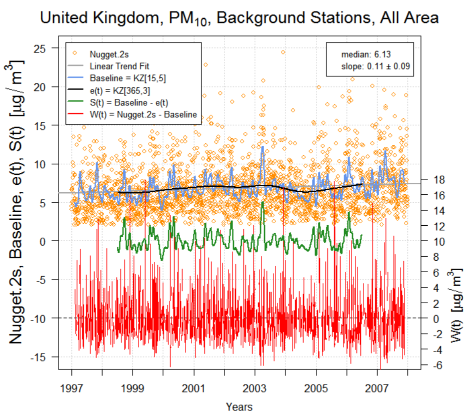
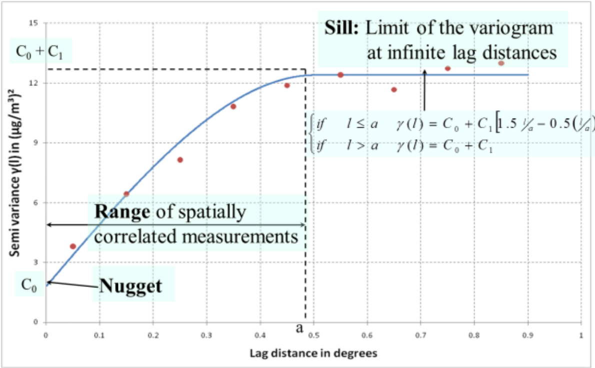
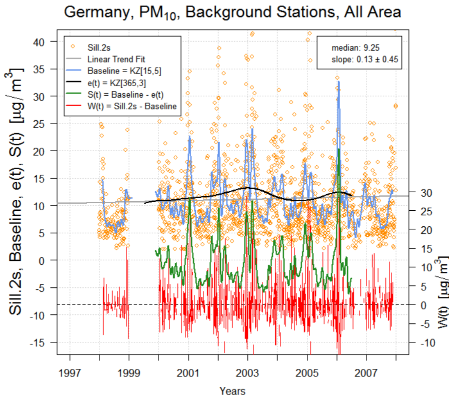
Trends of Daily Variogram Parameters Derived by Geostatistical Analysis of AirBase PM_{10} Time Series







Oliver Kracht and Michel Gerboles

European Commission - Joint Research Centre
I – 21026 Ispra (VA)
www.jrc.ec.europa.eu

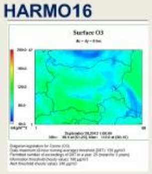
Joint Research Centre

Session T1:
Model Evaluation and Quality Assurance

HARMO 16, 8 - 11 September 2014
Varna - Bulgaria

16th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes

HARMO16



8-11 September 2014, Varna, Bulgaria

www.harmo.org/harmo16
harmo16@harmo.org

Objectives:

- Evaluate geo-referenced air quality measurements to obtain information about:
  - small scale variability and measurement uncertainty
  - spatial representativeness of monitoring stations

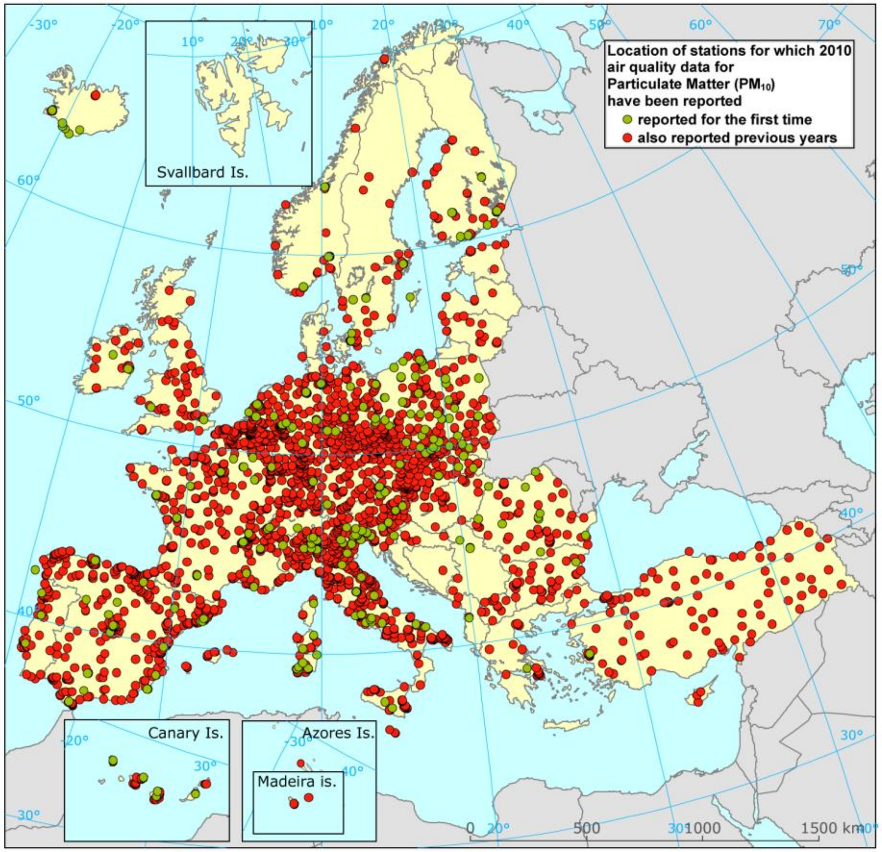
Methodology:

- Variography and time series analysis

Applications:

- AirBase records of daily PM_{10} values

Data availability in Airbase:

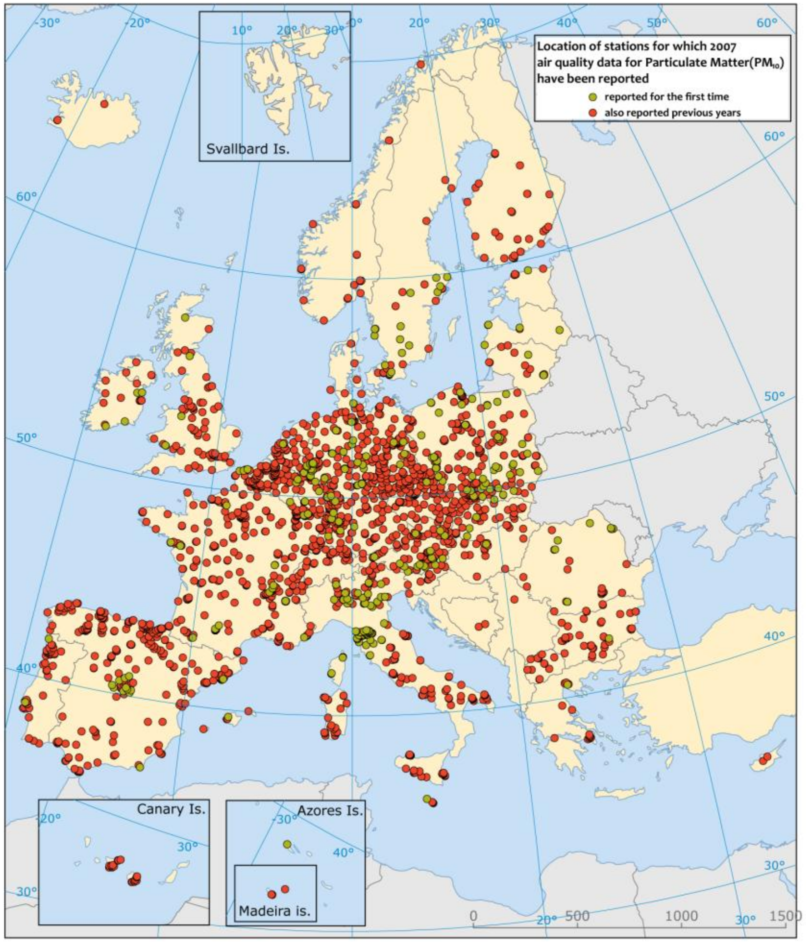


Source: ETC/ACM AirBase

february 2012

- Public air quality database system of the European Environment Agency (EEA)
- Monitoring data submitted by about 35 participating countries throughout Europe
- 140 pollutants, more than 6000 stations and 25000 time series with hourly and daily data of more than 30 years

Scope of this Exercise:



Source: ETC/ACC AirBase

- 1997 /2007 records from Airbase version 4
- Daily PM_{10} values
- Station type “Background”
- All area types (urban, suburban and rural)
- Examples for FR, DE, GB, AT, IT and NL

