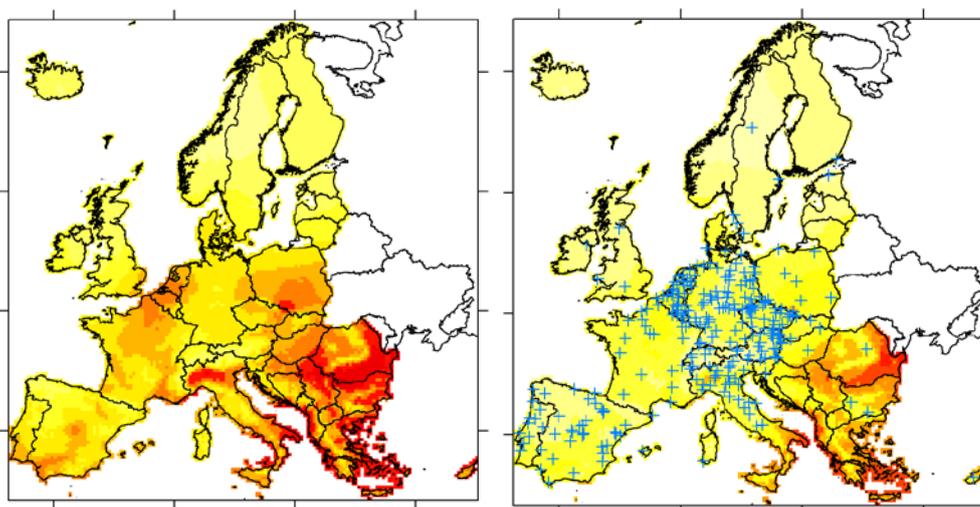


Spatio-temporal analysis and interpolation of PM10 measurements in Europe for 2009



ETC/ACM Technical Paper 2012/8
March 2013 – revised version

*Benedikt Gräler, Mirjam Rehr,
Lydia Gerharz, Edzer Pebesma*



The European Topic Centre on Air Pollution and Climate Change Mitigation (ETC/ACM) is a consortium of European institutes under contract of the European Environment Agency
RIVM UBA-V ÖKO AEAT EMISIA CHMI NILU VITO INERIS 4Sfera PBL CSIC

Front page picture:

Interpolated maps of yearly mean PM10 concentrations (left) in Europe for 2009 and the approximate annual standard deviation of the predictions calculated as by Denby et al. (2008b) using a temporal correlation coefficient of 0.3 (right). The method adopted is the metric covariance model using 100 neighbours (3-f in this paper). Blue crosses denote rural stations used to produce the maps. (Section 3.2 and Figure 5 of this paper).

Author affiliation:

Benedikt Gräler, Mirjam Rehr, Lydia Gerharz, Edzer Pebesma: Institute for Geoinformation (IfGI), University of Münster, Germany

DISCLAIMER

This ETC/ACM Technical Paper has not been subjected to European Environment Agency (EEA) member country review. It does not represent the formal views of the EEA.

© ETC/ACM, 2013.

ETC/ACM Technical Paper 2012/8 (revised version)

European Topic Centre on Air Pollution and Climate Change Mitigation

PO Box 1

3720 BA Bilthoven

The Netherlands

Phone +31 30 2748562

Fax +31 30 2744433

Email etcacm@rivm.nl

Website <http://acm.eionet.europa.eu/>

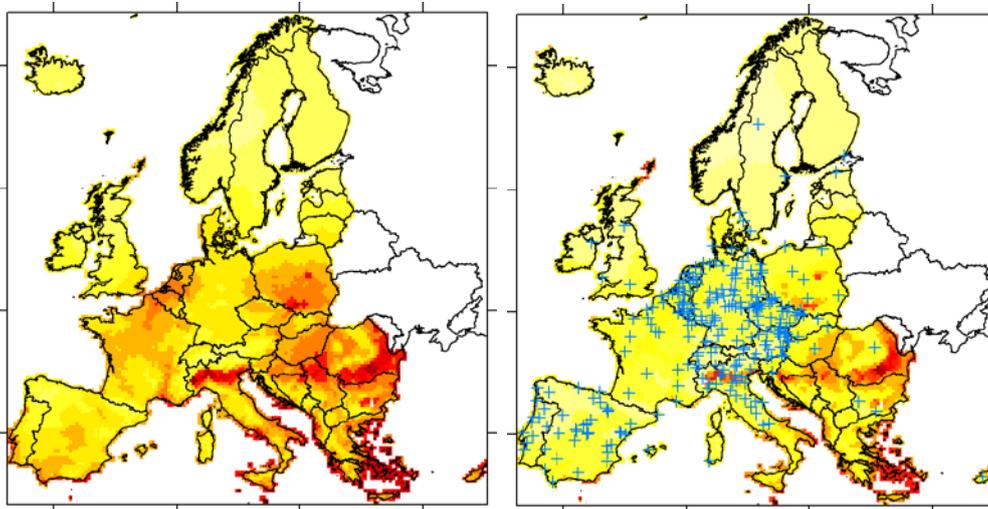
Spatio-temporal analysis and interpolation of PM10 measurements in Europe for 2009

Task 1022 - Subtask 2

ETC/ACM Technical Paper 2012/8 – revised version

Benedikt Gräler, Mirjam Rehr, Lydia Gerharz, Edzer Pebesma
ben.graeler@uni-muenster.de

March 2013



Content

Content	1
Revision note.....	2
Summary	2
1 Introduction	4
1.1 Aim and outline of the study	4
1.2 Spatio-temporal interpolation methods	4
1.3 Assessment methodology	5
1.4 Data	6
2 Details on spatio-temporal interpolation methods	7
2.1 Introduction, methods and data	7
2.2 Pre-processing	7
2.3 Time as a 3 rd dimension (f)	9
2.4 Separable covariance function (g).....	11
2.5 Product-sum covariance function (h)	12
2.6 Ordinary temporal interval kriging (i).....	12
2.7 Universal temporal interval kriging (j,k).....	13
2.8 Post-processing for residual kriging.....	13
3 Results	14
3.1 Statistical measures	14
3.2 Maps	17
4 Detecting asymmetric temporal dependencies	18
5 Discussion	21
5.1 How to select the temporal scale?	21
5.2 How robust are these methods?.....	21
5.3 Is temporal interval kriging suited for further analysis?	21
5.4 Do we need to consider asymmetric patterns?	23
6 Conclusions	24
References	26

Revision note

This is a revised version of the technical report published in January 2013 due to an error in the preparation of the EMEP model data used in some of the methods. Most figures and tables haven been updated. Corrections in Section 3 motivated further changes in Section 5.2, the conclusions and the summary.

Summary

This study continues the work described in Gräler et al. (2012a) and investigates the potential of additional approaches for spatio-temporal kriging of daily mean PM₁₀ concentrations as well as a re-assessment of the prediction quality for another year, i.e. 2009 next to 2005. All methods are applied to daily mean rural background PM₁₀ concentrations across Europe for the year 2009. Some of them incorporate daily EMEP model data and elevation data as supplementary data. Additionally, temporal interval¹ kriging approaches are investigated to assess their potential for analysing trends in yearly mean concentrations. Asymmetric spatio-temporal dependencies are also addressed.

The air quality indicators used in this study are daily and yearly mean PM₁₀ concentrations. Cross validation is used to assess the quality of the different techniques. Statistical measures are used to quantify the improvement for different indicators. Earlier studies (Denby et al. 2008a, Gräler et al. 2012a) showed that daily and spatio-temporal interpolations improve the statistical performance. In this study, we assess the robustness of these spatio-temporal interpolation methods by applying the metric, the separable and the product-sum model to 2009 data. Temporal interval kriging approaches based on a metric covariance model, where the interval width is a year, are evaluated for their capability to capture spatio-temporal dependence and to provide valid estimates of the prediction error variability.

The cross-validation results of the methods that are re-assessed in this study support the conclusion that within the spatio-temporal approaches the metric residual kriging performs overall best. This is as well supported by the statistical indicators derived from the 2005 data in Gräler et al. (2012a). However, the spatio-temporal metric covariance model does not outperform the pure spatial model in predicting yearly averages. On a daily level, the product-sum covariance model produces the best estimates applied to residuals of a multiple linear regression. The overall best methods on a daily level are the metric approaches applied to the raw data. However, an informed choice has to be made for the temporal scaling parameter that relates temporal with spatial distances in the metric covariance model approximating an isotropic space. Temporal interval kriging performed worse than the alternative approaches. However, the comparison of the predicting variances reveals considerable differences in magnitude and spatial variation. This fact stresses the importance to derive a valid uncertainty

¹ We perform block kriging in the isotropic 3-dimensional spatio-temporal (x,y,t) space, where the blocks are chosen to be lines along the time axis. To stress that no spatial averaging takes place, we refer to this form of block kriging as *temporal interval kriging*.

measure. Using proxies to the associated uncertainty of annual mean predictions might lead to false interpretations of temporal trends. The study on asymmetric dependence patterns due to prevailing environmental impacts (e.g. wind) did not reveal a meaningful asymmetric pattern for the residuals. The raw data seem to be weakly asymmetric along the North-South axis.

