Response of the authors:

The authors would like to thank the reviewers for their supportive and valuable feedback, questions, and comments. In the following, we are giving answers and explanations to their questions and comments. The text in the paper is adapted in different places according to the reviewer notes.

hess-2014-489-referee-report-1	**Response of the authors**
I feel a little bit uncomfortable with the issue that MERIS data are applied for validation as they only provide the water vapor above the clouds.	MERIS data measures the water vapor content up to the Earth's surface under cloud-free sky. If clouds exist, then the water vapor content is measured up to clouds top. Therefore, we used only 5 MERIS maps that were observed under clear skies.
Major remarks:	
page 3, line 7: "Various research suggested the assimilation of atmospheric measurements into these models to improve the quality of the data." As data assimilation is for a long time a standard operation in atmospheric modelling, this statement lacks information and does not help to introduce the research question (at least one citation is missing here). On the contrary, the statement asks for a comparison of the data fusion approach to a data assimilation approach, appealing by a finally physically consistent picture.	The sentence is modified and references are added to make the following clear: "The assimilation of measurements and measurement-based estimates into the model is a well-known research topic and it is beyond the scope of this work. We want to benefit from the model data to the support the measurement-based estimates of water vapor and to determine its values in regions where the measurements are missing."
page 3, line 10: "We want to comprehend if the model simulations of water vapor, in their current quality, together with other measurement-based estimates can provide complete knowledge about the atmospheric water vapor." The formulation is in my opinion misleading as the main data source in your case is not the mesoscale model but the remote sensing data. You improve these data by adding a mesoscale model. If you state it like this, you should first go for an improvement of the model simulation. By the which seems to be nearly cloud free. So better rephrase your research question.	The sentence is modified. We agree that complete knowledge means 3D grids with high spatial resolution, which is the further step to this research.
page 4 line 22: The Wang and Seaman (1997) citation is too old / inappropriate, as a) they apply precipitation, sea level pressure, wind, and temperature predictions for model evaluation and b) there has been a great development of mesoscale models in the recent years approaching the convective	The publication of Prein et al. 2015 is indeed an interesting review article about the state of mesoscale modeling. However, it was published just shortly after the finalization of the first review round. While primarily focusing on climate modeling the paper gives also a comprehensive

permitting scale (Prein et al., 2015). If you cite Wang and Seaman (1997), you should also discuss the influence of other parametrizations then that one for convection, also responsible for water vapor, like microphysics, planetary boundary layer and land surface. A simple search in Web of Science or Scholar Google shows some more appropriate papers dealing explicitly with water vapor in mesoscale models (e.g. Wilgan et al. (2015) or citations in it).	overview about the current state of the art of convection-permitting models. The key message of Prein et al. 2015 with respect to our study is that the downscaling of large-scale-models provides added value for the representation of several processes and the development of consistent spatial structures, in particular extreme precipitation and convective processes. However, there are still many open challenges, especially for the representation of turbulent processes and microphysics schemes. We would like to include the proposed reference (and replace the citation of Wang and Seaman, 1997) in a way that corroborates the general message of the paragraph which is that there are still many processes that lack a Therefore, we propose to change the sentence p4 l22 "The presence of convective motion or rapid dynamic effects are still a challenge for the performance of LAMs (Wang and Seaman, 1997). Hence, the model data can be considerably biased with respect to the actual state of the atmosphere. This, in addition to the configuration of the model domains, can significantly impact the simulation output [...]" to: "Despite manifold improvements over the last years, considerable uncertainties are still connected with the parameterization of physical processes in mesoscale-atmospheric models and biases of the driving model (Prein et al. 2015). This can significantly impact the simulation output [...]"
Minor comments:	
page 4 line 22 "The the" -> "The" page 11 line 2 "signalis" -> "signal is"	Done Done

hess-2014-489-referee-report-2	Response of the authors
1. I do not understand the plots (negative values for PWV???) Either explain more precisely what we see on the images or stick for all plots to what you did in Figure 1 – which is in my understanding the best way to plot PWV.	These are not absolute values. It is written in the caption that a linear trend is subtracted from the maps. This results in the negative values on the color bars. This is now added in text.
2. Driving model resolution – why ERA Interim when higher resolution ECMWF analyses are available? I think that it might be better to use shorter simulations driven by higher resolution analyses to show the benefits of your procedure for single cases. Then it is also necessary to use more than 42 vertical levels to better represent the tropospheric structure and especially the lower troposphere where most of the water vapor is.	Thank you for the comment. Comment 2 and 3 are discussing the point of improving the model output data through data assimilation before further using them. We totally agree on that point, and we are working on improving the model performance by testing input data at different spatial resolutions, paying attention to the physics and driving models. However, the method presented here used the model in its current quality together with measurement-based estimates in a data fusion approach that depends on the spatial properties of the input data. We benefit from the model particularly in regions with no measurements taking into account the spatial covariance properties from both the model and measurements to achieve the best results. If the model presents a bad scenario, then the output map are dominated by the measurement-based estimates. One point is important to consider, that is the systematic bias in the model. Therefore, we estimated the long wavelength signal using the measurement. We agree with the reviewer's opinion that a better resolution would be desirable for driving the local area model, especially when short term simulations are performed. However, although the operational analysis of ECMWF features a higher output precision then ERA-INTERIM, a smaller number of observations is ingested due to the need for real-time availability. Moreover, the model configuration of ECMWF's Analysis is not consistent over time i.e. model physics, data assimilation sources, and also resolution are subject to change. In our opinion, a reanalysis is much more consistent and of higher quality even if the output resolution is lower. For the year 2005 the resolution of ECMWF's Analysis is about 40 km which is about double as with respect to INTERIM.

	To make this clearer, we propose to change the following sentence (p. 8 l. 9) as follows: "Because neither gridded nor spectral nudging was activated, the local area model physics fully determine the propagation of moisture through the respective domains." is changed to "Neither gridded nor spectral nudging was activated in order to conserve the model's internal water balance. Hence the GCM boundary fluxes and the local area model physics solely determine the propagation of moisture through the respective domains." The selection of the number of layers was also based on the work of Berg et al. 2013. However, the layers are not equidistant as vertical resolution is increased for the lower troposphere where the major proportion of atmospheric vapor is located. We propose the following change of the sentence (p.7 l. 7.): "Vertically, the model divides into 42 layers and the model [...]" is changed to "Vertically, the model divides into 42 layers with variable distance. The resolution is increased for the lower troposphere where most of the atmospheric vapor resides. The model top is defined at 50 hPa."
3. When you merge model data with observations to create a “new” observation, it is necessary to make sure that the model provides the best possible result. This is only possible when you also use data assimilation to initialize the model. ERA Interim contains data assimilation – but I am sure that the WRF simulation would highly benefit from doing a re-assimilation of observations in the nests.	From the time slices that were used in this study one can see that the combined product of fusioned mesoscale model and point observations lead to an overall improved product of atmospheric water vapor. We agree that with better model quality the final results could be even better. However, the optimization of the mesoscale model is out of the scope of this study as the focus is on the methodology of the fusion approach. With the chosen model configuration, with 5 months of spin up time, the initial conditions would not play a too important role on the final results of summer 2005. Therefore, we did not consider the re-assimilation of observations.