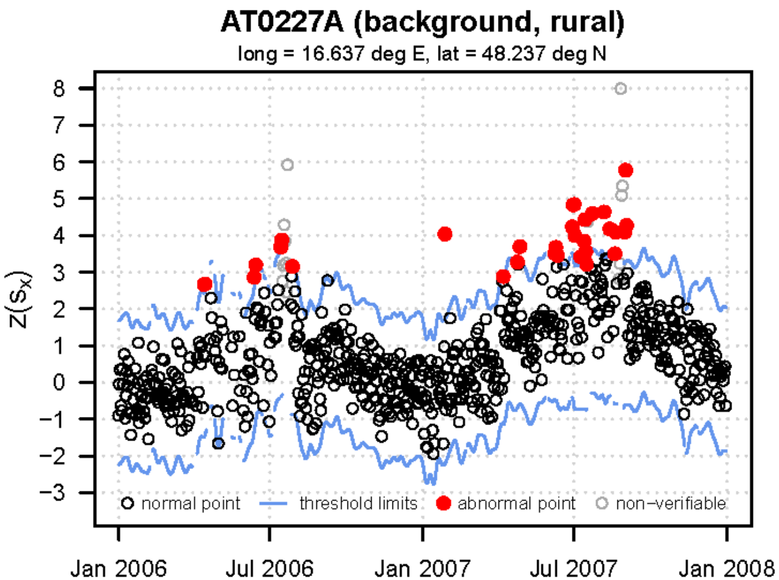
Workflow:

- Query AirBase for PM_{10} daily values of background stations (low small scale variability)
- Pre-treatment: spatio-temporal outlier screening
- Variogram fits (automated procedure, spherical model)
- Screening of variogram fits for internal consistency
- Kolmogorov-Zurbenko filtering method to separate the time series into their short-, mid and long-term components
- Trend analysis (linear regression)
- Significance of trends (conf. intervals and Mann-Kendall test)
- Improved confidence intervals by block bootstrap approach (to deal with serial autocorrelation in the time series)
- Comparison with (direct) field studies on measurement uncertainty

- Query AirBase for background stations PM_{10} (AirBase v.4)

Country	First Available Fit (year)	Last Available Fit (year)	Available Fits (count)
FR	2001	2007	2051
DE	1998	2007	3232
GB	1997	2007	3932
AT	2001	2007	2280
IT	2003	2007	1737
NL	2003	2007	1258

- Data pre-treatment: outlier screening



More info:

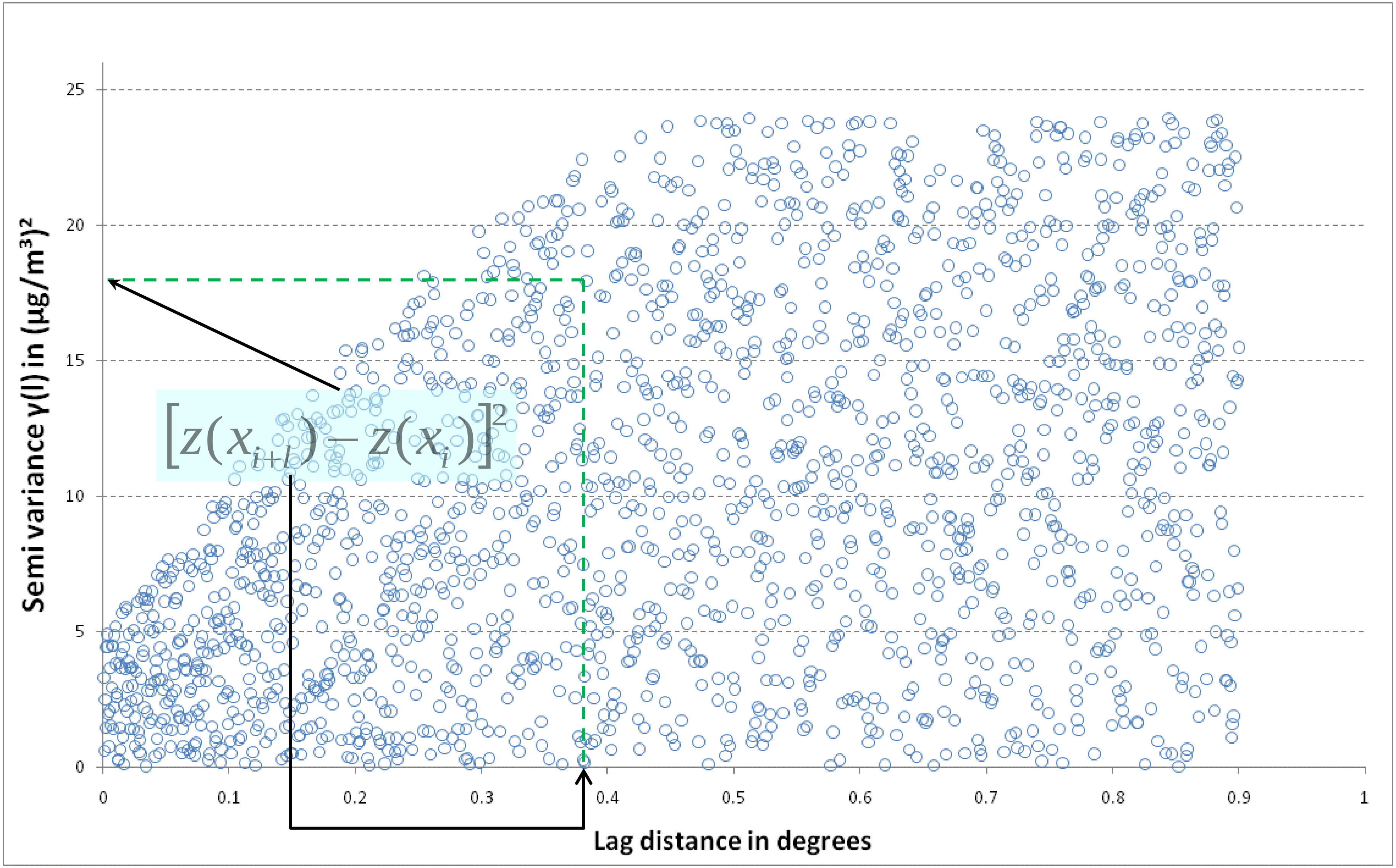
Kracht, O., H. I. Reuter and M. Gerboles, 2013: A Tool for the Spatio-Temporal Screening of AirBase Datasets for Abnormal Values. JRC Technical Reports, EUR 25787 EN, DOI 10.2788/81552, 209 pp.

Kracht, O. and M. Gerboles, 2013: Screening of Spatio-Temporal Anomalies in Long Term / Large Scale Air Quality Monitoring Time Series. JRC Technical Reports, EUR 26462 EN, DOI: 10.2788/60933, 382 pp.

Workflow:

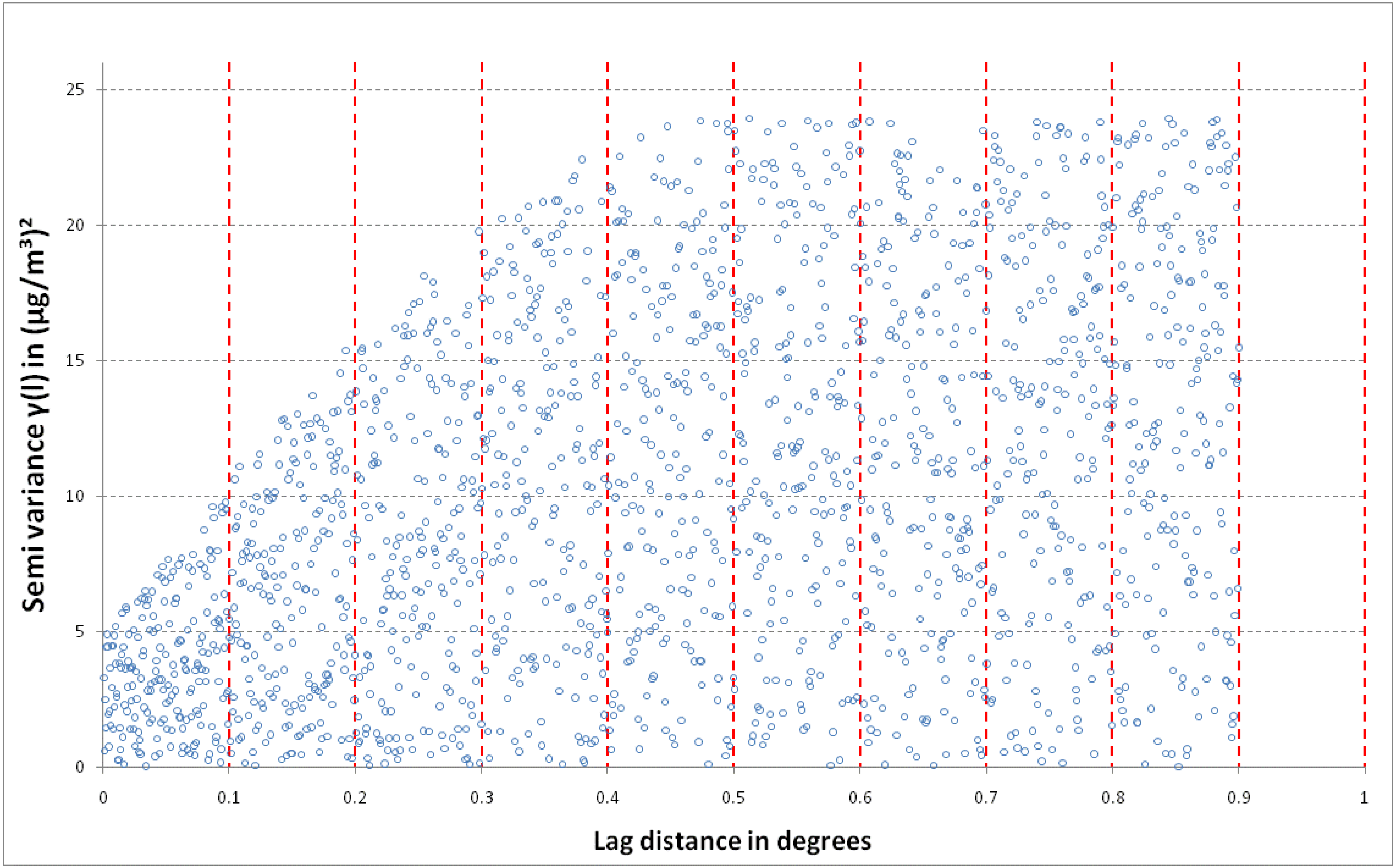
- Query AirBase for PM_{10} daily values of background stations (low small scale variability)
- Pre-treatment: spatio-temporal outlier screening
- **Variogram fits (automated procedure, spherical model)**
- Screening of variogram fits for internal consistency
- Kolmogorov-Zurbenko filtering method to separate the time series into their short-, mid and long-term components
- Trend analysis (linear regression)
- Significance of trends (conf. intervals and Mann-Kendall test)
- Improved confidence intervals by block bootstrap approach (to deal with serial autocorrelation in the time series)
- Comparison with (direct) field studies on measurement uncertainty

Variography



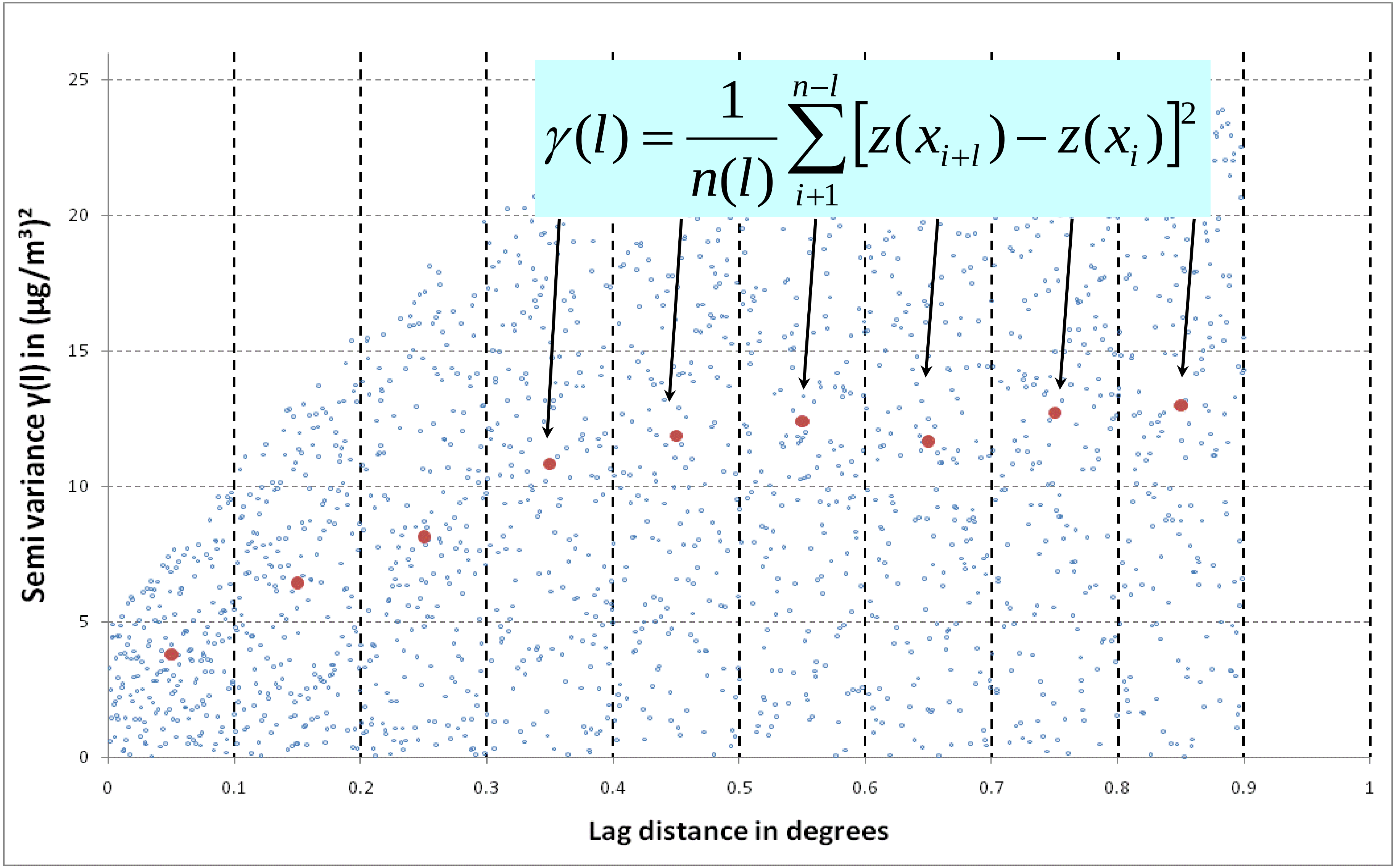
source: explanation of variography techniques, from M. Gerboles (2007): AQUILA Workshop presentation