Even though the spatio-temporal interpolation methods are only tested on two years (2005 and 2009), we believe that the spatio-temporal interpolation methods provide great potential for the production of PM_{10} concentration maps over Europe on a daily basis. The requirement of having daily EMEP model data may, however, limit the applicability of the approaches based on a multiple linear regression. Nevertheless, the metric approach applied to the raw observations performs best in 2009 and reasonable in 2005 on a daily level. The temporal interval kriging approaches did not perform very well in the cross-validation study, but their ability to provide correct prediction error estimates is very valuable in the uncertainty analysis. This becomes especially important when long term European-wide trends are of interest. Further research is needed to validate and re-adjust this approach. No strong asymmetry due to a prevailing direction could be discovered in the data set. However, this is only one type of asymmetry that can occur in spatio-temporal data. Gräler and Pebesma (2012b) showed that asymmetric copulas, capturing non-Gaussian dependence structures, best fitted the data for most non-zero temporal lags. These copulas are capable of mimicking dependencies that occur in processes with e.g. rapid decrease, but slow increase as present in air quality data.

1 Introduction

The increasing number of sensors and measurement stations with a finer temporal resolution call for sophisticated spatio-temporal interpolation methods. The first part of this study described in Gräler et al. (2012a) revealed significant temporal and spatio-temporal dependencies. A purely spatial interpolation approach ignores these dependencies and can be seen as a spatio-temporal interpolation where all temporal correlations are set to zero. Obviously, investigating this additional source of information has potential to improve the mapping of PM_{10} and spatio-temporal processes in general. This report revisits the methods with the best performance on the 2005 data, for the year 2009 to assess whether the performance improvement persists for 2009, and investigates further spatio-temporal kriging techniques to address asymmetric spatio-temporal dependencies as well as temporal interval kriging.

1.1 Aim and outline of the study

This study was conducted to further explore the potential of different spatio-temporal kriging methods for daily PM_{10} measurements across Europe and to check the robustness of methods developed in an earlier study. Therefore, a set of different approaches has been implemented in R (R Development Core Team, 2012). Daily means from AirBase (EEA: AirBase version 6) across Europe for the year 2009 serve as a sample data set. A cross validation study has been conducted and a set of different quality measures has been derived for every method. The cross validations have been exploited at two temporal resolutions:

- daily, comparing the predicted daily mean concentrations, and
- yearly, looking at the station-wise annual mean of daily predictions.

1.2 Spatio-temporal interpolation methods

Several spatio-temporal interpolation methods have been applied. The approaches are set-up following the notation and numbering introduced in Gräler et al. (2012a) extending work presented by Denby et al. (2008a), to ease the comparison between the ever growing set of new methods. Additionally, we performed the routinely applied interpolation procedure for annual means for the 2009 data to provide a benchmark in the cross-validation study.

The set of evaluated spatio-temporal kriging methods is applied to differently pre-processed daily mean PM_{10} concentrations as provided by AirBase. We consider the following two where the numbers match those in Gräler et al. (2012a):

1. the unaltered daily means (referred to as raw data)
3. residuals after multiple linear regression of the logs of observed daily mean concentrations with altitude, EMEP model daily data and a periodic temporal predictor

The kriging methods we revisit in this study include the three fully spatio-temporal methods (the numbering matches the previous report):

- f. a metric covariance function (see Section 2.3)
- g. a separable covariance function (see Section 2.4)
- h. a product-sum covariance function (see Section 2.5)

The separable covariance function assumes the spatial and temporal correlation processes to be disjoint for space and time where the product-sum introduces some interaction. The metric covariance function assumes time to be a third orthogonal dimension with identical correlation structure for spatial, temporal and spatio-temporal distances. Therefore, a rescaling of the temporal domain is necessary to reduce anisotropy in the metric kriging approach. A comparison of different heuristics to determine this scaling is presented as well.

In the last step of the interpolation, the regression predictions and their residuals are merged and back-transformed. The desired yearly averages are calculated per interpolated grid cell.

Additionally, we investigated the potential of temporal interval kriging to obtain better estimates of the kriging variance of the yearly mean concentrations. The original observations were used here, as the temporal interval kriging variance cannot be back-transformed easily if the data had been transformed before. Altitude and EMEP model data have been used as covariates in universal temporal interval kriging. The approaches are denoted by:

- i. ordinary temporal interval kriging (Section 2.6)
- j. universal temporal interval kriging with altitude (Section 2.7)
- k. universal temporal interval kriging with altitude and EMEP model data (Section 2.7)

1.3 Assessment methodology

The quality of all methods is assessed using statistical measures and applicability concerns in terms of processing time. The comparison of different interpolation methods takes place on a yearly (comparing yearly means for predictions and measurements) and daily (comparing daily means for predictions and measurements) level. The root mean squared error (RMSE), the bias (BIAS), mean absolute error (MAE) and Pearson's correlation coefficient (COR) between predictions and measurements at every location are calculated using leave-one-out cross validation. All methods are compared using the above measures where applicable. All methods are implemented in R (R Development Core Team, 2012; R 2.15.0) and use the R packages *spacetime* (Pebesma, 2012; package version 1-0.5) and *gstat* (Pebesma, 2004; package version 1.0-17).

1.4 Data

The daily mean PM₁₀ concentration test data are obtained from EEA's AirBase (version 6) for all European countries. To assess the robustness of the revisited approaches all methods are applied to daily mean values in 2009 from rural background stations. As in the yearly interpolation approach in Horálek et al. (2007), we used all measurement stations providing a data integrity of at least 75% throughout the year enabling a comparison with the routinely applied procedures. Thus, the number of measurement stations amounts to 289. However, when considering the complete space-time lattice (i.e. every space and time combination is available) as required by the separable and product-sum covariance functions, the data set yields 4931 missing values (4.7%). The metric covariance function can easily correct for these missing values and could in general be applied to data from stations with less than 75% coverage over time. Details will be given in the following chapter.

Daily mean PM₁₀ concentrations calculated by means of the EMEP model (rv3_8_1 EMEP/MS-CW model v.2011-06 according to the newly implemented nomenclature in the EMEP Status Reports (Fagerli et al. 2011)) using meteorological data and emission data for the year 2009 serve as additional data in the multiple linear regression. The 50 km EMEP Polar Stereographic grid is converted and projected to match the 10 km gridded interpolation domain in the standard EEA ETRS89-LAEA5210 projection. Additionally, altitude data is used from both the AirBase measurement stations and from the EEA 10 km grid originating from the 30 arcsec GTOPO altitude grid (WGS84) as applied in Horálek et al. (2007).

2 Details on spatio-temporal interpolation methods

2.1 Introduction, methods and data

We consider a range of different spatio-temporal kriging methods. The revisited methods were described in Gräler et al. (2012a) and are referred to with f , g and h while the additional methods continue this enumeration with i , j and k and are described in Section 2.6 and Section 2.7.

The interpolation methods consist of using:

- f. metric covariance function based on the complete 3D space with either 10, 100 or 1000 nearest neighbours (Section 2.3)
- g. separable covariance function (Section 2.4)
- h. product-sum covariance function (Section 2.5)
- i. ordinary temporal interval kriging (Section 2.6)
- j. universal temporal interval kriging with altitude (Section 2.7)
- k. universal temporal interval kriging with altitude and EMEP model data (Section 2.7)

The underlying data are:

- observed daily mean PM₁₀ concentrations (denoted as 1)
- simple log transforms of daily mean PM₁₀ concentrations (denoted as 1.1)
- residuals after multiple linear regression of the log-transformed concentrations (denoted as 3)

The spherical and exponential variogram models repeatedly appear to be best suited to capture the spatial and temporal dependencies and are therefore adopted in all methods. Their parameters are fitted using the functions provided by the R package *gstat* (Pebesma, 2004). The R scripts and data sets (up to the extent that they are free) that were used are available from the authors upon request.

To support the multi-linear regressions the PM₁₀ concentrations and their log-transforms undergo some pre-processing (Section 2.2). When adding the interpolated residuals again to the estimates produced by the multi linear regression some post-processing (Section 2.8) is needed to compensate for bias in the case when using log-transformed data.

2.2 Pre-processing

A multiple linear regression is applied to the daily mean PM₁₀ concentrations from the AirBase database and their log-transforms. The predictors include:

- the daily EMEP grid cell's prediction value for every station
- a temporal index to adjust for periodic yearly changes: $\cos\left(\frac{\text{day}}{365} \cdot 2\pi\right)$
- station altitude

In contrast to the yearly interpolation approach for rural areas (Horálek et al., 2007), meteorological data (wind speed, solar radiation) are not used here as predictors. The selection of predictors and their interactions is based on a forward stepwise approach and the model with the strongest explanatory value is selected.

Variations in the combination of the chosen predictors are also tested. This set-up is used for multiple linear regression of the log transformed daily mean PM₁₀ concentrations. An overview of the estimated regression coefficients for the different parameters can be found in Table 1 below.