4. Is it necessary to do such long simulations when you focus on single days? True – the deeper layer soil moisture needs long spin-up times – but how large is their influence on the daily evolving weather? The topmost layer directly influenced by the atmosphere reacts faster. And only this layer influences PWV.	We agree with the reviewer, that there might be weather conditions where local soil moisture is not an important driver for the local atmospheric water vapor content. However, for stable synoptic conditions with poor or no moisture advection, the state of the soil becomes quite important. This includes also the water content of the lower soil layers as they feed the transpiration process.
5. Model results are clearly dependent on resolution, physics, and driving model – and therefore lead to clearly different results for different synoptic situations. This is the more the problem when you drive the model with such a coarse resolution analysis and then look to single cases.	It is right that there are myriad configurations to run the mesoscale model. For that reason, we employ a setup which already proved to produce valuable results (Berg et al. 2013). Furthermore it would be desirable to extend the analysis to a larger set of scenes and also increase the spatial extent. However, this was not possible for the current study as the density of GNSS sensors is only sufficient for the Rhine Graben region and because of the breaking down of Envisat.
6. The results do not show the benefit of the methodology – the region you selected for the results is by far too small. I guess if you enlarge the region to e.g. at least Germany you might be able to find systematic pros and cons of the methodology	Yes, we agree that might be true. However, the focus of this work is provide water vapor maps of high spatial resolution, where the PS InSAR data are of great benefit. The fine scale signal has a spatial of length of several meters; therefore, an area of 100 km× 100 km is quite good for this research.
7. Some of the plots are too small (labels in figure 1, Figures 10, 12 and 13)	The figures are modified
Minor points: 1.) Give some information in the introduction what you think are the target groups that will work with your data.	This is added.
2.) Page 2, line 19: Why do you introduce the term “neutrosphere”?	The Earth’s atmosphere contains the ionospheric layer and an electrically-neutral layer called the neutrosphere (up to 50 km), where water vapor is located. The most of it is located near the surface but less than 1% can exist in altitudes above 12km (troposphere).
3.) Page 3, line 9: Replace “want to” by “would like to”	We think “want to” is better in this context.

4.) Page 4, line 22: Delete one “the”. What you describe in this sentence is exactly what you improve (optimize as well as possible) when data assimilation is applied.	“the” is deleted.
5.) Page 5, line 1: I would delete “redundant” – you never know whether a data set is redundant– this might change from case to case.	Yes, that’s correct. But, we are saying that it can be redundant, not asserting.
6.) Page 5, line 23: “space-bourne” instead of “space-based”?	“Space-based” is known and used in this context. So, we would like to keep it.
7.) Page 6, line 3: “Rhine valley” instead of “Rhine Graben”	Thank you. We prefer “Rhine Graben” since it is used in other publications.
8.) Page 6, line 17: “In cloudy conditions” instead of “Under cloud cover”	Modified
9.) Page 8, line 14: “Rhine valley” instead of “Rhine Graben”	We prefer “Rhine Graben” since it is used in other publications, too.

Manuscript prepared for Hydrol. Earth Syst. Sci. Discuss.
with version 2014/05/30 6.91 Copernicus papers of the LaTeX class copernicus.cls.
Date: 28 October 2015

Water vapor mapping by fusing InSAR and GNSS remote sensing data and atmospheric simulations

F. Alshawaf[1,3], B. Fersch[2], S. Hinz[1], H. Kunstmann[2], M. Mayer[3], and F. J. Meyer[4]

[1]Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
[2]Institute of Meteorology and Climate Research, Campus Alpin, KIT, 82467 Garmisch-Partenkirchen, Germany
[3]Geodetic Institute, KIT, 76131 Karlsruhe, Germany
[4]Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK 99775 USA

Correspondence to: F. Alshawaf (fadwa.alshawaf@kit.edu)

Abstract

Data fusion aims at integrating multiple data sources that can be redundant or complementary to produce complete, accurate information of the parameter of interest. In this work, data fusion of precipitable water vapor (PWV) estimated from remote sensing observations and data from the Weather Research and Forecasting (WRF) modeling system is applied to provide complete grids of PWV with high quality. Our goal is to correctly infer PWV at spatially continuous, highly-resolved grids from heterogeneous data sets. This is done by a geostatistical data fusion approach based on the method of fixed-rank kriging. The first data set contains absolute maps of atmospheric PWV produced by combining observations from Global Navigation Satellite Systems (GNSS) and Interferometric Synthetic Aperture Radar (InSAR). These PWV maps have a high spatial density and a millimeter accuracy; however, the data are missing in regions of low coherence (e.g., forests and vegetated areas). The PWV maps simulated by the WRF model represent the second data set. The model maps are available for wide areas, but they have a coarse spatial resolution and a yet limited accuracy. The PWV maps inferred by the data fusion at any spatial resolution show better qualities than those inferred from single data sets. In addition, by using the fixed-rank kriging method the computational burden is significantly lower than that for ordinary kriging.

1 Introduction

Water vapor is a vital constituent of the Earth's electrically neutral atmosphere (neutrosphere). Although the ratio of water vapor partial to total atmospheric pressure is typically below 4%, it is an important constituent in many respects. Due to the dynamic nature of the neutrosphere and the complex energy exchange with the Earth's surface, the spatio-temporal distribution of water vapor can be highly variable. Accurate information about its content and tendency is the main prerequisite for the prediction of clouds and precipitation. Water vapor is important for studies of climate and natural disasters such as floods, droughts or glacier melting. On the other hand, radio signals transmitted from spaceborne sensors are refracted when traversing the Earth's neu-

trosphere. The neutrospheric water vapor contributes to less than 10 % of the signal path delay; however, this error source is not easily eliminated. Accurate information about the water vapor concentration along the signal path is required, which is not always obtainable. Although many efforts have been made to produce accurate information about water vapor using ground-based, space-based or numerical methods, the available information is often limited in the temporal resolution, spatial resolution or accuracy (Bevis et al., 1992). Numerical atmospheric prediction models are increasingly used to provide simulations of the atmospheric parameters. Various research suggested the assimilation of atmospheric parameters, such as water vapor, estimated from Global Positioning System (GPS) or Interferometric Synthetic Aperture Radar (InSAR) into these models to improve the quality of the simulated parameters (Pichelli et al., 2010; Bennitt and Jupp , 2008). We want to comprehend if the model simulations of water vapor, in their current quality, can be used to even out the deficits of the measurement-based estimates, particularly in regions with no measurements. To achieve this purpose, a statistical data fusion approach is applied. The output water vapor maps can be used in tomographic approaches to provide 3D water vapor grids and in adjusting the parameters of numerical atmospheric prediction models. The remaining of this section presents the recent related research on water vapor using remote sensing data and atmospheric models.

The amount of remote sensing data available for monitoring the Earth and its atmosphere is growing in a rapid, continuous way. InSAR has proved its capability for detecting surface deformation, landslides, tectonic movements (Massonnet et al., 1993; Zebker et al., 1994) and for deriving digital elevation models (Zebker and Goldstein, 1986). The influence of water vapor in the observations can be reduced by averaging a large number of interferograms (Zebker et al., 1997) or by time series analysis that indicates the stable persistent scatterers (Ferretti et al., 2001; Hooper et al., 2007). Besides, InSAR has recently been used to derive the phase shift caused due to the propagation in the Earth's atmosphere from the interferograms or by time series analysis (Hanssen, 2001; Meyer et al., 2008; Pichelli et al., 2010; Alshawaf et al., 2012). Global Navigation Satellite Systems (GNSS), however, have been considered since the 1990s as an efficient microwave-based tool for atmospheric sounding (Bevis et al., 1992; Rocken et al., 1995). Since then, numerous methods exploited the GNSS observations to produce estimates

of the integrated atmospheric water vapor, and for generating water vapor maps (Luo et al., 2008; Jade and Vijayan, 2008; Karabatić et al., 2011). InSAR and GNSS signals are affected in a similar way by the atmosphere (Onn and Zebker, 2006). Therefore, Alshawaf et al. (2015b) presented a new approach to derive absolute, high-resolution maps of precipitable water vapor (PWV) by combining data from InSAR and GNSS. The SAR systems acquires the images at repeat cycles of multiples of days. Enivsat images, which are used in this work, are available at multiples of 35 days. The availability of the data over time can be increased by processing data from ascending and descending modes. In addition, new SAR missions have shorter repeat cycles, 11 days for TerraSAR-X and 6 days for Sentinel-1. The InSAR-based PWV estimates cannot be used to observe the variability of water vapor over short time, but they are important in different aspects. This geodetic-based method produces maps of the PWV at a high spatial resolution without additional costs. These data can be exploited, first, to model the spatial variations of atmospheric turbulent and nonturbulent effects. Second, they can be used to observe the variation of water content over long time periods to detect, for example, unusual trends. Third, they can be used to adjust/readjust the initial and boundary conditions in atmospheric prediction models.