Variography



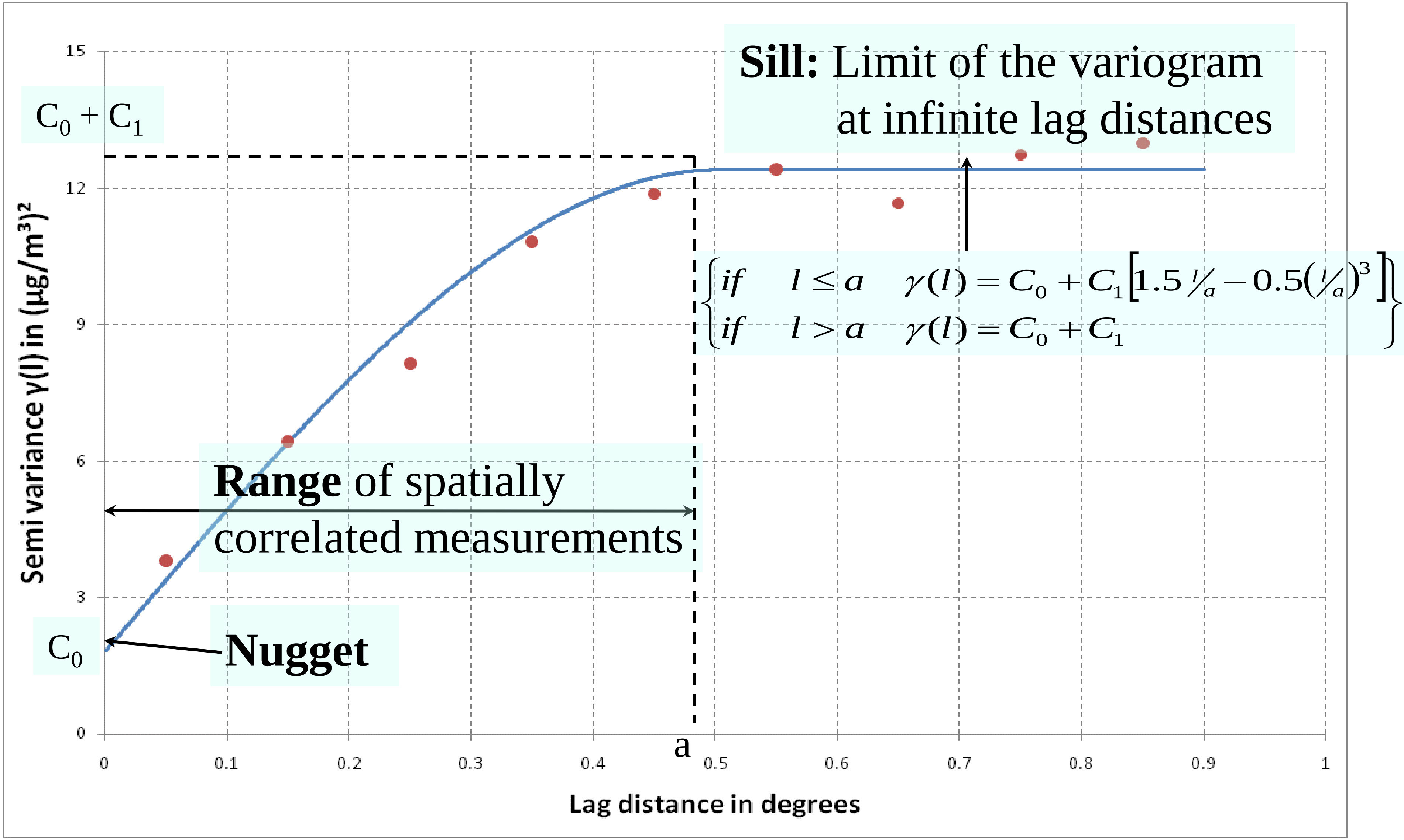
source: explanation of variography techniques, from M. Gerboles (2007): AQUILA Workshop presentation

Variography



source: explanation of variography techniques, from M. Gerboles (2007): AQUILA Workshop presentation

Variography



source: explanation of variography techniques, from M. Gerboles (2007): AQUILA Workshop presentation

Variography

European Commission

Primary interest within this study:

1) The **nugget variance** is reflecting fluctuations of the measurements at very short distance (towards 0).

$$s^2_{nugget} = s^2_{meas} + s^2_{sc}$$

uncertainty of measurement
variance associated with the sampling and analytical variability

micro-scale variance
variability that occurs at distances lower than the shortest sampling distance (continuity)

Variography

Primary interest within this study:

1) The **nugget variance** is reflecting fluctuations of the measurements at very short distance (towards 0).

$$s^2_{nugget} = s^2_{meas} + s^2_{sc}$$

2) The **sill** and the **range** parameter can inform about the extend of spatial correlation (spatial representativeness of monitoring stations).

Workflow:

- Variogram fits (automated procedure, spherical model)
- **Screening of variogram fits for internal consistency**

Constraints for accepting a valid variogram model fit:

(I)1 < nugget < 150 $(\mu g/m^3)^2$

(II) 0 < sill < 10^4 $(\mu g/m^3)^2$

(III) 0 < range < 2 deg

(IV) 0.04 < sill / nugget < $5 \cdot 10^3$

Workflow:

- Variogram fits (automated procedure, spherical model)
- **Screening of variogram fits for internal consistency**

Country	First Available Fit (year)	Last Available Fit (year)	Available Fits (count)	Accepted Fits (count)
FR	2001	2007	2051	1221
DE	1998	2007	3232	1555
GB	1997	2007	3932	2415
AT	2001	2007	2280	1189
IT	2003	2007	1737	890
NL	2003	2007	1258	1235

Time series $x(t)$

Time series decomposition

W(t): short-term component

- variations of weather
- short-term fluctuations in precursor emissions

S(t): mid-term (seasonal) component

- changes in the solar angle (induced variations of emissions & temperature dependencies)

e(t): long-term signal

- changes in overall emissions, pollutant transport, climate, economics, and environmental policies
- evolutions in the operational principles of the monitoring network