Table 1: Multiple linear regression coefficient estimates and standard errors for log transformed daily mean PM₁₀ concentrations in 2009 and 2005. The adjusted R² is given in square brackets. The 2005 values are recalculated from Gräler et al. (2012a) after the erratum has been published. The last two rows indicate the parameters for the interactions of EMEP model values and altitude with time.

multi linear regression coefficients for log(PM ₁₀) ...				
	... in 2009 [0.31]		... in 2005 [0.29]	
	Estimate	Std. Error	Estimate	Std. Error
Intercept	2.46117	0.0040	2.67177	0.0047
EMEP model	0.05331	0.0003	0.03293	0.0004
Time	0.03617	0.0055	-/-	-/-
Altitude	-0.00037	0.0053	-0.00060	0.0000
time:altitude	-0.00037	0.0000	-/-	-/-
Model:time	0.00250	0.0005	-/-	-/-
model:altitude:time	-/-	-/-	-0.00005	0.0000
model:altitude	-/-	-/-	0.00002	0.0000

Figure 1 shows the normal quantile-quantile-plot for the linear regression residuals of the model in 2009. This plot is well suited to visually judge the underlying assumption of normal distributed residuals. Perfectly normally distributed residuals would follow the dotted diagonal line. The residuals of the log-transformed data do seem to follow a normal distribution reasonably well, despite the lower tail. Therefore most of the results presented in this report refer to the log-transformed multiple linear regression.

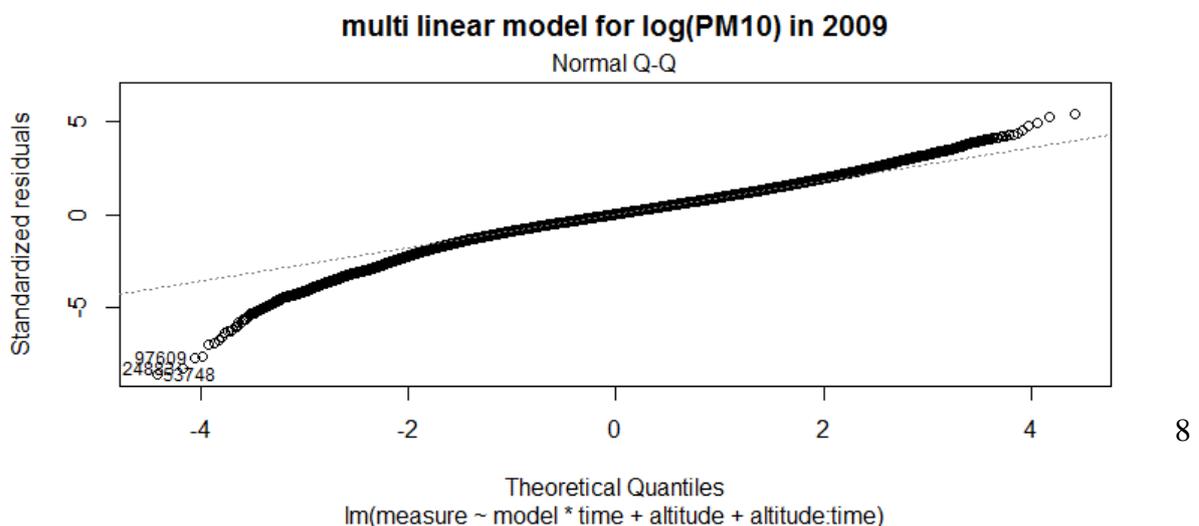


Figure 1: Diagnostic plot showing the normal quantile-quantile-plot for the 2009 regression model.

2.3 Time as a 3rd dimension (f)

Considering time to be the third orthogonal dimension leads to a natural three-dimensional extension of two-dimensional kriging. In order to allow for meaningful results using an isotropic covariance model, the temporal dimension has to be rescaled to align with the spatial directions. The choice of this parameter will influence the strength of the spatio-temporal covariance. We considered three different heuristics to determine this scaling:

1. Estimating pure spatial and pure temporal pooled (see Gräler et al. (2012a), Section 2.4) variograms yielded ranges of approximately 675 km and 4.1 days respectively. The fraction of both indicates a scaling of 164 km/day.
2. Optimising the variogram model fit. Different scales were applied and the RMSE between model and empirical variogram were calculated. The lowest RMSE was achieved for 60 km/day.
3. Selecting a set of scales and performing the full metric covariance kriging in a cross validation. The cross-validation results for the raw data and for the residuals are given in Table 2 and Table 3 respectively. The row denoted with ‘ID’ refers to numbers and letters assigned to the introduced methods in Section 2.1. The best cross-validation results were achieved for 180 km/day up to 360 km/day.

Table 2: Cross-validation results for different temporal scaling factors for the raw daily PM₁₀ concentrations and a metric covariance model. Green cells highlight the best value in every row, yellow cells are considered to perform almost as good as the best method. The temporal scale is given in the form day ~ km. The numbers in square brackets indicate the spatio-temporal neighbourhood size.

DAILY	raw PM10								
2009	metric [100] 1~60	metric [100] 1~120	metric [100] 1~164	metric [100] 1~180	metric [100] 1~240	metric [100] 1~300	metric [100] 1~360	metric [100] 1~420	metric [100] 1~480
ID	1-f	1-f	1-f	1-f	1-f	1-f	1-f	1-f	1-f
RMSE	9.55	9.30	9.35	9.24	9.24	9.23	9.23	9.41	9.42
BIAS	-0.22	-0.12	-0.12	-0.11	-0.11	-0.12	-0.12	-0.50	-0.49
MAE	5.86	5.63	5.67	5.57	5.56	5.56	5.55	5.78	5.79
COR	0.75	0.76	0.76	0.76	0.77	0.77	0.77	0.76	0.75
YEARLY									
RMSE	5.37	5.21	5.20	5.16	5.15	5.14	5.13	5.18	5.19
BIAS	-0.23	-0.14	-0.13	-0.12	-0.13	-0.13	-0.13	-0.52	-0.51
MAE	3.80	3.72	3.72	3.69	3.69	3.69	3.69	3.71	3.71
COR	0.63	0.63	0.63	0.64	0.64	0.64	0.64	0.64	0.64

Throughout the study we compare the two scales 164 km/day and 360 km/day. The first heuristic is rather straightforward to apply and the corresponding cross-validation results are compatible with the third criterion while the latter provides the overall best cross-validation results. With this rescaling, for any prediction location/time, a set of nearest observations forms automatically the set of the most correlated neighbours in space and time. This method relies on a single variogram with the variogram parameters for the different scales and data sets given in Table 4. A graphical representation of all four metric variograms is given in

Figure 2. These plots reveal some periodic pattern in the empirical variogram for the larger scale of 360 km/day letting the choice of 164 km/day appear more natural.

Table 3: Cross-validation results for different temporal scaling parameters for the residuals and a metric covariance model. Green cells highlight the best value in every row, yellow cells are considered to perform almost as good as the best method. The temporal scale is given in the form day ~ km. The numbers in square brackets indicate the spatio-temporal neighbourhood size.

DAILY	log – linear model – residual kriging								
2009	metric [100] 1~60	metric [100] 1~120	metric [100] 1~164	metric [100] 1~180	metric [100] 1~240	metric [100] 1~300	metric [100] 1~360	metric [100] 1~420	metric [100] 1~480
ID	3-f	3-f	3-f	3-f	3-f	3-f	3-f	3-f	3-f
RMSE	12.14	11.70	11.65	11.70	11.64	11.62	11.68	11.68	11.21
BIAS	-0.77	-0.64	-0.61	-0.64	-0.60	-0.60	-0.59	-0.60	-0.67
MAE	6.04	5.74	5.68	5.74	5.66	5.65	5.65	5.64	5.61
COR	0.68	0.69	0.70	0.69	0.69	0.69	0.69	0.69	0.71
YEARLY									
RMSE	4.69	4.58	4.54	4.58	4.53	4.50	4.49	4.48	4.48
BIAS	-0.80	-0.66	-0.63	-0.66	-0.62	-0.63	-0.62	-0.63	-0.70
MAE	3.47	3.39	3.35	3.39	3.34	3.33	3.32	3.32	3.34
COR	0.75	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76

Predictions can be made for any point in a space-time cube. The interpolation is no longer restricted to single time slices and every location is estimated based on its 100 nearest neighbours in the isotropic space-time cube. Depending on the temporal scale and local configuration of stations, the number of distinct stations will vary. However, the assumption of a metric covariance model neglects the distinction between a spatial and temporal domain. The nearest neighbours in the space-time cube are thought to have the strongest correlation with the predicted value. Additionally, 1000 and only 10 neighbours were selected to study the influence of the neighbourhood size. Increasing the neighbourhood ten-fold showed only little improvement while using only one tenth of neighbours performed surprisingly well. In this approach, missing values in a single time slice are made up for by using values from the preceding and following time slices. Thus, all the available data across space and time is incorporated. In general, this approach is not restricted to stations of at least 75% temporal coverage, and better results might be expected when all available data would be used.