Atmospheric modeling systems are standard approaches to simulate three-dimensional distributions of the neutrospheric water vapor at various temporal and spatial sampling. Dynamic local area models (LAM) are common tools for scaling down the coarse grids of global circulation models to meso-scale applicability. Several studies employed the Weather Research and Forecasting modeling system (WRF, Skamarock and Klemp, 2008) to compare the LAM simulations of PWV with GNSS point estimates (Mateus et al., 2010; Bender et al., 2008; Cimini et al., 2012) and PWV maps from MERIS (MEdium Resolution Imaging Spectrometer) (Alshawaf et al., 2012). These studies conclude that the medium to long scale (greater than 20 km) water vapor signals can be well predicted, whereas short scale fluctuations are often hardly captured in a realistic way.

Despite manifold improvements over the last years, considerable uncertainties are still connected with the parameterization of physical processes in mesoscale-atmospheric models and biases of the driving model (Prein et al., 2015). This, in addition to the configuration of the

model domains, can significantly impact the simulation output (Gong et al., 2010) as well as the model intrinsic water balance (Awan et al., 2011; Fersch et al., 2012; Fersch and Kunstmann, 2014). Therefore, the setup of the local area model is crucial, and it has to be proper for the study region and the research objectives.

Due to the availability of various data sources, which can be complementary or redundant, data fusion has received increasing attention in the Earth observation studies. The focus is put on the combination of multiple sources, which may be spatially, temporally, or spectrally inhomogeneous, to produce a more complete representation of a geophysical process. In this work, we use remote sensing data and numerical atmospheric models through a data fusion approach to provide improved information about the distribution of atmospheric water vapor. This information is important not only for weather forecasting and climate research, but also to better understand how the InSAR interferograms are affected by water vapor, and to select the most appropriate method for reducing this noise. In turn, reliable local water vapor maps can support adapting the WRF model configurations and, hence, may improve the model performance.

In the following, we present water vapor maps derived from microwave remote sensing data and numerical atmospheric models. Since the available data have different spatial levels of aggregation, it is important to discuss the change of support problem. Then, we present the data fusion approach based on the kriging or fixed-rank kriging techniques. We first describe the ordinary kriging and how it can be extended for fusing multiple data sets. Then, we present the reasons behind using the fixed-rank kriging. We use the data fusion approach for predicting maps of the atmospheric PWV from remote sensing data and atmospheric models.

2 Atmospheric water vapor

Several observation systems are commonly used to continuously monitor the vertical and horizontal distributions of water vapor in the atmosphere. These devices are used either from the ground, such as radiosondes and ground-based water vapor radiometers, or from space, such as space-based water vapor radiometers and infrared sensors. In this work, we employ microwave

remote sensing systems as well as numerical atmospheric models to provide accurate maps of the atmospheric water vapor at a high spatial resolution.

2.1 Water vapor from remote sensing data

Alshawaf et al. (2015b) presented a new approach to derive absolute, high-resolution maps of PWV by combining data from InSAR and GNSS. The data are collected in the region of Upper Rhine Graben in Germany and France over the period 2003–2008. Persistent Scatterer InSAR (PSI) using the Stanford Method for Persistent Scatterers (StaMPS, Hooper et al., 2007) was applied to derive PWV maps from the InSAR interferograms. These maps contain the water vapor signal of short scale spatial variations, while the elevation-dependent and long wavelength water vapor components are eliminated when forming the interferograms or by phase filtering. Therefore, GNSS-based PWV estimates were used to reconstruct the missing components. The approach for combining InSAR and GNSS data is presented in details in Alshawaf et al. (2015a) and Alshawaf et al. (2015b). Fig. 1 shows a map of PWV derived by combining PSI and GNSS data and the corresponding map extracted from MERIS observations. MERIS is a passive imaging spectrometer located on board the Envisat platform. It measures the solar radiation reflected from the Earth's surface or clouds. The ratio of the radiance values measured at the channels 14 and 15, located respectively at 885 nm and 900 nm, are used to determine the vertical PWV content in the neutrosphere (Fischer and Bennartz, 1997). MERIS provides maps of the PWV at a spatial resolution of 260 m $\times$ 290 m (full resolution mode). Under cloud weather conditions MERIS measurements are highly underestimated since the measured PWV represents only the water vapor existing between the sensor and clouds top; therefore, only 5 MERIS PWV images were available for this study.

The PSI method produces information where stable persistent scatterers are identified, which requires a high coherence between the SAR images. In forests and vegetated areas, the probability to identify persistent scatterers is low; therefore, in these regions only sparse points are found. The white areas within the left figure indicate regions of low coherence and the corresponding data from MERIS are masked out. The spatial correlation between the maps is 95 % and the root mean square (RMS) value of the differences is 0.68 mm. We can observe that the

persistent scatterers are dense in the urban areas, while they almost disappear in the low coherence regions. Since PWV data are spatial, their covariance function is exploited by geostatistical techniques to reasonably infer the PWV at regular grids. In order to improve the inferred PWV maps, especially in the areas where the PWV estimates are sparse, we apply data fusion of the remotely-sensed PWV maps with maps produced by the WRF model.

2.2 Water vapor from regional atmospheric models

As depicted in Fig. 2, the WRF model (version 3.1.1, Skamarock et al., 2008) was set up with a parent domain of 27 km × 27 km resolution and two nests with 9 km × 9 km and 3 km × 3 km, respectively. Feedback from the nests to their parent domain was not activated. Vertically, the model divides into 42 layers with variable distance. The resolution is increased for the lower troposphere where most of the atmospheric vapor resides. The model top is defined at 50 hPa. The selection of the physical modules is based on the study of Berg et al. (2013); accordingly, the WRF single moment (WSM) 5-class scheme (Hong et al., 2004) was selected for microphysics. Short and longwave radiation were computed with the community atmospheric model (CAM) scheme (Collins et al., 2004). The processes in the planetary boundary layer were represented by the Yonsai University scheme (Hong et al., 2006). The surface layer was simulated with the Monin–Obukhov scheme, and the Noah land-surface-model (Chen and Dudhia, 2001) was applied for the surface physics. Sub-grid convective processes were included with the Kain-Fritsch parametrization (Kain, 2004). The global dynamic boundary conditions were ingested from the European Center for Medium-Range Weather Forecasts (ECMWF) ERA-INTERIM reanalysis at 6 hours interval (Uppala et al., 2008). In ERA-INTERIM, a broad range of different data sources is assimilated. For the atmospheric moisture analysis, ground based station observations, radiosonde profiles, and GPS radio occultation are exploited. Additionally, total column water vapor information from the Special Sensor Microwave/Imager (SSM/I) and the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) is assimilated (Dee et al., 2011). MERIS retrievals of column water vapor are not ingested into ERA-INTERIM and thus they depict an independent data set for our approach.

The WRF simulations cover the period between July 2004 and September 2005, such that the first 5 months were considered as spin-up. The PWV content was determined at every output time-step (10 min) by a vertical integration of all moisture fields from the land surface to the model top. Two output time slices were compared with the simultaneous MERIS observations. The long scale signal is modeled by a linear trend and subtracted from the maps; hence, negative values are observed on the color bars. From the compared maps shown in Fig. 3, we observe that the spatial atmospheric patterns are not always correctly resembled by the model. On 27 June 2005 (09:51 UTC), WRF and MERIS PWV maps are strongly correlated with a coefficient of 0.8, whereas the analysis of 5 September 2005 (09:51 UTC) shows a lower spatial correlation (0.71). While the patterns east to the Upper Rhine valley are reasonably resembled, an unexpected discontinuity exists in the area around 7.7° E, 48.7° N.

At the lateral boundaries, WRF ingests the mixing ratio concentration from the global model. Thus, for the presented simulation the global climate model lateral boundary conditions were applied to the first (outer) domain. Neither gridded nor spectral nudging was activated in order to conserve the model's internal water balance. Hence the GCM boundary fluxes and the local area model physics solely determine the propagation of moisture through the respective domains. For the analysis of 27 June, 2005, the atmospheric conditions were rather unexcited and varied slowly resulting in a good agreement between MERIS and WRF data. On 5 September, a quickly moving frontal system with a strong west to east gradient and a notch in the atmospheric vapor over the Upper Rhine Graben characterized the study region. It is not clearly distinguishable if the structure and dynamics of the ERA-INTERIM boundaries or the WRF model configurations are responsible for the discontinuity in PWV.