$$X(t) = \underbrace{e(t) + S(t)}_{\textit{Baseline}} + W(t)$$

Baseline: defined as the sum of the long-term and seasonal component

Time series decomposition

4) $W(t) = x(t) - Baseline$

short-term component

3) $S(t) = Baseline - e(t)$

mid-term (seasonal) component

2) $e(t) = KZ_{365,3}(x)$

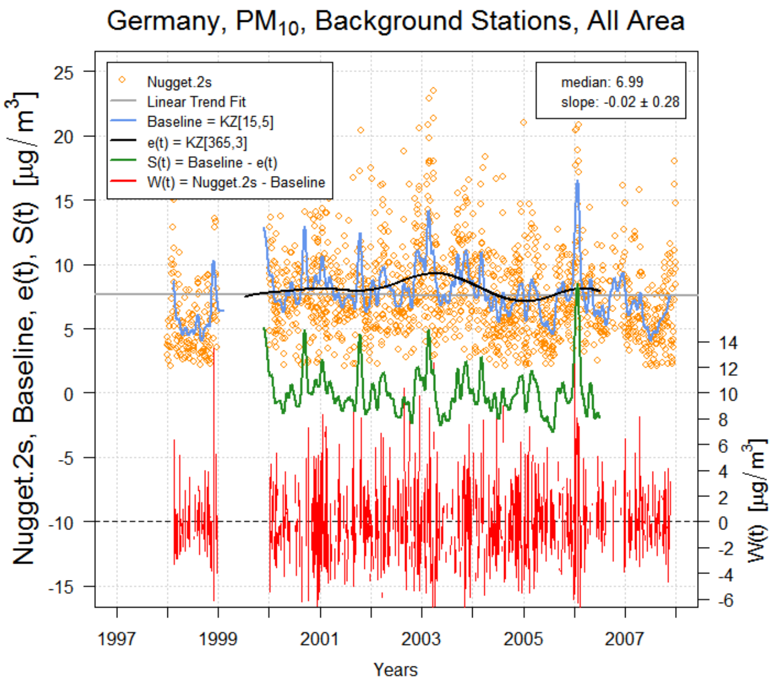
long-term signal

$$X(t) = \underbrace{e(t) + S(t)}_{Baseline} + W(t)$$

1) $Baseline = KZ_{15,5}(x)$

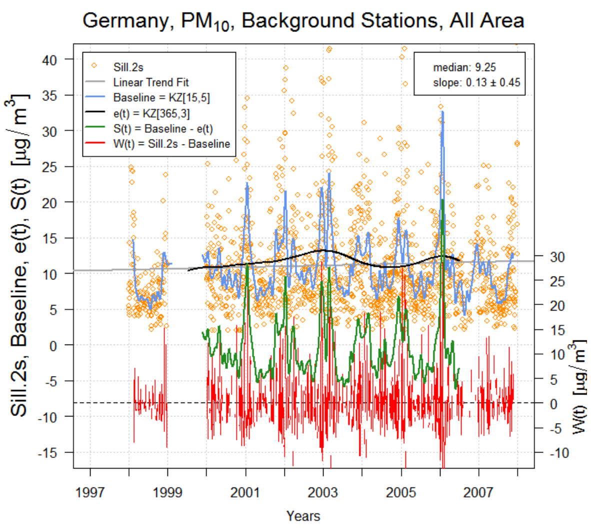
Results & Observations

Nugget time series

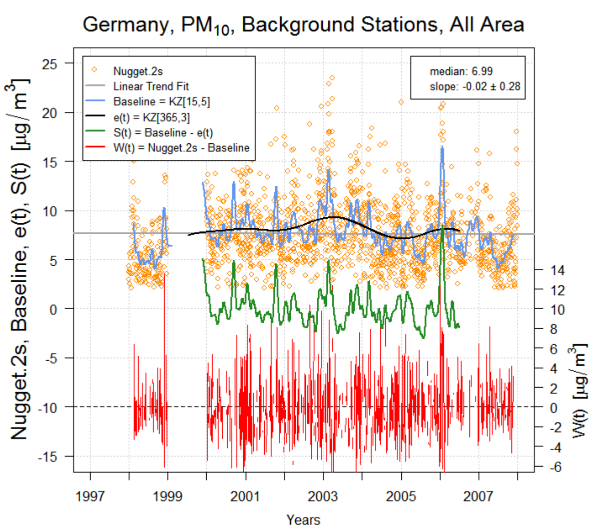


Results & Observations

Sill

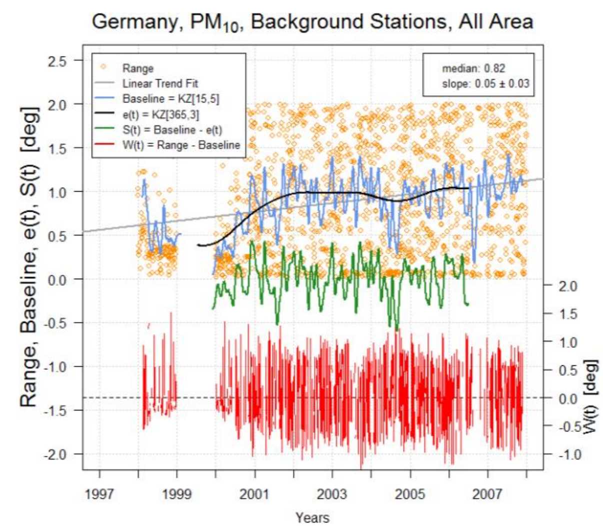


Nugget

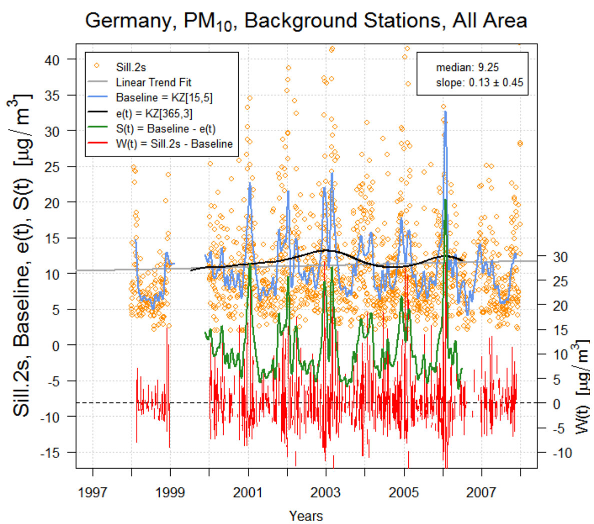


Results & Observations

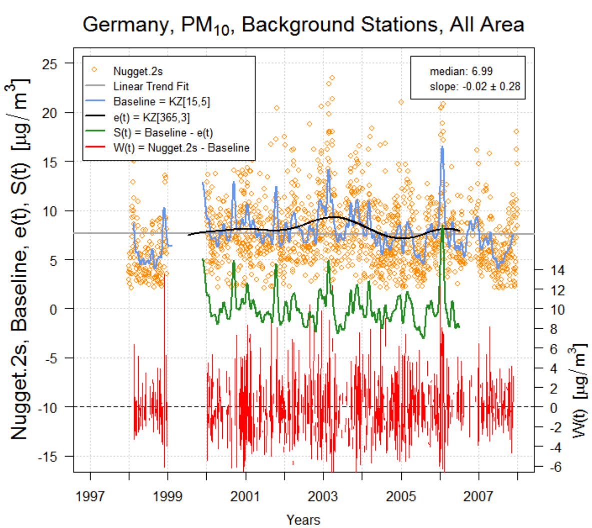