Table 4: Set of variogram parameters used within this study for the metric covariance model.

	raw data		residuals	
	164 km/day	360 km/day	164 km/day	360 km/day
nugget	64.56	66.25	0.11	0.12
partial sill	120.18	109.06	0.18	0.17
range [km]	798	856	577	660

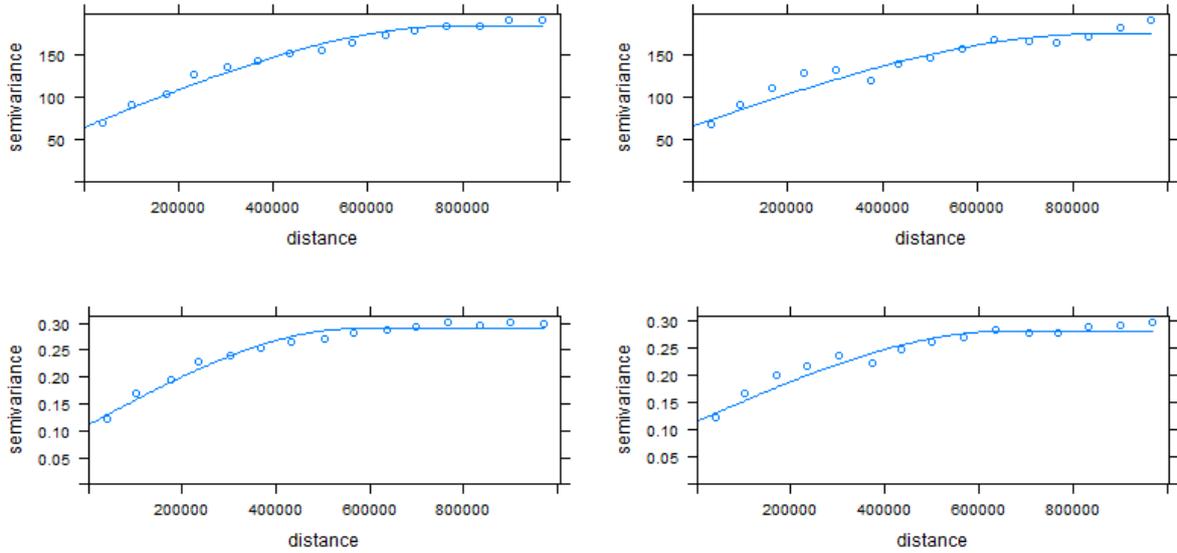


Figure 2: Variogram of the metric kriging approach of the residuals of the raw (first row) and log-transformed (second row) PM_{10} concentrations. Distance corresponds to meters, where a temporal difference of 1 day equals 164 km (left column) and 360 km (right column).

2.4 Separable covariance function (g)

Extending the kriging approach from space to space-time alters the covariance function. In space and under the assumptions of isotropy and stationarity, the covariance is a function $C(h)$ of the separating distance h between two locations. A spatio-temporal covariance function is thought of as a function of a spatial and a temporal distance $C(h, t)$. In general, this function is hard to estimate. A separable covariance function is assumed to fulfil $C(h, t) = C_S(h) \cdot C_T(t)$ and the distinct spatial and temporal effects can be estimated with relative ease. The interpolation is performed based on the complete spatio-temporal data set provided for the variogram estimation. This allows the prediction of values at any location and point in time.

A drawback of this approach potentially affecting the predictions occurs when one has to deal with incomplete data sets. The estimation of the kriging weights relies on identical and complete time series for every spatial location. In order to adopt for a large number of missing values in the AirBase data set, a monthly stepwise spatio-temporal interpolation is performed. Hence, a complete space-time lattice needs only to be present in each month and not for the entire year. The interpolation is then carried out for every day in each month. However, the spatial and temporal variograms used in the monthly interpolation steps are fitted once for the full set of observations only. The spatial variogram is the pooled variogram (see Gräler et al. (2012a), Section 2.4) over all spatial data (Spherical model: nugget: 0.09, partial sill: 0.24, range: 675 km). The temporal variogram is also the pooled variogram but only takes the temporal lags into account (Exponential model: nugget: 0, partial sill: 0.28, range: 1.87 days).

2.5 Product-sum covariance function (h)

A powerful extension to the above is the introduction of the product-sum covariance function. It relies on the assumption that the spatio-temporal covariance function $C(h, t)$ can be written as $C(h, t) = k_1 C_S(h) \cdot C_T(t) + k_2 C_S(h) + k_3 C_T(t)$ where $k_1 > 0$ and $k_2, k_3 \geq 0$ are real coefficients. We follow an estimation approach introduced by De Iaco (2001) where the coefficients k_1 , k_2 and k_3 depend on the spatial, temporal and spatio-temporal sills. Figure 3 shows the empirical and the fitted spatio-temporal variogram model. The extended covariance model gains flexibility in modelling the spatio-temporal process. This approach has a similar technical drawback as the one before: the estimation can only be based on complete spatio-temporal datasets. Thus, we follow the same month-wise interpolation workaround using disjoint temporal neighbourhoods which might as well affect the predictions.

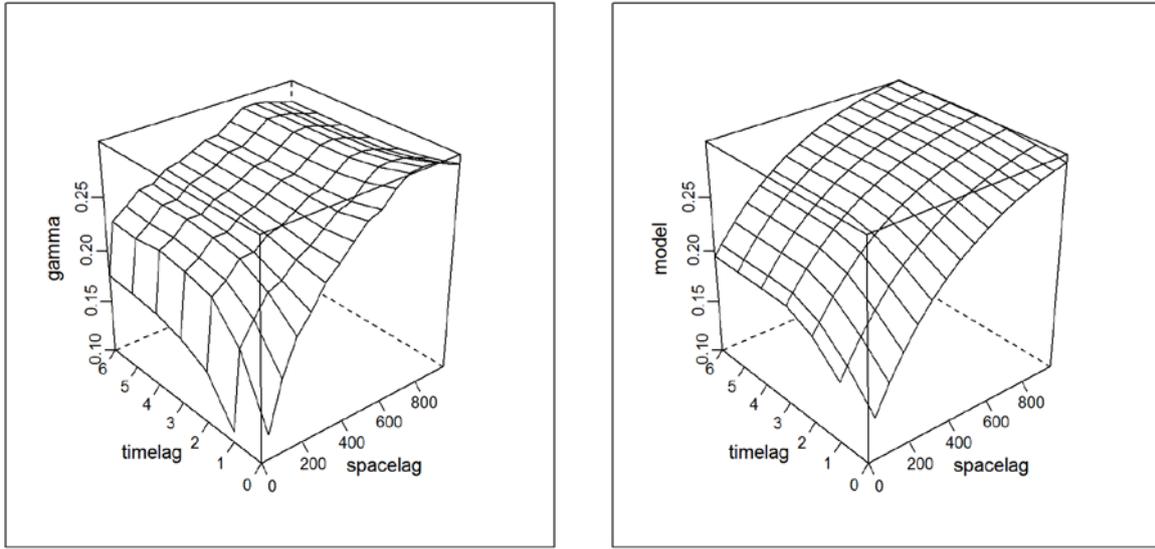


Figure 3: Empirical (gamma) and fitted spatio-temporal variogram (model) of the residuals of log-transformed PM_{10} concentrations used in the separable covariance model. The “timelag” corresponds to days where the “spacelag” denotes km.

2.6 Ordinary temporal interval kriging (i)

The procedures mentioned so far lack the ability to provide an uncertainty measure for the yearly predictions that are calculated as the mean of the daily kriging estimates. A method that allows the calculation of the kriging variance for a prediction over a larger support (i.e. areas instead of points) is block kriging (Cressie and Wikle, 2011). Block kriging is used to derive estimates for spatial areas (grid cells or polygons) from point data, where “usual” point kriging produces estimates for points and not for areas. In our setting, we define blocks as pillars in the 3-dimensional isotropic spatio-temporal space where their height corresponds to one year. However, monthly mean estimates could be derived analogously. Although the method is mathematically equivalent with block kriging, to avoid the suggestion that we produce *spatial (block) averages* we refer to this method as *temporal interval kriging*. Due to the enormous number of observations involved in this approach, we had to limit the spatial extent to the closest ten neighbours. All available temporal observations of these neighbours are used to derive a prediction for the temporal interval. Unfortunately, the property of providing a direct kriging variance for the block support does not hold for log transformed

data. Thus, we only applied this approach to the raw daily mean concentrations using the raw data variograms from Table 4.

2.7 Universal temporal interval kriging (j,k)

Universal temporal interval kriging extends ordinary temporal interval kriging assuming a linear trend for the mean based on provided variables. We provided the altitude alone and together with the EMEP model data to evaluate these approaches. The temporal predictor through the cosine function is not used in this case, as its influence in the raw data regression is not significant. These procedures are also only applied to raw daily mean PM₁₀ concentrations. Table 5 lists the variogram parameters used in the universal temporal interval kriging approaches for an exponential variogram model.

Table 5: Set of variogram parameters used for different universal temporal interval kriging approaches. The temporal scale is indicated as day ~ km.

	metric altitude 1~164	metric altitude 1~360	metric alti+EMEP 1~164	metric alti+EMEP 1~360
Nugget	51.45	55.11	54.40	58.01
Partial sill	148.07	138.18	141.03	131.69
Range parameter [km]	487	570	194	588

2.8 Post-processing for residual kriging

After the residuals have been interpolated, they are added again to the estimates produced by the multiple linear regression. In general, the kriging predictor is asymptotically unbiased and will give the correct estimate. In our case, the estimates are correct for the log-transformed concentrations. The non-linear back transformation using \exp alters the weights of the linear prediction on the non-log scale resulting in biased prediction of the mean (unbiased for the median). Predictions made on log-transformed data from the pre-processing are back-transformed using $Y(s) = \exp(Z(s) + \sigma(s)^2/2)$ including the correction factor to lower the effect of the unbiased estimator (Denby et al., 2008b). Here, $Z(s)$ denotes the estimate at some location s combining the multiple linear regression model with the kriged residuals. The kriging variances are back transformed by $V(s) := (\exp(\sigma^2(s)) - 1) \cdot \exp(2 Z(s) + \sigma^2(s))$ (Equation 2, Denby et al., 2008b). The annual PM₁₀ averages are given by the means per grid cell of all daily interpolated values and the annual variance is roughly approximated by the formula given by Denby et al. (2008b, Appendix A). Section 5.3 illustrates the potential of temporal interval kriging to provide correct estimates of the annual prediction variance.

