Thunis et al. (2013) for ozone and in Pernigotti et al. (2013) for PM_{10}. The uncertainty $U_{95,r}^{RV}$ for PM_{10} – based on the reference gravimetric method.

Parameters for measurement uncertainty calculation

Up to Delta 5.2:

Pollutant	$U_{95,r}^{RV}$	RV	α	N_p	N_{np}
PM_{10}	0.28	50 µg.m^{-3}	0.13	40	1
Ozone	0.09	120 µg.m^{-3}	0.79	4 (a)	1 (a)

(a) In older versions of the Delta, "NA" was stated.

Then, parameters for annual averages have been updated.

Since Delta 5.3:

Pollutant	$U_{95,r}^{RV}$	RV	α	N_p	N_{np}
PM_{10}	0.28	50 µg.m^{-3}	0.13	30	0.25
Ozone	0.09	120 µg.m^{-3}	0.79	11	3

MQO – highly sensitive to these parameter values.

Parameters for annual average uncertainty reduction

Annual average uncertainty reduction – for PM_{10}, discussed in Pernigotti et al. (2013), the update discussed in Delta 5.3 User‘s Guide.

For ozone, introduced in Delta 5.2 and Delta 5.3 User‘s Guides, but not discussed.

For PM_{10}, Pernigotti et al. state: „*If time series were composed by N measurements of independent consecutive air samples then the uncertainty of the average concentration would be reduced by a factor sqrt(N). But this is not the case because values within an air pollutant time series are correlated by the errors of the measurement method and by the trends in consecutive pollutant measurements.“*

(Pernigotti, D., Thunis, P., Gerboles, M., Belis. C., 2013, ‘Model quality objectives based on measurement uncertainty: Part II: PM10 and NO2’, Atmospheric Environment 79, 869-878).

Further, they introduce an alternative approach.

Parameters for PM_{10} ann. average uncertainty reduction

Instead of the equation

$$U(\overline{O}) = \frac{U(O_i)}{\sqrt{N}} = ku_r^{RV}\sqrt{\frac{(1-\alpha^2)\cdot O_i^2 + \alpha^2 \cdot RV^2}{N}}$$

they ***introduce another one***, i.e.

$$U(\overline{O}) = ku_r^{RV}\sqrt{\frac{(1-\alpha^2)}{N_p}\overline{O}^2 + \frac{\alpha^2 \cdot RV^2}{N_{np}}}$$

Further, they ***estimate*** the values of N_p and N_{np} (considered now as constants) ***empirically***, using so-called „GDE method“ (based on uncertainty concepts of „Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods“), based on ***5 pairs of yearly averaged AirBase data***.

In Delta 5.3 User‘s Guide, an ***update*** of N_p and N_{np} is introduced (without provided details), in order „to reflect uncertainties associated to the β-ray measurement technique“.

1. Motivation

2. MQO for annual averages vs. for percentiles

3. Observation uncertainty in MQO

4. Suggestions concerning MQO for percentiles

Suggestions concerning MQO for percentiles

In principle, a similar approach like in the case of the annual average could be used, i.e.

$$U(O_{perc}) = ku_r^{RV} \sqrt{\frac{(1-\alpha^2)}{N_{p,perc}} \overline{O}^2 + \frac{\alpha^2 \cdot RV^2}{N_{np,perc}}}$$

In such a case, $N_{p,perc}$ and $N_{np,perc}$ should be in principle dependent on the pollutant and on the level of the percentile.

For estimating $N_{p,perc}$ and $N_{np,perc}$, similar ***empirical*** approach like for annual average might be used. Currently, ***55 pairs*** of PM_{10} percentile (as well as annual average) data for 2014 are stored in the AQ e-reporting database (operated by EEA).

Parallel to the percentile parameter estimation, the parameters for PM_{10} annual average might be updated.

For O_3 percentile, similar parameter estimation like for O_3 annual average might be used.

Thank you for your attention.