3 Change of support problem

Spatial data, for which close observations correlate more than distant ones, can be collected at points or areal units. The former are called point-level data or simply point data and the latter are areal-level or block data (Gelfand et al., 2001). In geostatistics, this defines the spatial support of the data. When both data types are available, data fusion can be applied to infer the underlying

process at any level of support. The change of support problem is concerned with the inference of the underlying process at point- or block-levels different from those at which the data are available. This also includes fusing data at different support levels. Based on the available input data and the desired output grid, there are four prediction possibilities: points to points, points to blocks, blocks to points, or blocks to blocks. These prediction possibilities may be collected under the umbrella of kriging (Cressie, 1993).

For block data that can be expressed as an average of point data as if it is collected within the block, such as rainfall, temperature, surface elevation, and atmospheric water vapor, the following model is appropriate

$$Y(B_i) = \frac{1}{|B_i|} \int_{B_i} Y(\boldsymbol{s}) \mathrm{d}\boldsymbol{s} \tag{1}$$

where $Y(B_i)$ and $Y(\boldsymbol{s})$ define the block and point data, respectively (Fig. 4). B_i refers to the block over which the data are aggregated and $|B_i|$ is the volume (or cardinality) of the data. The block-level covariance can then be related to the point-level covariance as follows:

$$C(B_i, B_j) = \mathrm{cov}\left(\frac{1}{|B_i|} \int_{B_i} Y(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}, \frac{1}{|B_j|} \int_{B_j} Y(\boldsymbol{v}) \mathrm{d}\boldsymbol{v} \right) \tag{2}$$

$$= \frac{1}{|B_i||B_j|} \int_{B_i} \int_{B_j} C(\boldsymbol{u}, \boldsymbol{v}) \, \mathrm{d}\boldsymbol{u} \, \mathrm{d}\boldsymbol{v} \tag{3}$$

where $C(B_i, B_j)$ is the block-to-block or block covariance function and $C(\boldsymbol{u}, \boldsymbol{v})$ is the point covariance function.

4 Spatial data fusion using kriging methods

4.1 Ordinary kriging

In geostatistics, a spatial process can be inferred over a continuous spatial domain by exploiting the covariance function as an important source of information. Predictions are obtained based either on single or multiple sets. Kriging is a geostatistical interpolation technique that infers values at new locations by considering spatial correlations (Cressie, 1993). The spatial density of the data points has to be enough to capture the covariance structure of the process. This information is represented by a variogram or covariance function, which is used to determine the predictions. If the considered spatial data set is denoted by $\boldsymbol{Z}$, then the kriging estimator $\hat{Y}(\boldsymbol{s}_0)$ at the location $\boldsymbol{s}_0$ is determined as follows:

$$\hat{Y}(\boldsymbol{s}_0) = \boldsymbol{a}' \tilde{\boldsymbol{Z}} \tag{4}$$

where the vector $\boldsymbol{a}$ contains the kriging weighting coefficients and $\tilde{\boldsymbol{Z}}$ is the centered data set (see Eq. (7)). The best linear unbiased estimator is found by solving the following constrained minimization problem:

$$\begin{aligned} \min_{\boldsymbol{a}} \quad & \mathrm{E}\Big\{(\hat{Y}(\boldsymbol{s}) - Y(\boldsymbol{s}))^2\Big\} \quad \text{subject to} \\ & \mathrm{E}\{\hat{Y}(\boldsymbol{s})\} = \mathrm{E}\{Y(\boldsymbol{s})\} \end{aligned} \tag{5}$$

The constraint is added to guarantee that the estimator is unbiased with respect to the true process $Y(\boldsymbol{s})$. A semivariogram function that reflects the spatial correlations is required to solve the minimization problem, which is determined from the detrended data, Eq. (A6).

The kriging method extends the spatial process using the following linear model:

$$\boldsymbol{Z}(\boldsymbol{s}) = \underbrace{\mathbf{T}(\boldsymbol{s}) \cdot \boldsymbol{\alpha} + \nu(\boldsymbol{s})}_{Y(\boldsymbol{s})} + \underbrace{\epsilon(\boldsymbol{s})}_{noise} \tag{6}$$

where $\epsilon(\boldsymbol{s})$ is an independent error term, which is assumed to be a white noise process with a mean zero and variance σ_ϵ^2. $\mathbf{T}(\boldsymbol{s}) \cdot \boldsymbol{\alpha}$ defines a deterministic linear trend, $\mathbf{T}$ has a size of $N \times 3$ and each row has the following entries: [1 longitude($\boldsymbol{s}$) latitude($\boldsymbol{s}$)]. N is the number of observations and $\boldsymbol{\alpha}$ is a vector of the least squares regression coefficient. $\nu(\boldsymbol{s})$ captures the spatial covariance structure of the process, and it is assumed to have a mean zero and generally a non-stationary covariance function. Before inferring the signal at a new location, it is required to center the data by estimating and subtracting the linear trend, i.e.,

$$\tilde{\boldsymbol{Z}} = \boldsymbol{Z} - \mathbf{T}\hat{\boldsymbol{\alpha}} \quad \text{with} \quad \hat{\boldsymbol{\alpha}} = (\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}'\boldsymbol{Z} \tag{7}$$

The detrended signal $\tilde{\boldsymbol{Z}}$ is used to determine the predictions in Eq. (4) and the deterministic signal is calculated from $T(\boldsymbol{s}_0)\hat{\boldsymbol{\alpha}}$. The sum of the two terms gives the total estimated value of $Y(\boldsymbol{s}_0)$. In the next section, a similar strategy is followed to solve for the best unbiased estimator using two data sets as presented in Braverman et al. (2009).

4.2 Spatial statistical data fusion

The spatial statistical data fusion (SSDF) is a method that statistically combines two data sets to optimally infer the quantity of interest and calculate the corresponding uncertainties at any predefined grid (Nguyen, 2009; Braverman et al., 2009). This method extends the kriging technique described above to find the optimal estimator using multiple data sets. Let the underlying process $Y(\boldsymbol{s})$ to be estimated at the location $\boldsymbol{s}$ from the data in $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ with the size N_1 and N_2, respectively. The estimator $\hat{Y}(\boldsymbol{s})$ at the location $\boldsymbol{s}$ is obtained from the two data sets as follows:

$$\hat{Y}(\boldsymbol{s}) = \boldsymbol{a}_1'\tilde{\boldsymbol{Z}}_1 + \boldsymbol{a}_2'\tilde{\boldsymbol{Z}}_2 \tag{8}$$

where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the fusion weighting coefficients, and $\tilde{\boldsymbol{Z}}_1$ and $\tilde{\boldsymbol{Z}}_2$ are detrended data sets of $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$, respectively. Following Eq. (5) and Eq. (8), the Lagrangian function L for the

minimization problem under the unbiasedness constraint is

$$L = \boldsymbol{a}_1'\boldsymbol{\Sigma}_{11}\boldsymbol{a}_1 + \boldsymbol{a}_2'\boldsymbol{\Sigma}_{22}\boldsymbol{a}_2 + 2\boldsymbol{a}_1'\boldsymbol{\Sigma}_{12}\boldsymbol{a}_2' - 2\boldsymbol{a}_1'\boldsymbol{c}_1 - 2\boldsymbol{a}_2'\boldsymbol{c}_2 + 2m(\boldsymbol{a}_1'\mathbf{1}_{N_1} + \boldsymbol{a}_2'\mathbf{1}_{N_2} - 1) \tag{9}$$

where $\boldsymbol{\Sigma}_{ii} = \mathrm{cov}(\tilde{\boldsymbol{Z}}_i)$, $\boldsymbol{\Sigma}_{ij} = \mathrm{cov}(\tilde{\boldsymbol{Z}}_i, \tilde{\boldsymbol{Z}}_j)$, and $\boldsymbol{c}_i = \mathrm{cov}(\tilde{\boldsymbol{Z}}_i, Y(\boldsymbol{s}))$ are the covariance functions. $\mathbf{1}_{N_i}$ is a vector with all entries one and a length N_i, and m denotes the Lagrange multiplier. The last term of L accounts for the unbiasedness constraint. By differentiating L with respect to $\boldsymbol{a}_1, \boldsymbol{a}_2, m$ and assigning the results to zero, we get in the following system of equations:

$$\begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \mathbf{1}_{N_1} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \mathbf{1}_{N_2} \\ \mathbf{1}_{N_1}' & \mathbf{1}_{N_2}' & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ m \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ 1 \end{bmatrix} \tag{10}$$

and hence

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ m \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \mathbf{1}_{N_1} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \mathbf{1}_{N_2} \\ \mathbf{1}_{N_1}' & \mathbf{1}_{N_2}' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ 1 \end{bmatrix} \tag{11}$$