3 Results

The overall picture of the cross-validation results follows the observations made in Gräler et al. (2012a) for the data set from 2005. Note that only the best performing spatio-temporal methods are revisited and this report only makes a comparison between these. The metric covariance model, as the best performing method, reduces the RMSE of the annual means by 5% compared to the routinely applied annual mean predictions.

3.1 Statistical measures

The sequential cross-validation leaving each station out once and predicting the measurements from the remaining ones is the key to compare the different interpolation approaches. For every cross-validation a set of statistical measures is calculated: the root mean square error (RMSE) providing the standard deviation of the errors, the bias (BIAS) and mean absolute error (MAE) indicating the mean and mean absolute deviation of the prediction. Furthermore, Pearson's correlation coefficient (COR) is calculated. The full set of all cross validation results is given in Table 6 comparing the listed methods on a daily and yearly basis. The row denoted with 'ID' refers to numbers and letters assigned to the introduced methods in Section 2.1. Table 6 provides the results for:

- kriging on raw data using a metric variogram model
- kriging on log-transformed raw data using a metric variogram model
- different kriging approaches using residuals of the multiple linear regression of log-transformed PM_{10} concentrations (including the routinely applied annual interpolation)
- different temporal interval kriging approaches for temporal scales of 164 km/day and 360 km/day

Each interpolation approach is a composition of the pre-processing strategy (indicated by 1, 1.1 or 3, see Section 2.2) and the kriging method (indicated by f – k, see Section 2.3 – Section 2.7) given in the blue header cells as 'ID'. Thus, some kriging techniques appear several times in Table 6, but under different pre-processing strategies. Every column defines a unique interpolation approach. The number of neighbours used in the interpolation is given in square brackets. The different temporal scales for the metric variogram are denoted in the form of day ~ km. In this table, we only used 164 km/day and 360 km/day. The full bandwidth of cross-validation results on the temporal scale are shown in Table 2 and Table 3 in Section 2.3.

The two direct kriging methods (first two columns) do not use residuals of a multi linear interpolation and are applied to the raw measurements (*1-f*); the following two columns use log-transformed data (*1.1-f*). The third set of columns describes our kriging approaches for the residuals obtained from a multiple linear regression model applied to the log-transformed data (*3-f*, *3-g*, *3-h*). The last two sets of columns present the results for the temporal interval kriging approaches (*3-i*, *3-j*, *3-k*). Naturally, comparisons of temporal interval kriging predictions with daily observations cannot be made.

Table 6: Statistics for the cross-validation results for the major spatio-temporal approaches. Green cells highlight the best value in every row, yellow cells are considered to perform almost as good as the best method. IDs given green letters denote methods newly introduced in this report. (NA = not applicable). The temporal scale is given in the form day ~ km. The numbers in square brackets indicate the spatio-temporal neighbourhood size. The temporal interval kriging approaches have only a spatially restricted neighbourhood and rely on the 10 spatially nearest neighbours and all the temporal data points in this region ([sp 10]).

DAILY	raw PM10		log(PM10)		log – linear model – residual kriging						temporal interval kriging 1~164			temporal interval kriging 1~360			
2009	metric [100] 1~164	metric [100] 1~360	log metric [100] 1~164	log metric [100] 1~360	annual means	metric [10] 1~164	metric [1000] 1~164	metric [100] 1~164	metric [100] 1~360	sep. COV	p-s. COV	metric ordinary [sp10]	metric alti [sp10]	metric alti+EMEP [sp10]	metric ordinary [sp10]	metric alti [sp10]	metric alti+EMEP [sp10]
ID	1-f	1-f	1.1-f	1.1-f	none	3-f	3-f	3-f	3-f	3-g	3-h	3-i	3-j	3-k	3-i	3-j	3-k
RMSE	9.35	9.23	9.59	9.60	NA	11.57	11.78	11.65	11.68	12.14	11.41	NA	NA	NA	NA	NA	NA
BIAS	-0.12	-0.12	-0.35	-0.31	NA	-0.79	-0.57	-0.61	-0.59	-0.48	0.05	NA	NA	NA	NA	NA	NA
MAE	5.67	5.55	5.79	5.76	NA	5.69	5.74	5.68	5.65	5.79	5.60	NA	NA	NA	NA	NA	NA
COR	0.76	0.77	0.74	0.74	NA	0.70	0.69	0.70	0.69	0.66	0.70	NA	NA	NA	NA	NA	NA
YEARLY																	
RMSE	5.20	5.13	5.43	5.39	4.34	4.62	4.56	4.54	4.49	5.02	4.84	6.12	5.49	5.17	6.08	5.30	5.06
BIAS	-0.13	-0.13	-0.34	-0.30	0.11	-0.82	-0.60	-0.63	-0.62	-0.99	-0.34	-0.53	-0.30	-0.55	-0.48	-0.31	-0.58
MAE	3.72	3.69	3.93	3.87	3.23	3.41	3.39	3.35	3.32	3.76	3.50	4.57	4.04	3.87	4.53	3.88	3.81
COR	0.63	0.64	0.60	0.59	0.76	0.76	0.76	0.76	0.76	0.70	0.72	0.46	0.64	0.67	0.46	0.65	0.68

Comparing the values of the repeatedly applied methods on the 2009 and 2005 data, all cross-validation statistics are in same order of magnitude. On the daily level, approaches perform typically better for 2005 than for 2009 while it is the opposite case for the yearly level. Comparing the results on a yearly level, the metric kriging approaches using residuals of a multiple linear regression on log-transformed measurements constantly outperform all the other spatio-temporal ones. Even limiting the number of neighbours to 10 does not distort the predictions meaningfully. Again as in Gräler et al. (2012a), the increased number of 1000 neighbours does not gain any valuable improvement compared to 100 neighbours but drastically increases the computational efforts. Nevertheless, no improvement can be shown for the spatio-temporal covariance models compared to a pure spatial interpolation of annual

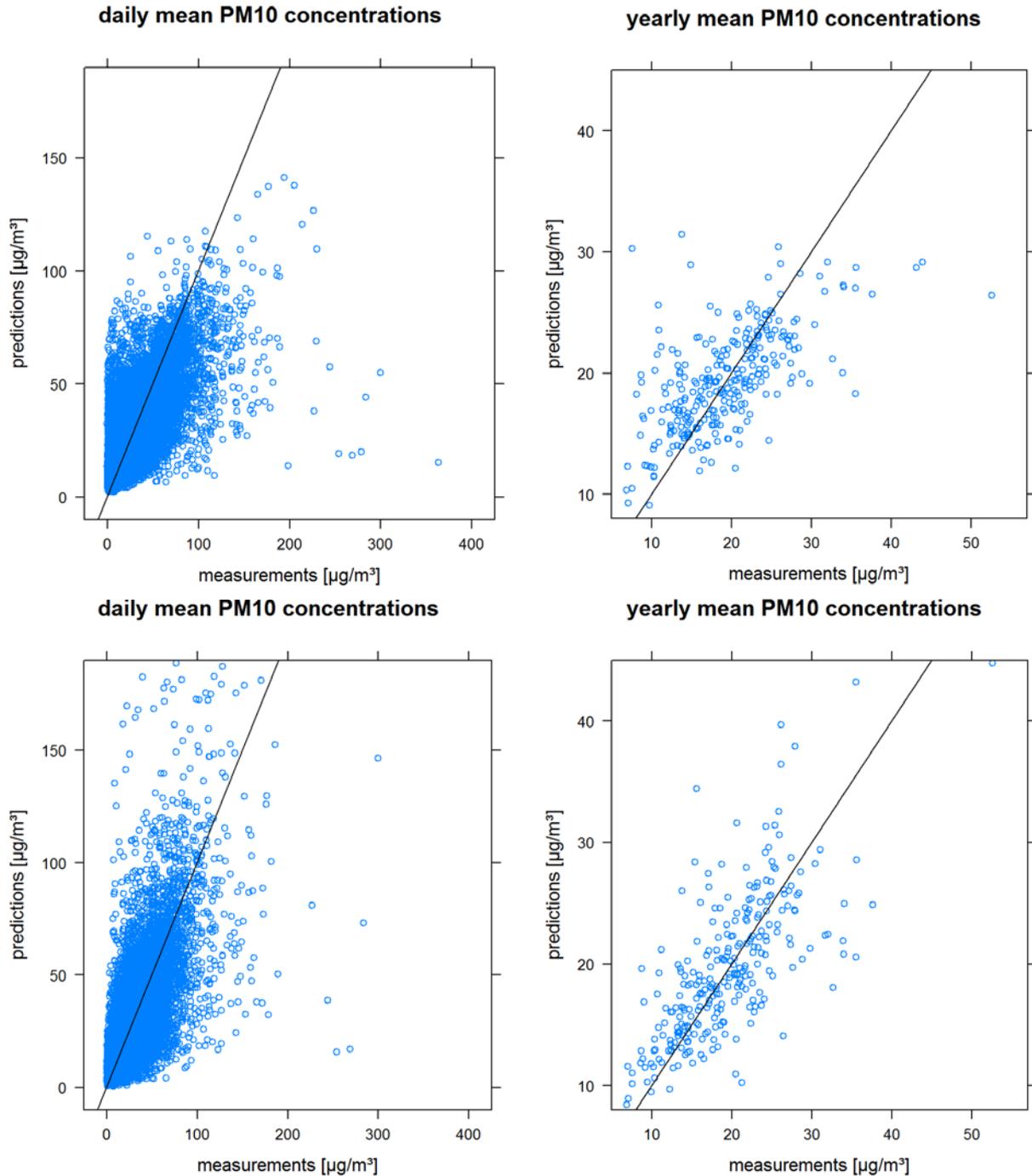


Figure 4: Cross validation scatterplots on a daily and yearly level for the raw metric (1-f) with 1 day ~ 360 km (upper row) and the product-sum model (lower row). The straight lines indicate the main diagonal, the perfect match.

means. A different picture appears on a daily level where in this study the metric model on the raw data (1-f) outperforms any other approach in most values of the cross-validation. In accordance with the study on 2005 data, the product-sum covariance model (3-h) produces better estimates than the simpler separable covariance model (3-h).