Range



Sill

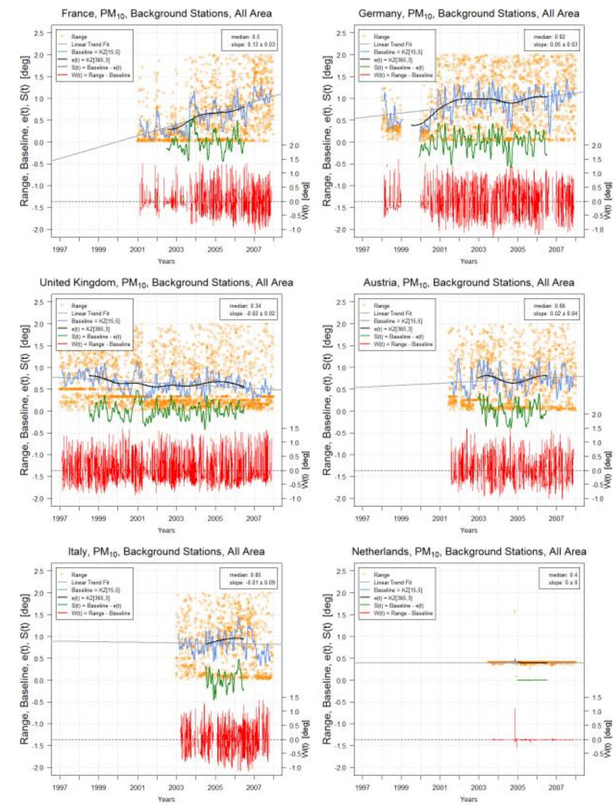


Nugget

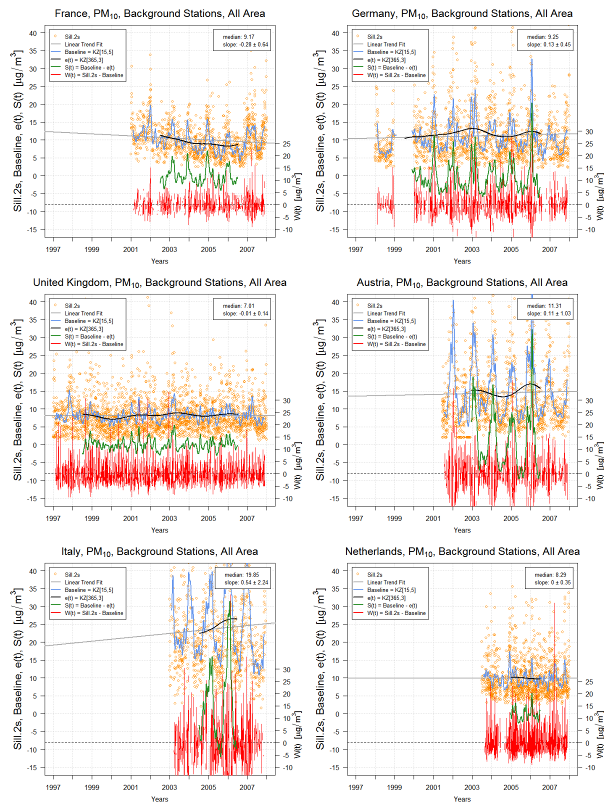


Results & Observations

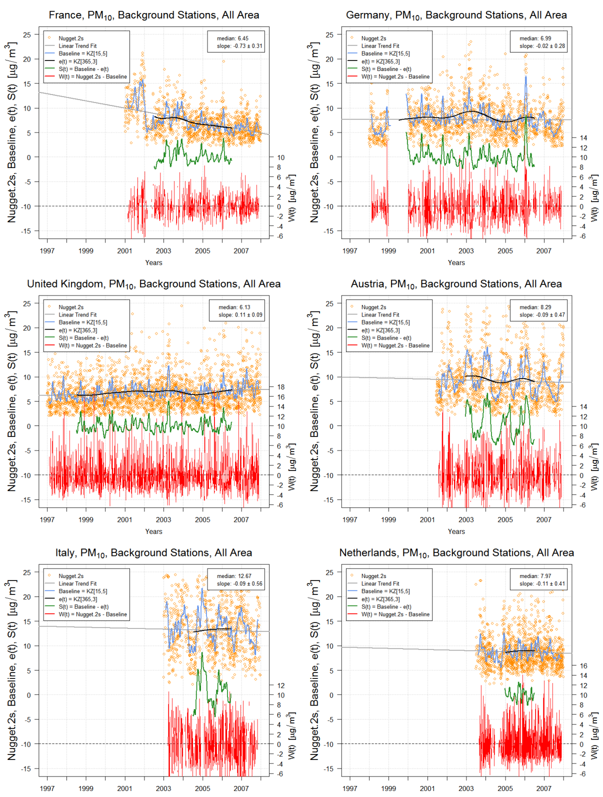
Range



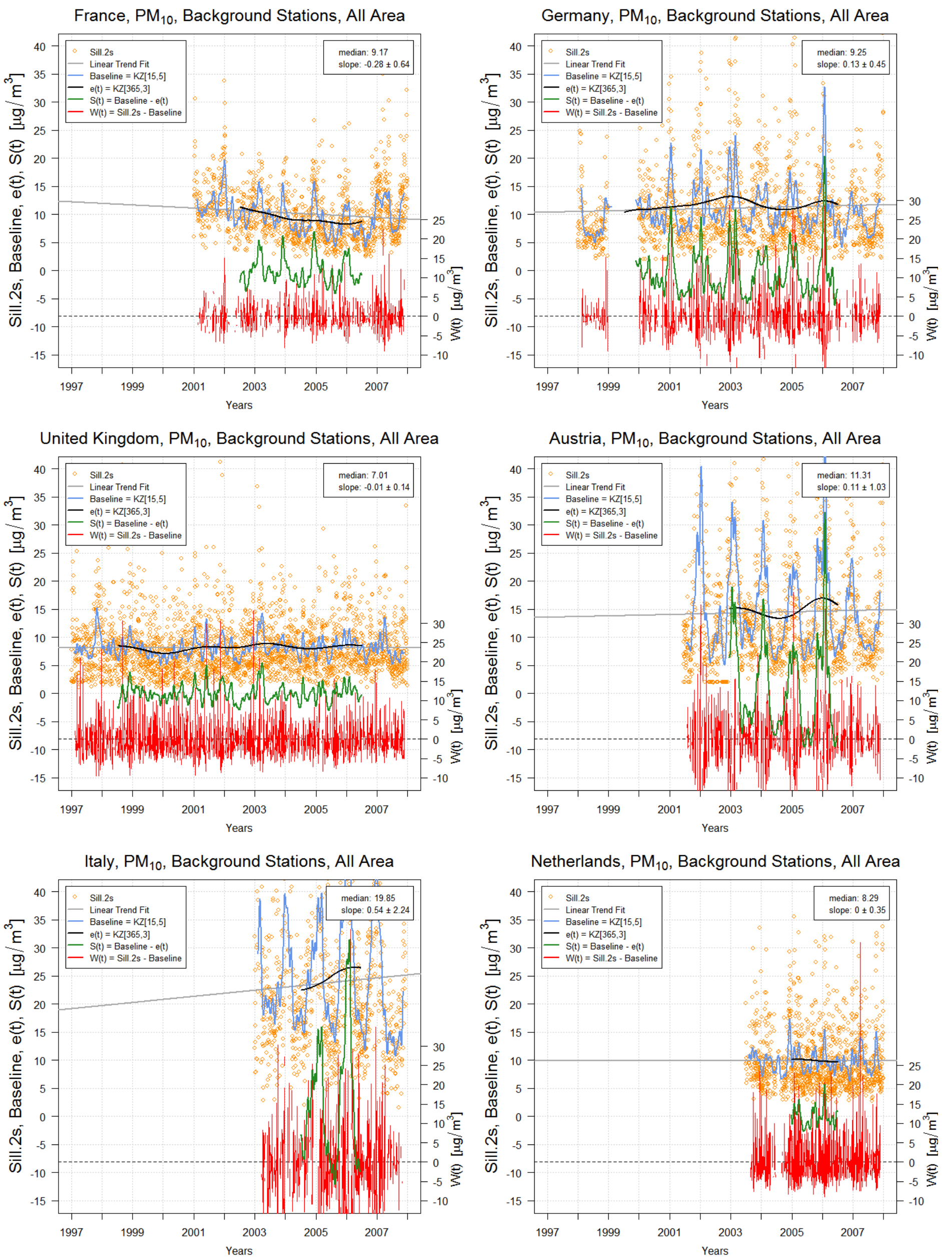
Sill



Nugget



Results & Observations



Sill parameter

W(t):

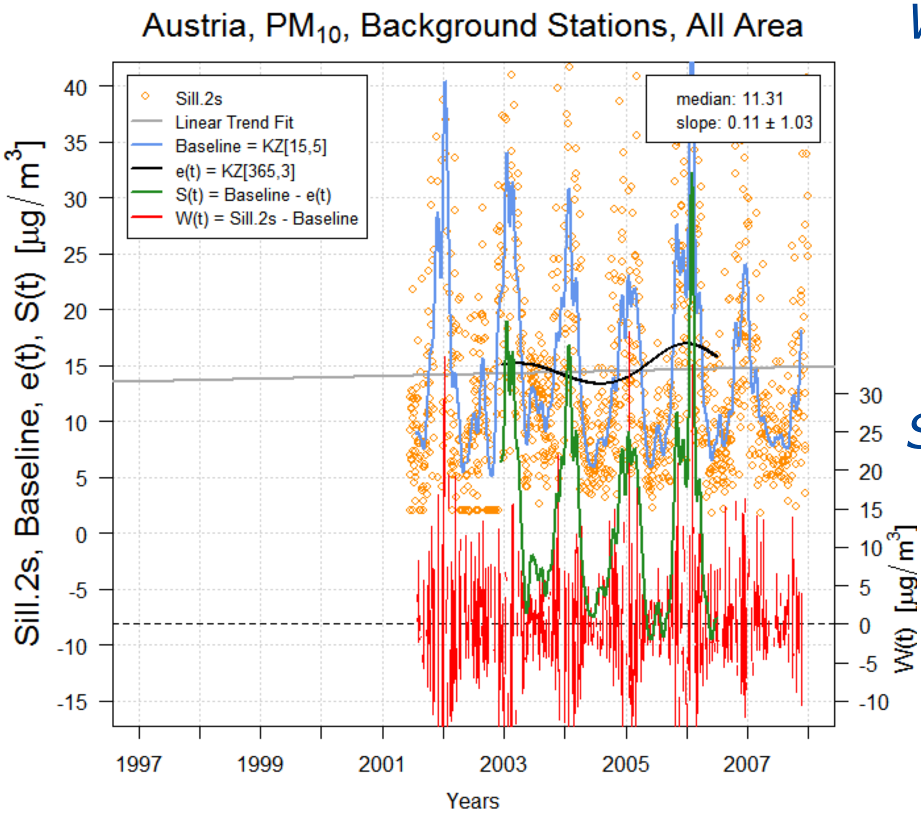
- non-stationarities (changes of variance over time)
- indicates temporal variations in the macro-scale spatial correlation structures

S(t):

- pronounced cyclic behaviour
- phase relationship (winter increase) for FR, DE, AT and IT
- macro-scale spatial variability increases in winter because of
  - local emission caused by heating, sanding and PM_{10} re-suspension
  - limited air mixing