There are several important discussion points for the solution in Eq. (11). The covariance matrices $\boldsymbol{\Sigma}_{ij}$ should be determined without assuming that the underlying process is isotropic or stationary. This is important for atmospheric parameters, particularly the atmospheric water vapor that shows spatial anisotropy as observed from the spatial autocorrelation function in Fig. 5. The covariance function $\boldsymbol{c}_i$ should account for the change of the support between the input and the output data. For massive data sets, the size of the covariance matrix is huge and the solution in Eq. (11) is not feasible anymore. Also, the covariance matrices should be modeled such that they would allow data prediction to any level of aggregation. The Fixed-rank kriging covariance model suggested by Cressie and Johannesson (2008) provides a comprehensive solution for these problems for single data sets and the generalized model for fusing multiple data sets was presented by Nguyen (2009) and Braverman et al. (2009). In the next section, we describe the Fixed-rank kriging method and the associated covariance model. Then, we describe how the data fusion approach is applied to our data sets.

4.3 Fixed-rank kriging

The Fixed-rank kriging (FRK) approach splits the spatial process into two or three components depending on the spatial wavelength, i.e,

$$Y(\boldsymbol{s}) = \underbrace{\mathbf{T}(\boldsymbol{s}) \cdot \boldsymbol{\alpha}}_{\text{linear trend}} + \underbrace{\mathbf{S}(\boldsymbol{s}) \cdot \boldsymbol{\eta} + \zeta(\boldsymbol{s})}_{\nu(\boldsymbol{s})} \tag{12}$$

The model in Eq. (12) is called the spatial random effects (SRE) model (Cressie and Johannesson, 2008). The first component represents a deterministic linear trend that reflects the large scale spatial variations. The second component $\mathbf{S}(\boldsymbol{s}) \cdot \boldsymbol{\eta}$ captures the relatively smooth spatial variations, which form the covariance structure of the process. That is, $\text{cov}(\mathbf{S}(\boldsymbol{u}) \cdot \boldsymbol{\eta}, \mathbf{S}(\boldsymbol{v}) \cdot \boldsymbol{\eta}) = \mathbf{S}(\boldsymbol{u})\,\mathbf{K}\,\mathbf{S}'(\boldsymbol{v})$ with $\mathbf{K}$ the covariance function of $\boldsymbol{\eta}$. This component is modeled by a linear combination of spatial random effects at multiple spatial scales. The vector $\boldsymbol{\eta}$ contains r hidden spatial random effects, which are estimated from the data at predefined nodes. Therefore, we should be able to estimate $\boldsymbol{\eta}$ regardless of the aggregation level of the input data. When neglecting the last term in Eq. (12), the weighted sum $\sum_{j=1}^{r} S_j(\boldsymbol{s})\eta_j$ should give the detrended value of Y at the location $\boldsymbol{s}$.

The weights stored in the matrix $\mathbf{S}$ for each location $\boldsymbol{s}$ depend on the distance between $\boldsymbol{s}$ and each node. The weighting function $S(\boldsymbol{s})$ has the following form:

$$S(\boldsymbol{s}) = \begin{cases} \left[1 - (||\boldsymbol{s} - \boldsymbol{m}_i||/r_i)^2\right]^2, & \text{for } ||\boldsymbol{s} - \boldsymbol{m}_i|| \leq r_i, \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$\boldsymbol{m}_i$ is the node location, and r_i is a predefined effective radius. The formula in Eq. (13) represents a bisquare bell-shaped function that has its maximum value at $\boldsymbol{m}_i$ and decreases smoothly until it reaches zero outside the circle. To demonstrate, a schematic diagram for the nodes setup is shown in Fig. 6. Within the domain of the data, four nodes, $\boldsymbol{m}_1, \cdots, \boldsymbol{m}_4$, are defined with a corresponding radius. In Fig. 6, if $\boldsymbol{s}$ is located within the radius of a certain node, it gets a positive weight, otherwise the weight is zero. Hence, $\mathbf{S}(\boldsymbol{s}) = [0, 0, 0, S(\boldsymbol{s})]$.

The last component in Eq. (12) accounts for the variations of the process that has not been captured so far (Kang and Cressie, 2011). The component ζ is assumed to be an uncorrelated Gaussian process with a mean zero and a variance σ_ζ^2.

Based on the model in Eq. (12), the FRK estimator is found when $\boldsymbol{\eta}$ and ζ are determined, i.e.,

$$\begin{aligned}\hat{Y}(\boldsymbol{s}_o) &= \boldsymbol{S}_p(\boldsymbol{s}_o)\cdot\hat{\boldsymbol{\eta}} + \hat{\zeta}(\boldsymbol{s}_o) \\ &= \boldsymbol{S}_p(\boldsymbol{s}_o)\mathbf{K}\mathbf{S}'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{Z}} + \boldsymbol{\sigma}_\zeta^2\mathbf{E}(\boldsymbol{s}_o = \boldsymbol{s})\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{Z}}\end{aligned} \tag{14}$$

where $\boldsymbol{S}_p(\boldsymbol{s}_o)$ is the weighting matrix for the prediction location and $\boldsymbol{\Sigma}$ is the covariance matrix of the input data. The matrix $\mathbf{E}$ in Eq. (14) has a value of one if $\boldsymbol{s} = \boldsymbol{s}_o$ and zero elsewhere. $\hat{Y}$ represents the detrended estimator. $\hat{\boldsymbol{\eta}}$ and $\hat{\zeta}$ are the optimal a postriori estimates of $\boldsymbol{\eta}$ and ζ, respectively (Braverman et al., 2011). In order to get the total value of $\hat{Y}_t$, we calculate

$$\hat{Y}_t(\boldsymbol{s}_o) = \boldsymbol{T}(\boldsymbol{s}_o)\cdot\hat{\boldsymbol{\alpha}} + \hat{Y}(\boldsymbol{s}_o) \tag{15}$$

The steps followed to obtain the predictions based on the FRK method are summarized in Fig. 7. The methods to estimate the noise variance σ_ϵ^2, the covariance matrix $\mathbf{K}$, and the variance of the fine-scale signal σ_ζ^2 are shown in Appendix A.

We classify the spatial variations of the atmospheric water vapor signal into three components: long wavelength, medium to short wavelength, and uncorrelated fine scale. Therefore, we split the water vapor signal using the linear model in Eq. (12) and use the FRK method for prediction.

We applied the OK and FRK to estimate the zenith-directed wet delay derived from remote sensing data. For the FRK, the matrix $\mathbf{S}$ is constructed using the node setup shown in Fig. 8. The nodes or center locations of 93 basis functions are established at three spatial resolutions, the first resolution is 40 km, the second resolution is 20 km, and the third resolution is 10 km. The semivariogram and the fitted spherical variogram model are shown in Fig. 9 (a), while the covariance matrix determined using the FRK method in shown in Fig. 9 (b). The predicted maps with 3 km $\times$ 3 km resolution are shown in Fig. 10. Due to the lack of ground truth data

that should be used to estimate the bias in the model data, we do not add the long-wavelength component into the figures to enable unbiased comparison. We observe similar results from both ordinary kriging and fixed-rank kriging that agree with the original WRF map. The spatial correlation coefficients with the corresponding WRF data are approximately 85 and 83 % for FRK and OK, respectively. For using OK, we assumed the signal spatially isotropic to ease the computations; therefore, the OK prediction map shows results sightly different from the FRK. The most impressive point here is the computational time reported for both algorithms. The FRK algorithm is fast, so that it requires significantly shorter time to produce the predictions. Most of the time is invested in the calculations of the covariance model parameters and constructing the matrices $\mathbf{S}$ and $\boldsymbol{\Sigma}$. We implemented the OK algorithm such that the predictions are found iteratively. Also, to estimate a value at the location $\boldsymbol{s}$, we do not use the entire data, but only those which exist within a predefined radius around the prediction location. Nevertheless, the OK algorithm requires computational time with an order of magnitude higher than that required by the FRK method, on the same machine.