Cross validation scatter plots on a daily and yearly level are given in Figure 4. They are based on the raw PM_{10} daily mean values and metric kriging over 100 neighbours and temporal scaling of 360 km/day for the upper row and the product-sum model applied to residuals after the multiple linear regression for the lower row. The product-sum model shows a reasonable fit on the yearly level where the daily cross-validations still have pretty large deviations for both approaches.

3.2 Maps

We chose the overall best performing spatio-temporal procedure according to the statistical measures on a yearly and daily level (metric covariance model with 100 neighbours and a temporal scale of 360 km/day, 3-f) to produce a map of yearly mean rural background concentrations in 2009 over Europe. This map is shown alongside with the maps produced for the 2005 data recalculated from the erratum for Gräler et al. (2012a) in Figure 5 (left column), also produced with the metric covariance model with 100 neighbours, but with a temporal scale of 120 km/day, 3-f. This method was selected then also as best performer according to the statistics on yearly level. The standard deviation maps shown in Figure 5 (right column) approximate the standard deviation of the predicted annual means. They are derived as described by Denby et al. (2008b, Appendix A) using a temporal correlation coefficient of 0.3. However, temporal interval kriging as introduced in Section 2.6 provides more correct values than the approximate approach suggested by Denby et al. (2008b). See Section 5.3 for a discussion on temporal interval kriging.

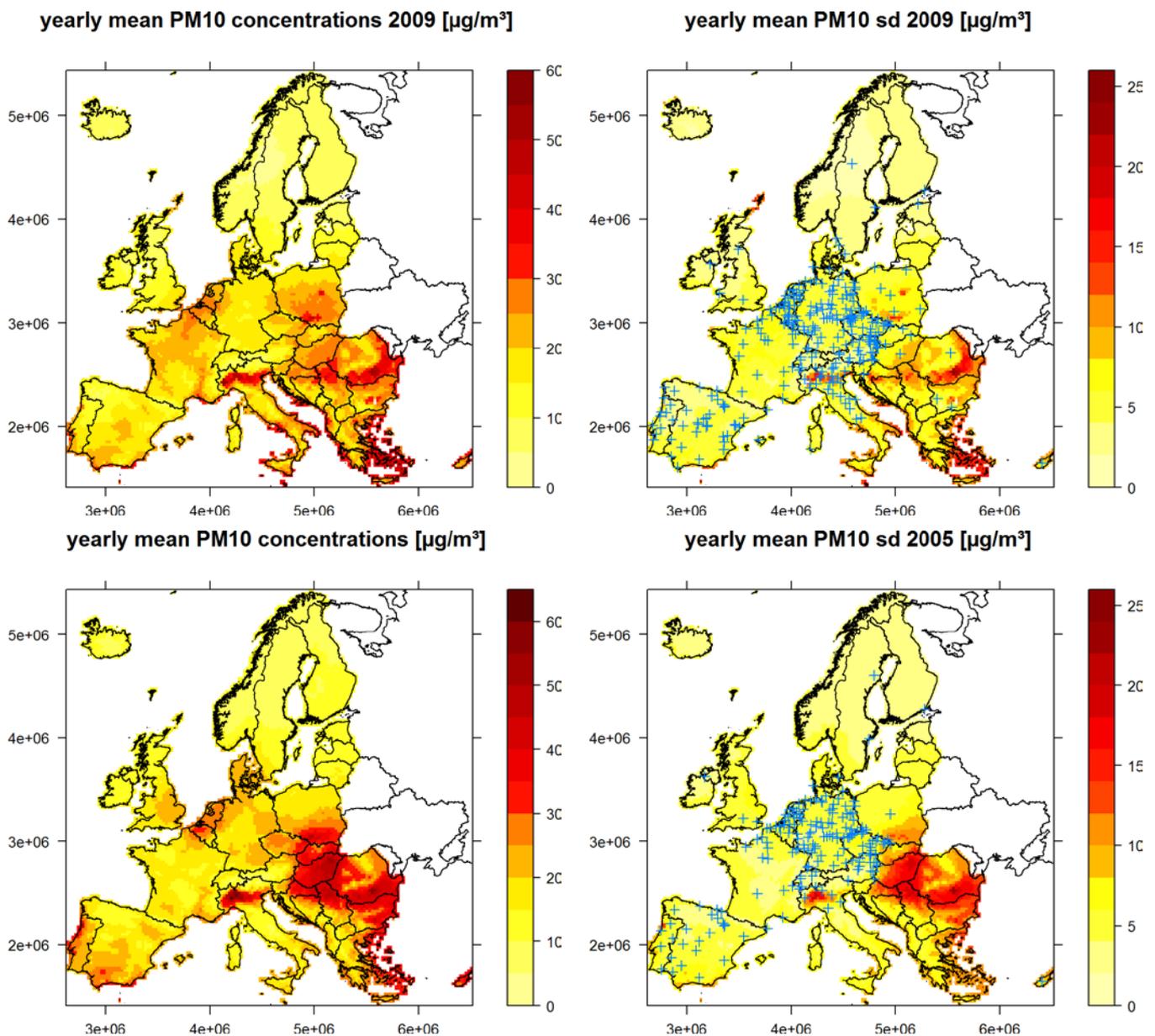


Figure 5: Interpolated maps of yearly mean PM₁₀ concentrations (left) in Europe for 2009 (top row) and 2005 (bottom row) and the approximate annual standard deviation of the predictions calculated as by Denby et al. (2008b) using a temporal correlation coefficient of 0.3 (right). The method adopted is the metric covariance model using 100 neighbours (3-f). Blue crosses denote rural stations used to produce the maps. The 2005 maps have been recalculated after the erratum for Gräler et al. (2012a) had been published.

The yearly mean PM₁₀ concentration maps for 2009 show a similar picture for Western, Central and Northern Europe and reduced concentrations in Eastern Europe. As the back-transformation of the logs relies on the kriging variance as well, areas with a high kriging variance show higher estimates. The kriging variance is especially high in the extrapolated eastern part of Europe due to the lack of stations. Therefore, the back transformed predictions appear to be exaggerated and are associated with a high uncertainty. The new stations appearing in the maps of 2009 reduce this effect considerably and the annual average as well as the standard deviation estimates decrease. A reduction of this order of magnitude is mainly due to this statistical effect.

4 Detecting asymmetric temporal dependencies

This chapter explores temporal asymmetric dependencies. Methods addressed so far assumed symmetrical dependence structures, which could distort the predictions if asymmetries were present. Assuming dominant environmental influences such as a prevailing wind direction raises the question of asymmetric temporal dependencies. That is, a field of moving air pollutants travelling in a particular direction will induce stronger dependencies between observations made at location s_0 and location s_1 when these lie in the main wind direction. In order to accomplish this effect, we started with a metric covariance model and sheared the temporal dimension into different spatial directions for a set of displacement lengths indicating the strength of this effect. A schematic sketch is shown in Figure 6 illustrating how the sheared spatio-temporal space reduces some spatio-temporal cross distances (i.e. $d((s_0, t_0), (s_1, t_1))$) while it enlarges others (i.e. $d((s_1, t_0), (s_0, t_1))$). Note that distances in the spatial and temporal domain are kept constant. Thus, the temporal distances within one station are maintained, as well as the spatial distances between the stations at each individual moment of time. The sheared space implies taking a Lagrangian frame of reference instead of a Eulerian frame of reference, when constant and uni-directional flow is assumed.

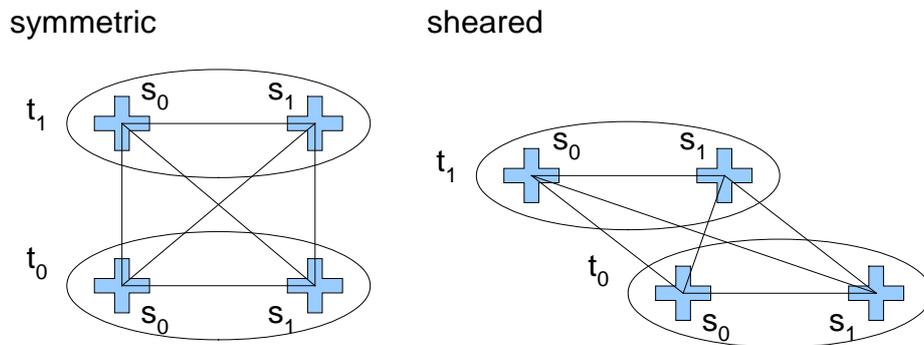


Figure 6: Schematic sketches of the sheared versus the symmetric spatio-temporal space.