Results & Observations

Sill parameter

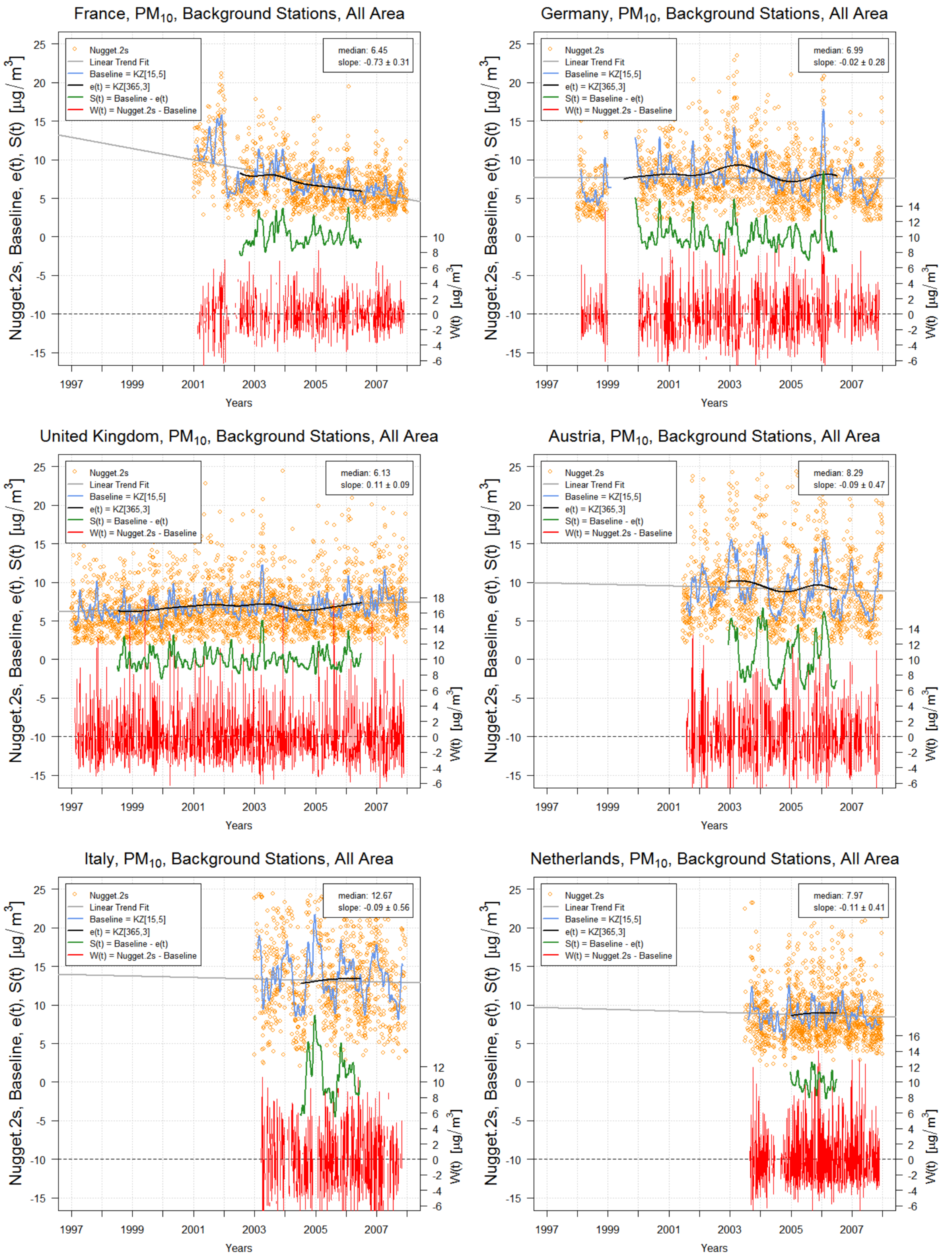


W(t):

- non-stationarities (changes of variance over time)
- indicates temporal variations in the macro-scale spatial correlation structures

S(t):

- pronounced cyclic behaviour
- phase relationship (winter increase) for FR, DE, AT and IT
- spatial variability increases in winter because of
  - local emission caused by heating, sanding and PM_{10} re-suspension
  - limited air mixing



Results & Observations

Nugget parameter

W(t):

- non-stationarities (changes of variance over time)
- indicates temporal variations in the macro-scale spatial correlation structures

S(t):

- phase relationship (winter increase) for <u>AT and IT</u>
- **small**-scale spatial variability increases in winter
- short term influences and seasonal fluctuations of measurement uncertainty (?)

Results & Observations

Comparison with PM_{10} field uncertainty

Country	Parameter	Median Value ($\mu g/m^3$)
FR	Nugget (2s)	6.45
DE	Nugget (2s)	6.99
GB	Nugget (2s)	6.13
AT	Nugget (2s)	8.29
IT	Nugget (2s)	12.67
NL	Nugget (2s)	7.97
FR	Sill (2s)	9.17
DE	Sill (2s)	9.25
GB	Sill (2s)	7.01
AT	Sill (2s)	11.31
IT	Sill (2s)	19.85
NL	Sill (2s)	8.29

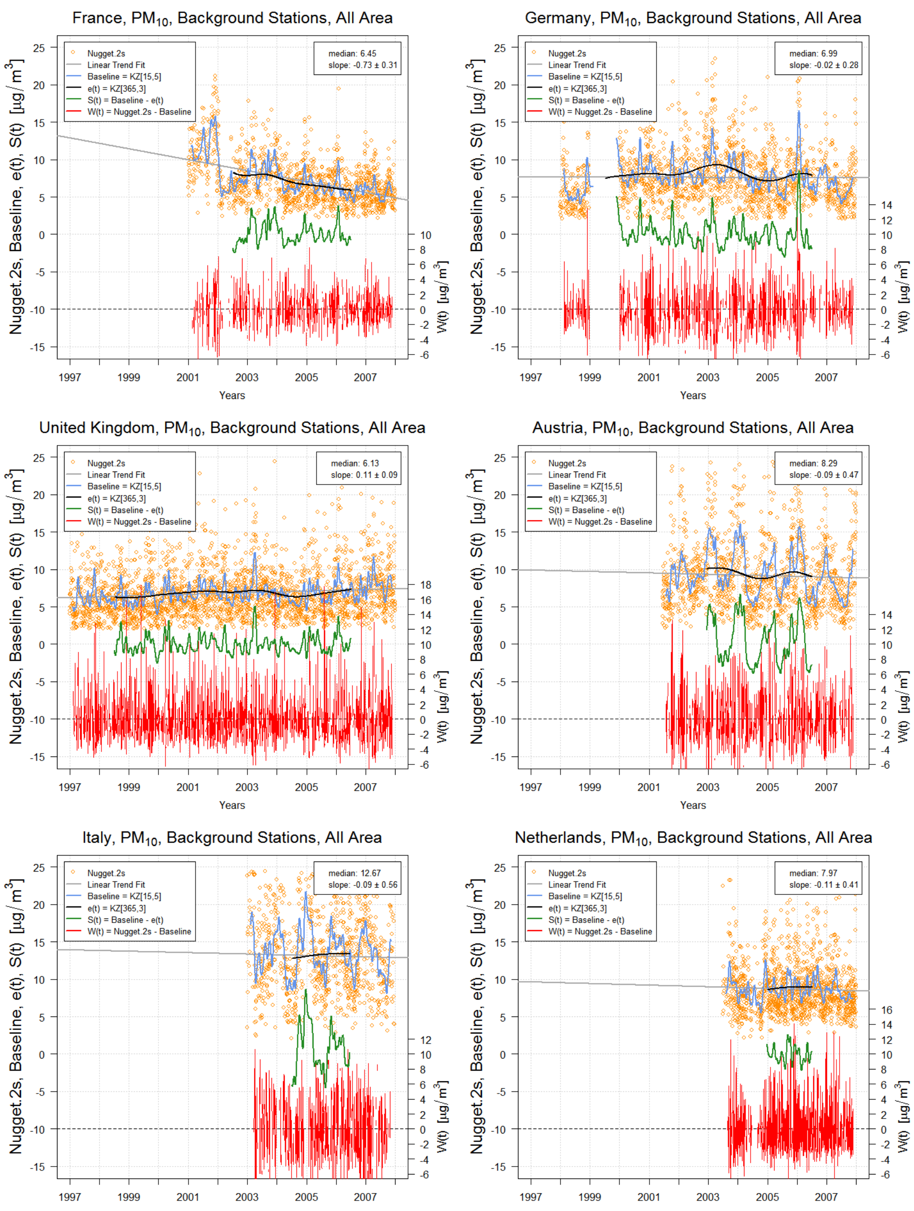
Comparison with PM_{10} field uncertainty

	Sampler type, flow, filter or line temperature	$U_{relative}$* (field)	Mean $U_{relative}$ (field)	Nugget (2s) $U(PM_{10}) / PM_{10}$
AT (BG-UR)	Digitel, 30 m^3/h, Whatman QMA	11 %	16%	17%
	Teom FDMS, 1 m^3/h	18 %		
	Digitel 30 m^3/h, Pall Quartz	31 %		
	Eberline FH 62 IR, 1 m^3/h	17 %		
	Digitel 30 m^3/h, Ederol Glass fibre	9 %		
	Teom FDMS, 1 m^3/h	9 %		
DE (BG-UR)	Leckel SEQ 47/50, 2.3 m^3/h, Schleicher und Schuell glass fibre GF10	6 %	12%	11%
	Thermo FH 62 IR, 1 m^3/h.	24 %		
	Thermo FH 62 IR sharp, 1 m^3/h.	23 %		
	Digitel 30 m^3/h, Whatman QMA	3 %		
	Digitel 30 m^3/h, Whatman QMA	6 %		
FR (BG-UR)	Teom FDMS 1400 AB, 1 m^3/h	23 %	17%	12%
	Environnement MP 101M-RST, 1 m^3/h, dynamic heating	22 %		
	Partisol, 1 m^3/h, Pallflex PTFE membrane 2 µm	9 %		
	Teom FDMS Vers. C, 1 m^3/h	18 %		
	Teom FDMS Vers. C, 1 m^3/h	14 %		
IT (BG-UR)	Opsis SM 200, 1 m^3/h,	15 %	21%	28%
	Teom FDMS Vers. C, 1 m^3/h	31 %		
	Zambelli explorer plus, 1 m3/h, Pall PTFE	28 %		
	Tecora, 1 m3/h, Pall T6020	17 %		
	Tecora 2.3 m^3/h, Millipore AQFA quartz	13 %		

***** ***More info*:***

F. Lagler, C. Belis and A. Borowiak (2011): A Quality Assurance and Control Program for $PM_{2.5}$ and PM_{10} Measurements in European Air Quality Monitoring Networks, JRC Technical Reports, EUR 24851 EN, DOI 10.2788/31647.