In the next section, we describe the extension of the FRK method for predicting the atmospheric PWV by fusing remote sensing data and the WRF model.

5 Data fusion for water vapor estimation

In this section, we fuse the PWV maps derived from the remote sensing data and WRF model. Since we classify the spatial variations of the atmospheric water vapor signal into long wavelength, medium to short wavelength, and uncorrelated fine scale components, we use the following model setup for prediction.

5.1 Model setup

PWV maps will be derived from the remote sensing data, denoted $\boldsymbol{Z}_1$, and those from the WRF model, denoted $\boldsymbol{Z}_2$ with the sizes N_1 and N_2, respectively. $\boldsymbol{Z}_1$ contains the point PWV estimates from remote sensing data and $\boldsymbol{Z}_2$ contains the block WRF data. Following the SME

model in Eq. (12), the two data sets can be expressed as

$$\begin{bmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{T}_1 \\ \overline{\mathbf{T}}_2 \end{bmatrix} \boldsymbol{\alpha} + \begin{bmatrix} \mathbf{S}_1 \\ \overline{\mathbf{S}}_2 \end{bmatrix} \boldsymbol{\eta} + \begin{bmatrix} \boldsymbol{\zeta}_1 \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \end{bmatrix} \tag{16}$$

The regression coefficient $\boldsymbol{\alpha}$ should be estimated jointly from both data sets. However, we do not have apriori information about the biases; therefore, we estimate $\boldsymbol{\alpha}$ in this contribution independently for each data set. The matrices $\mathbf{S}_1$ and $\overline{\mathbf{S}}_2$ contain the weights of each location for each data set. To distinguish between point and block data, we used the notation $\overline{\mathbf{S}}_2$ for block-level data. The model components for point and block data are given in Table 1. The WRF data are available at a resolution of $3\,\text{km} \times 3\,\text{km}$; therefore, the highly variable signal of water vapor is smoothed. Hence, we do not add the component $\boldsymbol{\zeta}$ for the model data.

To solve the system in Eq. (11), we determine the covariance structure associated with each SRE model in Eq. (16), i.e.,

$$\boldsymbol{\Sigma}_{11} = \text{var}(\tilde{\boldsymbol{Z}}_1) = \mathbf{S}_1 \mathbf{K} \mathbf{S}_1' + \sigma_\zeta^2 \mathbf{V}_\zeta + \sigma_{\epsilon_1}^2 \mathbf{V}_{\epsilon_1} \tag{17}$$

$$\boldsymbol{\Sigma}_{22} = \text{var}(\tilde{\boldsymbol{Z}}_2) = \tilde{\mathbf{S}}_2 \mathbf{K} \tilde{\mathbf{S}}_2' + \sigma_{\epsilon_2}^2 \mathbf{V}_{\epsilon_2} \tag{18}$$

$$\boldsymbol{\Sigma}_{12} = \text{cov}(\tilde{\boldsymbol{Z}}_1, \tilde{\boldsymbol{Z}}_2) = \mathbf{S}_1 \mathbf{K} \tilde{\mathbf{S}}_2' = \boldsymbol{\Sigma}_{21}' \tag{19}$$

where $\sigma_\zeta^2 \mathbf{V}_\zeta$ and $\sigma_\epsilon^2 \mathbf{V}_\epsilon$ are diagonal covariance matrices for ζ and ϵ, respectively. Note that when computing the cross covariance functions $\boldsymbol{\Sigma}_{12}$ and $\boldsymbol{\Sigma}_{21}$, the only part of the signals that is assumed correlated is $\boldsymbol{\eta}$. In order to solve Eq. (11), we need not only to specify the covariance matrices of the input data, but also to find the covariance between the observations and the spatial process at the prediction locations. The covariance terms are obtained from:

$$\begin{aligned} \boldsymbol{c}_1 &= \text{cov}(\tilde{\boldsymbol{Z}}_1(\boldsymbol{s}), Y(\boldsymbol{s}_o)) \\ &= \boldsymbol{S}_p(\boldsymbol{s}_o) \mathbf{K} \mathbf{S}_1'(\boldsymbol{s}) + \sigma_\zeta^2 \mathbf{E}(\boldsymbol{s} = \boldsymbol{s}_o) \end{aligned} \tag{20}$$

$$\boldsymbol{c}_2 = \text{cov}(\tilde{\boldsymbol{Z}}_2, Y(\boldsymbol{s}_o)) = \boldsymbol{S}_p(\boldsymbol{s}_o) \mathbf{K} \tilde{\mathbf{S}}_2' \tag{21}$$

The matrix $\mathbf{E}$ in Eq. (20) has a value of one if $\boldsymbol{s} = \boldsymbol{s}_o$ and zero elsewhere. By solving for $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ in Eq. (11) and substituting the results in Eq. (8), the estimator $\hat{Y}(\boldsymbol{s}_o)$ becomes

$$\hat{Y}(\boldsymbol{s}_o) = \left(\boldsymbol{S}_p(\boldsymbol{s}_o)\mathbf{K} \begin{bmatrix} \mathbf{S}_1' \\ \tilde{\mathbf{S}}_2' \end{bmatrix} + \begin{bmatrix} \sigma_\zeta^2 \mathbf{E} \\ 0 \end{bmatrix} \right) \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\boldsymbol{Z}}_1 \\ \tilde{\boldsymbol{Z}}_2 \end{bmatrix} \tag{22}$$

The mean squared prediction error (MSPE) corresponding to $\hat{Y}$ can be obtained from

$$\text{MSPE} = \boldsymbol{a}_1' \boldsymbol{\Sigma}_{11} \boldsymbol{a}_1 + \boldsymbol{a}_2' \boldsymbol{\Sigma}_{22} \boldsymbol{a}_2 + 2\boldsymbol{a}_1' \boldsymbol{\Sigma}_{12} \boldsymbol{a}_2 - 2\boldsymbol{a}_1' \boldsymbol{c}_1 - 2\boldsymbol{a}_2' \boldsymbol{c}_2 \tag{23}$$

Using the FRK covariance model in Eq. (19) makes the matrix inversion of Eq. (22) scalable. That is, the matrix inversion can be achieved by applying a recursive block-wise inversion as follows:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{O}_1 & \mathbf{O}_2 \\ \mathbf{O}_3 & \mathbf{O}_4 \end{bmatrix} \tag{24}$$

where

$$\begin{aligned} \mathbf{O}_1 &= \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} \\ \mathbf{O}_2 &= -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ \mathbf{O}_3 &= -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} \\ \mathbf{O}_4 &= (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{aligned}$$

and $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are matrices of any size, and $\mathbf{A}, \mathbf{D}$ must be square. The inversion of individual matrices in Eq. (24) is achieved by applying the formula of Sherman–Morrison–Woodbury, which is made possible due to the FRK covariance structure,

$$\begin{aligned} \boldsymbol{\Sigma}_{ii}^{-1} &= (\mathbf{D}_i + \mathbf{S}_i \mathbf{K}_i \mathbf{S}_i')^{-1} \\ &= \mathbf{D}_i^{-1} - \mathbf{D}_i^{-1}\mathbf{S}_i(\mathbf{K}^{-1} + \mathbf{S}_i'\mathbf{D}_i^{-1}\mathbf{S}_i)^{-1}\mathbf{S}_i'\mathbf{D}_i^{-1} \end{aligned} \tag{25}$$

The computations require the inversion of the matrices $\mathbf{K}$ and $(\mathbf{K}^{-1}+\mathbf{S}_i'\mathbf{D}_i^{-1}\mathbf{S}_i)$, where each of them has the size $r \times r$ with r significantly smaller than the data size. Note that $\mathbf{D}_i$ is a diagonal matrix, for which the inversion is achieved by inverting the diagonal elements. Using the FRK covariance model makes the computational burden for the matrix inversion linear with the data size (Cressie and Johannesson, 2008).