In our set-up, we distorted the space into eight directions (North, North-East, East, ..., North-West) and in steps of 50 km displacement in the spatial plane. A new empirical variogram has been calculated for each of the distorted spaces and a variogram model has been fitted. This experiment was conducted for raw PM_{10} concentrations and residuals at 164 km/day and 360 km/day temporal scales each. Comparing the set of empirical variograms and their model representations did not reveal any strong asymmetric effect in the data. In fact, most plots showed practically identical patterns for all 8 directions for the same strength of displacement. The strongest effect occurred for the raw data with a displacement of 350 km in the spatial plane between two days. The variograms for all directions for this case are illustrated in Figure 7. From visual inspection, the variograms along the North-South axis show a reduced nugget effect that occurs at both the empirical and the fitted variogram. This indicates an improved model for this North-South and South-North shearing direction.

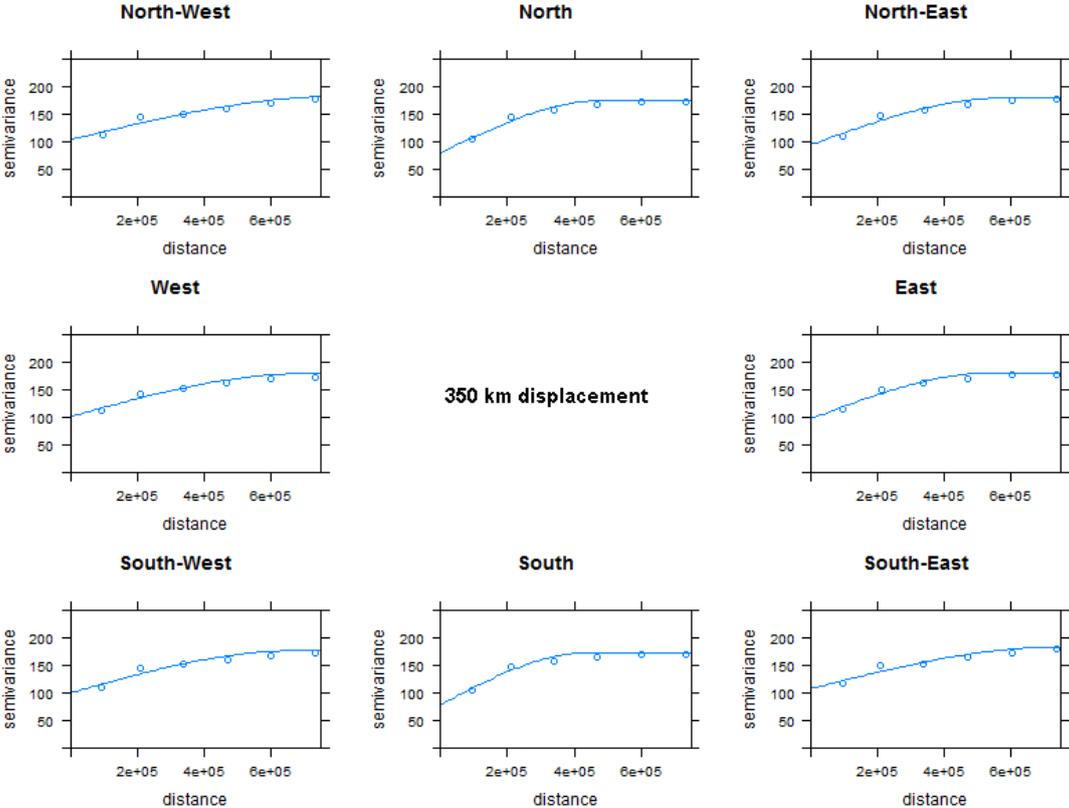


Figure 7: Strongest temporal asymmetric dependence structures shown for 8 directions and a displacement of 350 km of the raw observations.

5 Discussion

5.1 How to select the temporal scale?

Extending our studies from Gräler et al. (2012), we further explored the introduced spatio-temporal approaches. One parameter that needs to be set for the metric covariance model is the temporal scaling that relates *days* to *meters* yielding a meaningful metric in the covariance space. We investigated a set of scaling parameters and applied the full cross-validation procedure. The simple heuristic of taking the ratio of the purely spatial and purely temporal range performed already well in the cross-validation. However, trying distances beyond this value of 164 km/day improved the cross-validation statistics with an optimum at 360 km/day. A visual inspection of the empirical and modelled variogram at 360 km/day raised questions as some periodic patterns occurred (see Figure 2).

5.2 How robust are these methods?

Following the conclusion of Gräler et al. (2012a) based on 2005 data, we investigated the robustness of the earlier proposed methods for another year, 2009. The three types of methods selected have been the best performing ones in the preceding study. As expected, the more flexible product-sum model (3-g) again outperforms the separable covariance model (3-h) which assumes no interaction in the spatio-temporal cross-correlations. Different from the earlier study, the metric covariance model (3-f, 100 neighbours, 360 km/day temporal scaling) is not the overall best method on a yearly nor on a daily level. The best approach on a daily level is spatio-temporal kriging on the raw data (1-f) while the best approach for annual mean predictions is the pure spatial interpolation of annual means. However, all methods are close in terms of their cross-validation statistics. The ranking of spatio-temporal interpolation methods for 2009 was very similar to that of the 2005 data. However, the spatio-temporal approaches do not outperform the pure spatial interpolation of annual means on a yearly level. In terms of producing a map of annual averages, no statistical improvement is present but the daily predictions allow the analysis of the process at a higher temporal resolution. The actual days on which a threshold is exceeded could for instance be analysed, as opposed to only the total number of days over the year as currently done. To further investigate the potential of spatio-temporal interpolation procedures, the interaction of log-transformations, linear model and kriging should be investigated for air quality data.

5.3 Is temporal interval kriging suited for further analysis?

We investigated a set of temporal interval kriging approaches on their potential in predicting the yearly average PM_{10} concentrations and evaluated their performance in the cross-validation set-up. The main advantage of temporal interval kriging is its property to deliver valid prediction variances. The prediction variances we show in Figure 5 are only an approximation following Denby et al. (2008b). Figure 9 shows the prediction standard deviations as evaluated during the cross validation procedure for the adopted approximation method (top, see Section 3.2) and the temporal interval kriging approach (bottom). The coloured dots indicate stations for which the prediction standard deviation has been calculated. It is not only the magnitude of the values that defers; it is as well the spatial spread of the uncertainty. In the temporal interval kriging approach, the standard deviations follow the kriging property of the variance to depend only on the prediction locations and as such represent in some sense the network density. In the adopted approximation method the differences are not only caused by the approximation itself, but also caused by the fact that the

back transformation of afore log-transformed values depends on the magnitude of the prediction value. Thus, stations being predicted to be high show a higher variability. These two maps illustrate the importance of considering the right uncertainty measures in analysing environmental phenomena.

As temporal interval kriging provides by construction the correct variability for its estimates, it provides a sound basis for annual or monthly trend analysis including the uncertainty of the interval means. Obviously, the model selection and the fit of the model needs further research to produce as accurate estimates as the best performing daily interpolation methods.