	Data quality objective (Directive 2008/50/EC)	Limit values	Corresponding expanded uncertainty
PM_{10}	**25 %** (relative expanded uncertainty at 95 % confidence level)	**Yearly: 40 µg/m³** **Daily: 50 µg/m³**	**10 µg/m³** **12.5 µg/m³**



Results & Observations

Nugget trend analysis

- requires specific care because of inherent serial autocorrelation
- combined linear regression with a block-bootstrap-based approach
- variable block length following a geometric distribution with a mean value of 30 days (CF-Interval 1) and 365 days (CF-Interval 2)
- 95% CF intervals by applying a coverage factor of 1.96 to the empirical standard deviation of the bootstrapped slope estimates

Results & Observations

Nugget and Sill trend analysis

Country	Parameter	Median Value ($\mu g/m^3$)	Trend Slope ($\mu g/m^3$/year)	95% CF-Interval 1 ($\mu g/m^3$/year)	95% CF-Interval 2 ($\mu g/m^3$/year)
FR	Nugget (2s)	6.45	-0.73	± 0.24	± 0.31
DE	Nugget (2s)	6.99	-0.02	± 0.18	± 0.28
GB	Nugget (2s)	6.13	0.11	± 0.07	± 0.09
AT	Nugget (2s)	8.29	-0.09	± 0.41	± 0.47
IT	Nugget (2s)	12.67	-0.09	± 0.51	± 0.56
NL	Nugget (2s)	7.97	-0.11	± 0.30	± 0.41
FR	Sill (2s)	9.17	-0.28	± 0.44	± 0.64
DE	Sill (2s)	9.25	0.13	± 0.36	± 0.45
GB	Sill (2s)	7.01	-0.01	± 0.11	± 0.14
AT	Sill (2s)	11.31	0.11	± 1.04	± 1.03
IT	Sill (2s)	19.85	0.54	± 1.89	± 2.24
NL	Sill (2s)	8.29	0.00	± 0.45	± 0.35

(Trend is significant for the nugget parameter of FR and GB only.)

Results & Observations

Nugget and Sill Mann-Kendall trend test

Country	Parameter	Median Value ($\mu g/m^3$)	Kendall's tau	95% CF-Interval 1	95% CF-Interval 2
FR	Nugget (2s)	6.45	-0.285	± 0.111	± 0.194
DE	Nugget (2s)	6.99	-0.031	± 0.083	± 0.145
GB	Nugget (2s)	6.13	0.084	± 0.052	± 0.069
AT	Nugget (2s)	8.29	-0.035	± 0.103	± 0.110
IT	Nugget (2s)	12.67	-0.015	± 0.099	± 0.080
NL	Nugget (2s)	7.97	-0.027	± 0.073	± 0.061
FR	Sill (2s)	9.17	-0.098	± 0.121	± 0.192
DE	Sill (2s)	9.25	0.008	± 0.087	± 0.100
GB	Sill (2s)	7.01	0.014	± 0.043	± 0.052
AT	Sill (2s)	11.31	0.016	± 0.116	± 0.101
IT	Sill (2s)	19.85	0.005	± 0.139	± 0.122
NL	Sill (2s)	8.29	-0.018	± 0.065	± 0.046

(Trend is significant for the nugget parameter of FR and GB only.)

Nugget and Sill trend analysis

Country	Parameter	Median Value ($\mu g/m^3$)	Trend Slope ($\mu g/m^3$/year)	95% CF-Interval 1 ($\mu g/m^3$/year)	95% CF-Interval 2 ($\mu g/m^3$/year)
FR	Nugget (2s)	6.45	-0.73	± 0.24	± 0.31
DE	Nugget (2s)	6.99	-0.02	± 0.18	± 0.28
GB	Nugget (2s)	6.13	0.11	± 0.07	± 0.09
AT	Nugget (2s)	8.29	-0.09	± 0.41	± 0.47
IT	Nugget (2s)	12.67	-0.09	± 0.51	± 0.56
NL	Nugget (2s)	7.97	-0.11	± 0.30	± 0.41
FR	Sill (2s)	9.17	-0.28	± 0.44	± 0.64
DE	Sill (2s)	9.25	0.13	± 0.36	± 0.45
GB	Sill (2s)	7.01	-0.01	± 0.11	± 0.14
AT	Sill (2s)	11.31	0.11	± 1.04	± 1.03
IT	Sill (2s)	19.85	0.54	± 1.89	± 2.24
NL	Sill (2s)	8.29	0.00	± 0.45	± 0.35

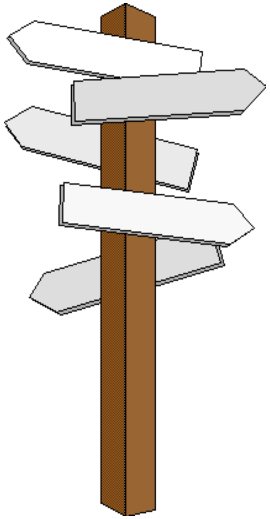
- A negative trend in the nugget time series can indicate:
  - improvement of the measurement uncertainty of the monitoring stations
  - reduction in small scale variability (change in the nature or quantity of emissions, transported pollution, atmospheric reactions)
- Other reasons causing either negative or positive trends:
  - increase / decrease of the number of monitoring stations
  - changes in station classifications

Conclusions

- Several variogram parameter time series are revealing a pronounced cyclic behaviour in the seasonal component S(t), and non-stationarities (change of variance over time) in the short-term component W(t).
- These effects clearly indicate the presence of temporal variations in the macroscale spatial correlation structures (sill).
- The strength of total seasonal variation of both the nugget and the sill values seems to be associated with the strength of topographic roughness and dissection of the country.
- A closer look at the nugget effect time series is of high practical interest, as it can provide information about the evolution of the mean measurement uncertainty of the related air pollutant.
- Further investigations are needed to determine if the trends of nugget variance are caused by a decrease / increase of the measurement uncertainty or by long term variations of air pollution and / or meteorological factors.

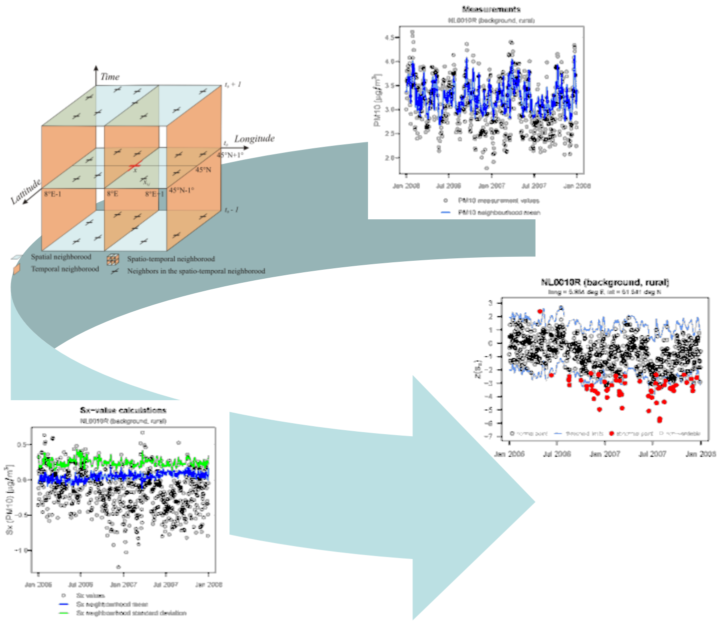
Outlook

- We consider that the investigation of variogram parameter time series and trend analysis deserves further attention. Future working steps might be:
  - Improving the robustness of the variogram model fit algorithms and implement automatic procedures able to screen the goodness of fit.
  - Processing new data compilations (AirBase v.8) in order to obtain more complete and longer time series.
  - Repeating the statistical analysis based on these updated data basis and geostatistical computations.
  - Linking these results with the spatial representativeness of monitoring stations.



European Commission

Thank you for your attention!



Joint Research Centre