5.2 Application to the data

In this section, we build PWV maps from remote sensing and WRF model data using a spatial statistical data fusion method. The first PWV map, derived by combining GNSS and PSI, has 169 688 data points. The WRF model provides a block-level map of 1296 cells of the size 3 km $\times$ 3 km. The data to be fused have different qualities, huge size, different spatial support, and gaps in the remote sensing data. The output grid is defined at 3 km $\times$ 3 km (block-level support) and MERIS PWV maps are used for evaluation.

Following the work flow in Fig. 7, we first estimate the long wavelength trends and remove them from the data using Eq. (7). By comparing the PWV from the WRF model and remote sensing data, we found it is most likely that the model data have a bias. Due to the lack of apriori information about the bias and the absence of accurate ground truth data to estimate it, we estimated $\boldsymbol{\alpha}$ independently for each data set. The centered maps are shown in Fig. 11.

Second, the matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ are constructed for the first data set (remote sensing data) and the second data set (model data). The node setup is shown in Fig. 8. The number of nodes must be the same for both data sets and they are selected such that $\mathbf{S}$ does not contain columns of zeros, otherwise the corresponding node has to be removed. If PWV data are available at point-level, a weighting value is calculated for each point with respect to all nodes. However, the WRF model simulates data at block-level, hence we superimpose the model grid with a lattice of regular points such that each cell in the WRF grid contains 9 points. A weighting value is calculated for each point, these values are averaged to get one weighing value for each WRF cell to form the matrix $\overline{\mathbf{S}}_2$. Building the matrix $\boldsymbol{S}_p$ for the prediction locations is done in a similar way, either at point-level or block-level, depending on the output grid.

In the third step, the covariance parameters $(\mathbf{K}, \sigma_\zeta^2, \sigma_\epsilon^2)$ are estimated from the centered data $\tilde{\boldsymbol{Z}}_1$ and $\tilde{\boldsymbol{Z}}_2$. The error variances for both data sets, $\mathbf{K}$ and σ_ζ^2 are estimated as described in Appendix A. Note that when the two data sets are combined to infer a single process, i.e., PWV, one $\mathbf{K}$ is estimated for all data sets.

Results

So far, all components required to produce the predictions using Eq. (22) are obtained. In Fig. 12, we show the prediction maps obtained by applying FRK to individual data sets as well as the map obtained by data fusion. The figure also shows the MSPE maps associated with each prediction map. We compare the interpolations obtained by applying FRK to single data sets with those obtained by SSDF and we compare both with the MERIS data. The results show that the map obtained by data fusion correlates more consistently with the map predicted only from PSI + GNSS (Table 2). In the PWV map generated by WRF, shown in Fig. 11, the area in the lower left corner shows artifacts that do not reflect the correct values of PWV as observed from the MERIS PWV map, Fig. 3c and d. Applying FRK to the WRF data does not remove these artifacts from the prediction map. However, in the map obtained by the fusion of both data sets, the artifacts in the lower corner disappeared, but the corresponding MSPE values are large for this region. The MSPE values corresponding to the SSDF predictions are generally smaller, and we should note that in the regions of sparse observations, the corresponding MSPE values tend to increase. For regions of sparse observations in the PWV map (Fig. 11), i.e., the areas in the west of the Rhine valley or in the lower right corner, the map from the WRF model contributes to improve the estimation of the PWV values in the prediction map. The region in the lower right corner has a higher topography and the wet delay values are expected to decrease as we observe from the map of WRF. In the prediction map obtained by applying FRK to PWV from PSI and GNSS, the predicted values tend to increase since the data in this area are sparse and partially biased. By applying the SSDF approach, the data available from WRF influence the predictions such that the PWV values in this area are more reasonable and they decrease by moving to the lower right corner. In a similar way, the data from WRF improve the predictions

in the region around (7.8° E, 49.25° N), where only sparse PWV data exist. The data from the model, however, affects the prediction in the lower left corner such that they are smaller than those observed in the MERIS map.

In addition, we show the PWV profiles over the line drawn horizontally at the latitude 49.37° N in Fig. 12 (h). It is observed from the plots that the predictions made by data fusion are affected more by the data from WRF in region A, where the remote sensing data are sparse. However, in region B, the WRF data are significantly overestimated. In the prediction map made by data fusion, these data have a lower effect in than those received from the remote sensing data. The map received by applying the data fusion shows the best spatial correlation with the data from MERIS and the smallest RMS value (see Table 2).

In the above example, the data from remote sensing have a more significant influence on the output. In Fig. 13, we show another example where the model highly affects the predicted map. The predicted map based on model data shows a better spatial correlation and a lower uncertainty values compared to the map predicted using remote sensing data. In this case, the fusion map is more affected by the model data. The spatial correlation coefficients and the values of uncertainty are given in Table 2. In the first example (Fig. 12), the effect of the remote sensing data on the prediction map is significant. The other example in Fig. 13 and Table 2 show that the model has a larger effect on the output map.

6 Conclusions and outlook

We presented a method to obtain the atmospheric PWV over any aggregation level by the fusion of remote sensing data and atmospheric models. The PWV maps derived by combining data from PSI and GNSS are available at discrete points that are absent in regions of low coherence. On the other hand, the WRF model provides simulations of PWV in the atmosphere on regular grids at a coarse spatial resolution. Both the quality of the model data, and the model skills for representing meso-scale atmospheric structures should be improved. The quality of the prediction maps should be improved by data fusion. For data fusion, the method of spatial statistical data fusion, first presented in (Nguyen, 2009), was employed. This method is based on the

fixed-rank kriging approach that attempts to solve the problems of computational complexity of huge data sets, change of support, and bias. We inferred PWV data on a grid of 3 km × 3 km and compared the results with PWV maps inferred from MERIS data on the same grid. The results show a strong correlation between data fusion maps and those maps from MERIS. The difference between both maps shows uncertainty values of less than 1 mm, which is lower than that obtained from inferring data based on single sets.

To further improve the results, we suggest the following. The matrix $\mathbf{S}_i$ has so far been constructed for each data source by defining a set of spatial nodes. The number of the nodes is empirically adjusted such that the covariance function computed for the data set based on the estimated matrix $\mathbf{K}$ approximates the empirical covariance. In future work, the size and the locations of nodes have to be optimized by minimizing the difference between the empirical and the estimated covariance functions. We should also estimate the biases for each data set (if exist), so that they can be accounted for in the fusion approach. The data fusion approach can be extended such that more than two data sets are used, for example, by including the MERIS maps in the fusion. With the increasing number of satellite missions and improved atmospheric models, we are able to produce complete, accurate information about the Earth's atmosphere based on data fusion approaches. Moreover, the improved PWV maps can be iteratively assimilated to the local area atmospheric model to generate more accurate 3-dimensional water vapor fields. Also, testing other combinations of physical schemes within the WRF model can further improve the resulting water vapor maps. In this paper, we compared the prediction maps with the data from MERIS; however, in future work, the results should be validated using Bootstrapping, or Jackknifing techniques.

Appendix A: Estimation covariance parameters

Predicting the stochastic component of the atmospheric signal using kriging requires obtaining the covariance function $\mathbf{\Sigma}$ and fitting a covariance model. Using the FRK covariance model, we need to estimate the matrix $\mathbf{K}$, the noise variance σ_ϵ^2, and the variance of the fine-scale signal σ_ζ^2. The first method proposed to estimate $\mathbf{K}$ is called binned method-of-moments (MM)

(Cressie and Johannesson, 2008; Nguyen, 2009). This approach derives the empirical estimator for $\mathbf{\Sigma}$ and obtains $\mathbf{K}$ such that $||\hat{\mathbf{\Sigma}} - \mathbf{\Sigma}||_F$ is minimum, where $||\cdot||_F$ refers to the Frobenius norm.