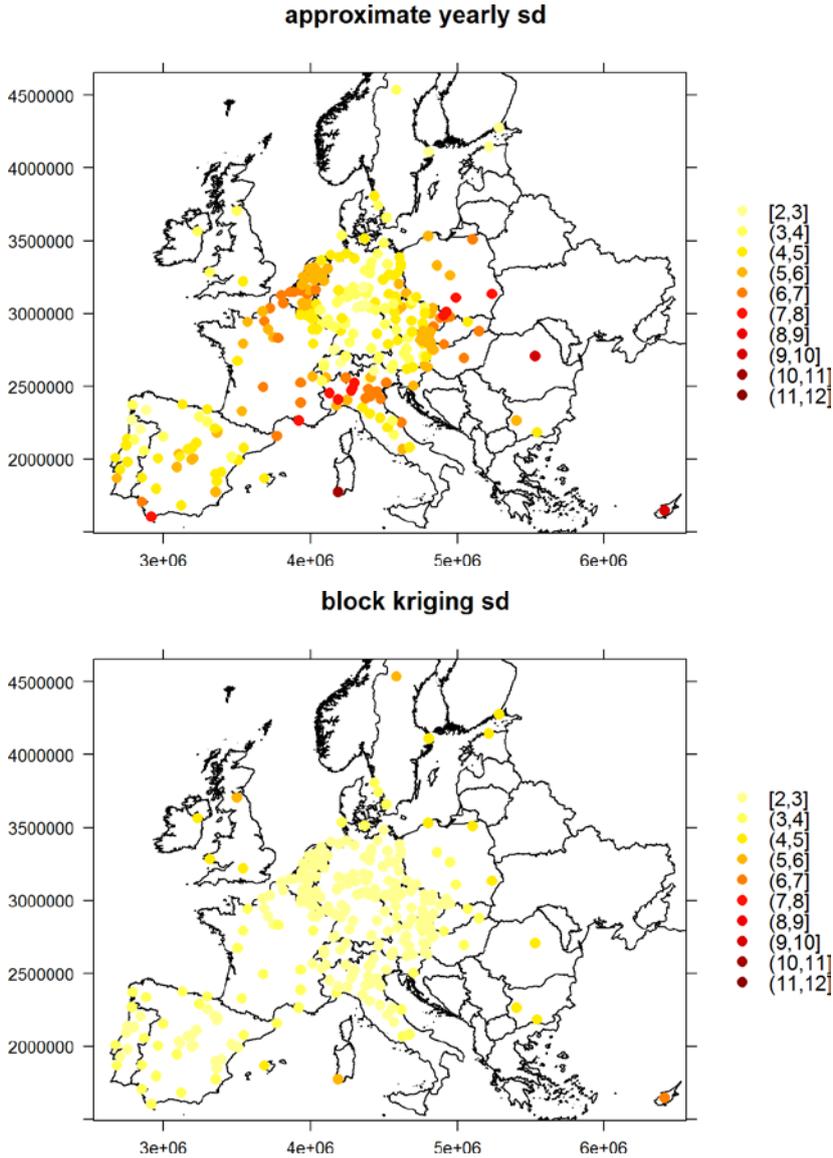


Figure 9: Comparison of prediction standard deviation estimate and block kriging prediction standard deviation. Dots indicate stations for which the prediction standard deviation has been calculated during the cross-validation procedure.

5.4 Do we need to consider asymmetric patterns?

The study described in Chapter 4 did not reveal any strong asymmetric patterns, i.e. no prevailing transport effect for the residuals in method 3-f and a weak asymmetry for the raw measurements in method 1-f along the North-South axis. On the one hand one may assume that the EMEP model, which includes wind data, is capable of capturing the transport effect to a certain degree producing residuals with a symmetric dependence structure. On the other hand the spatio-temporal extend of the study area and the various environmental conditions one will find across Europe throughout the year may generate a mixture of asymmetric dependence structures that to some degree cancel out each other. When one extends the analysis to non-stationary random fields in space and/or time, the asymmetric dependence structures might become evident again. Finally, the assumption of constant, uni-directional flow may be overly simplistic, and using wind fields to assess more realistic asymmetries may yield better results.

However, Gräler and Pebesma (2012b) used a vine-copula approach to model the spatio-temporal dependencies of the raw measurements. In their work, they could achieve the best fit for asymmetric copulas looking at lag classes of one or two days and several spatial distances. These asymmetric copulas allow the introduction of different strengths of dependence. Pairs of values where, for example, the station reports a high value the first day and a low one the following day may exhibit a different strength of dependence than for a low value at the first day and a high one the second day. These asymmetries refer, e.g., to a slow increase but rapid decay of concentrations as present in air quality data due to slowly accumulating toxics and rapid decays due to sudden weather events. In the decay phase, high values are stronger associated with low values the following day, while low values may rather unlikely be succeeded by very high ones due to a typically slow increase phase and thus the strength of dependence of this one day time gap differs for different combinations of measurement values. In the Gaussian realm, this scenario will always be modelled with identical strength of dependence.

This brief study did not reveal any asymmetric dependencies due to prevailing environmental effects but an in parallel conducted study showed asymmetric dependencies in a different realm. Thus, we can reject the issue of a dominant asymmetric effect but need to consider additional types of asymmetric dependencies to further improve the model. These improvements would reduce the smoothing and symmetrising effect of the Gaussian realm.

6 Conclusions

The cross-validation results show that the different heuristics to estimate the temporal scaling have only a minor influence on the prediction. The model selection and variogram parameter estimation will have a stronger effect. The exploration of further heuristics would help to study the spatio-temporal cross dependencies, but the impact on the predictions would remain minor.

Results are comparable with Gräler et al. (2012). The cross-validation statistics on an annual level are in general better for the 2009 data. Only the approaches not relying on a regression in 2009 outperform the spatio-temporal approaches in 2005 on a daily level. This study showed that the results from the cross-validation are not strictly robust against year-to-year variability for these years, but the results remain close. In especially, the pure spatial interpolation procedure to derive annual estimates from annual means could not repeatedly be outperformed. However, we believe that the spatio-temporal interpolation methods provide great potential for the production of PM₁₀ concentration maps over Europe on a daily basis. The availability of daily EMEP model data may, however, limit the applicability of the regression based approaches. This and the previous study have only been conducted for rural background data. In general, any of the above methods are as well applicable to urban stations but further studies are needed to assess their performance.

The set of temporal interval kriging approaches did not perform very well in the cross-validation study, but in principle their ability to provide precise prediction error estimates is very valuable in uncertainty analysis. This becomes especially important when long term European wide trends are of interest. Section 5.3 and Figure 9 illustrate the differences appearing in the uncertainty analysis using different methods. The temporal interval kriging approach provides valid kriging variances associated with the interval predictions where the alternative approach only provides a proxy of the variability of the predictions. The differences are roughly of a factor of 2 and the spatial patterns differ completely. However, further research on the model would be needed to validate and re-adjust this approach. The active discussion on trends in air quality measures will benefit from the improved uncertainty measure, as it provides a valid variability of the prediction. Besides annual estimates, the temporal interval kriging approach can as well be used to generate, for instance, monthly estimates to study within year variability. Alternatively, stochastic simulations may be used to improve the uncertainty assessment (e.g. for models using log transformed observations).

In a very comprehensive first assessment, no strong global asymmetry effect due to a prevailing transport direction could be discovered in the data set. Especially when EMEP modelling data is incorporated this appears to be the case and may have its cause in the model being capable of capturing asymmetry in transport effects. Nevertheless, the raw data showed a reduced nugget effect in both the empirical and fitted variogram along the North-South axis indicating an improved fit. However, the spatial and temporal extend of the data set may blur the effect of climatological, meteorological and seasonal asymmetric dependencies. Furthermore, this is only one type of asymmetry that may occur in spatio-temporal data. Gräler and Pebesma (2012b) showed that asymmetric copulas best fitted the data for many non-zero temporal lags. These copulas are capable of mimicking dependencies that occur in processes with e.g. a slow increase, but rapid decay. We suggest exploring further different types of asymmetric dependencies and non-stationarity to improve the kriging part of the complete exposure assessment. Some of the non-stationarity effects potentially disturbing the predictions might be reduced by a reassessment of station types as some rural background stations show a deviating pattern from the majority.

An approach circumventing the effect of the exaggeration of the back-transform from the log scale is to estimate the median instead of the mean. The median does not get a bias by the log transform and the back transform does not have to incorporate the prediction variance (Pebesma and De Kwaadsteniet 1997).

References

- Cressie, N., C. Wikle (2011). *Statistics for Spatio-Temporal Data*. Wiley, New York.
- De Iaco, S. (2001). Space–time analysis using a general product–sum model. *Statistics Probability Letters*, 52(1), pp. 21 - 28
- Denby, B., J.Horálek, P. de Smet, F. de Leeuw and P.Kurfürst (2008a).European scale exceedance mapping for PM10 and ozone based on daily interpolation fields. ETC/ACC Technical paper 2007/8.
http://acm.eionet.europa.eu/reports/ETCACC_TP_2007_8_spatAQmaps_dly_interpol
- Denby B., M. Schaap, A. Segers, P. Builtjes and J. Horálek (2008b). Comparison of two data assimilation methods for assessing PM10 exceedances on the European scale. *Atmospheric Environment*. 42, 7122-7134.
- EEA. AirBase, European air quality database, version 6. EEA. <http://airbase.eionet.europa.eu/>
- Fagerli H, et al (2011). Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe in 2009. EMEP Report 1/2011.
http://emep.int/publ/reports/2011/status_report_1_2011.pdf
- Gräler, B., L. E. Gerharz, & E. Pebesma (2012a): Spatio-temporal analysis and interpolation of PM10 measurements in Europe. ETC/ACM Technical Paper 2011/10, January 2012.
http://acm.eionet.europa.eu/reports/ETCACM_TP_2011_10_spatio-temp_AQinterpolation
- Gräler, B. and E. Pebesma (2012b): Modelling Dependence in Space and Time with Vine Copulas. Presented at: Geostats 2012, Oslo, Norway, 11-15 June 2012.
http://ifgi.de/graeler/publications/SpatioTemporalCopulas_GraelerPebesma.pdf
- Horálek, J., B. Denby, P. de Smet, F. de Leeuw, P. Kurfürst, R. Swart and T. van Noije, (2007). Spatial mapping of air quality for European scale assessment. ETC/ACC Technical paper 2006/6.
http://air-climate.eionet.europa.eu/reports/ETCACC_TechPaper_2006_6_Spat_AQ
- Pebesma, E., accepted. spacetime: Spatio-Temporal Data in R. *Journal of Statistical Software*.
<http://CRAN.R-project.org/package=spacetime>
- Pebesma, E. (2004). Multivariable geostatistics in S: the *gstat* package. *Computers & Geosciences*, 30: 683-691. <http://CRAN.R-project.org/package=gstat>
- Pebesma, E. and J.W. de Kwaadsteniet (1997): Mapping Groundwater Quality in the Netherlands. *Journal of Hydrology*. 200, pp. 364-386
- R Development Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
<http://www.r-project.org>