Another approach proposed by Katzfuss and Cressie (2009) targets to determine the covariance parameters using the algorithm of maximum likelihood estimation (MLE). Furthermore, they estimated the covariance paramters using the expectation-maximization (E-M) algorithm (Dempster et al., 1977) to reduce the computational burden. This algorithm provides estimates not only of $\mathbf{K}$ but also of σ_ζ^2, where the solution for the MLEs is found iteratively. Within each iteration the algorithm performs two steps, the expectation and maximization. In the following, we present a description of how to obtain the maximum likelihood estimates of the covariance model parameters via the E-M algorithm.

Assuming that the observations in $\tilde{\boldsymbol{Z}}$ follow a multivariate Gaussian distribution, that is $\tilde{\boldsymbol{Z}} \sim N(0, \mathbf{\Sigma})$. Let the parameters of interest $\mathbf{K}$ and σ_ζ^2 be summarized in the vector $\mathbf{\Theta}$, then the likelihood function $L(\mathbf{\Theta})$ (Katzfuss and Cressie, 2009)

$$\begin{aligned} -2\log L(\mathbf{\Theta}) &= -2f(\tilde{\boldsymbol{Z}}; \mathbf{\Theta}) \\ &= \log\ \det \mathbf{\Sigma} + \tilde{\boldsymbol{Z}}' \mathbf{\Sigma}^{-1} \tilde{\boldsymbol{Z}} + c \\ &= \log\ \det \mathbf{\Sigma} + \ \mathrm{tr}\ (\Sigma^{-1} \tilde{\boldsymbol{Z}} \tilde{\boldsymbol{Z}}') + c \end{aligned} \tag{A1}$$

where $c = (N/2)\log 2\pi$ is a constant independent of $\mathbf{\Theta}$ and hence it cancels out in the maximization step. $\mathrm{tr}(\cdot)$ denotes the trace operator of a square matrix, with $\mathrm{tr}\ (\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$.

In the expectation step of the algorithm, we calculate

$$Q(\mathbf{\Theta}; \mathbf{\Theta}^{[t]}) = \mathrm{E}_{\mathbf{\Theta}^{[t]}}\{-2\log L(\boldsymbol{\eta}, \boldsymbol{\zeta}; \mathbf{\Theta}) | \tilde{\boldsymbol{Z}}\} \tag{A2}$$

given that:

$$\begin{aligned} -2\log L(\boldsymbol{\eta}, \boldsymbol{\zeta}; \mathbf{\Theta}) = \log\ \det \mathbf{K} + \ \mathrm{tr}\ (\mathbf{K}^{-1} \boldsymbol{\eta}\boldsymbol{\eta}') + N\log\sigma_\zeta^2 \\ + \sigma_\zeta^{-2}\ \mathrm{tr}\ (\boldsymbol{\zeta}\boldsymbol{\zeta}') + N\log\sigma_\epsilon^2 + \sigma_\epsilon^{-2}\ \mathrm{tr}\ (\boldsymbol{\epsilon}\boldsymbol{\epsilon}') \end{aligned}$$

Then Eq. (A2) becomes

$$Q(\boldsymbol{\Theta};\boldsymbol{\Theta}^{[t]}) = -\frac{1}{2}\Big[\log\,\det\mathbf{K} + \,\mathrm{tr}\,(\mathbf{K}^{-1}\mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\eta}\boldsymbol{\eta}'|\tilde{\boldsymbol{Z}}\}) + N\log\sigma_\zeta^2 + \sigma_\zeta^{-2}\,\mathrm{tr}\,(\mathbf{V}_\zeta^{-1}\mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\zeta}\boldsymbol{\zeta}'|\tilde{\boldsymbol{Z}}\}) + N\log\sigma_\epsilon^2 + \sigma_\epsilon^{-2}\,\mathrm{tr}\,(\mathbf{V}_\epsilon^{-1}\mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'|\tilde{\boldsymbol{Z}}\})\Big] \tag{A3}$$

We should remind the reader that the parameters to be estimated here are $\mathbf{K}$ and σ_ζ^2, while σ_ϵ^2 is estimated from the robust semivariogram, as described later. To proceed with the solution, it is required to quantify the conditional expectations in Eq. (A3). Using the standard formula required for calculating conditional expectations for multivariate normal distribution, the expectations will have the following form (Katzfuss and Cressie, 2009)

$$\mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\eta}\boldsymbol{\eta}'|\tilde{\boldsymbol{Z}}\} = \boldsymbol{\Sigma}_\eta^{[t]} + \boldsymbol{\mu}_\eta^{[t]}\boldsymbol{\mu}_\eta^{\prime[t]}$$
$$\mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\zeta}\boldsymbol{\zeta}'|\tilde{\boldsymbol{Z}}\} = \boldsymbol{\Sigma}_\zeta^{[t]} + \boldsymbol{\mu}_\zeta^{[t]}\boldsymbol{\mu}_\zeta^{\prime[t]}$$

with

$$\boldsymbol{\mu_\eta}^{[t]} = \mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\eta}|\tilde{\boldsymbol{Z}}\} = \mathbf{K}^{[t]}\mathbf{S}'{\boldsymbol{\Sigma}^{t}}^{-1}\tilde{\boldsymbol{Z}}$$
$$\boldsymbol{\mu_\zeta}^{[t]} = \mathrm{E}_{\boldsymbol{\Theta}^{[t]}}\{\boldsymbol{\zeta}|\tilde{\boldsymbol{Z}}\} = \sigma_\zeta^{2[t]}\mathbf{V}_\zeta{\boldsymbol{\Sigma}^{[t]}}^{-1}\tilde{\boldsymbol{Z}}$$
$$\boldsymbol{\Sigma}_\eta^{[t]} = \mathrm{cov}_{\boldsymbol{\Theta}^{[t]}}(\boldsymbol{\eta}|\tilde{\boldsymbol{Z}}) = \mathbf{K}^{[t]} - \mathbf{K}^{[t]}\mathbf{S}'{\boldsymbol{\Sigma}^{[t]}}^{-1}\mathbf{S}\mathbf{K}^{[t]}$$
$$\boldsymbol{\Sigma}_\zeta^{[t]} = \mathrm{cov}_{\boldsymbol{\Theta}^{[t]}}(\boldsymbol{\zeta}|\tilde{\boldsymbol{Z}}) = \sigma_\zeta^{2[t]}\mathbf{V}_\zeta - \sigma_\zeta^{2[t]}\mathbf{V}_\zeta{\boldsymbol{\Sigma}^{[t]}}^{-1}\sigma_\zeta^{2[t]}\mathbf{V}_\zeta$$

After the expectation step, we perform a maximization step. The parameters $\mathbf{K}$ and σ_ζ^2 in Eq. (A3) should be selected such that $Q(\cdot)$ is maximized. The partial derivative is taken with respect to both parameters and the result is assigned to zero. Finding the derivative here is rather simple since $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$ do not show dependency on each other, as observed from Eq. (A3). The

updating scheme of the E-M algorithm in each iteration is

$$\mathbf{K}^{[t+1]} = \mathbf{K}^{[t]} + \mathbf{K}^{[t]} \left(\mathbf{S}' {\boldsymbol{\Sigma}^{[t]}}^{-1} \left(\tilde{\boldsymbol{Z}} \tilde{\boldsymbol{Z}}' {\boldsymbol{\Sigma}^{[t]}}^{-1} - \mathbf{I}_N \right) \mathbf{S} \right) \mathbf{K}^{[t]} \tag{A4}$$

$$\sigma_\zeta^{2[t+1]} = \sigma_\zeta^{2[t]} + \sigma_\zeta^{2[t]} \text{ tr } \left(\frac{1}{N} {\boldsymbol{\Sigma}^{[t]}}^{-1} \left(\tilde{\boldsymbol{Z}} \tilde{\boldsymbol{Z}}' {\boldsymbol{\Sigma}^{[t]}}^{-1} - \mathbf{I}_N \right) \mathbf{V}_\zeta \right) \sigma_\zeta^{2[t]} \tag{A5}$$

We keep updating the solution until the algorithm converges. One criterion to monitor convergence is to calculate the norm of the difference between the current and last update of the vector $\boldsymbol{\Theta}$ (which is of size $r^2 + 1$). That means $||\boldsymbol{\Theta}^{[t+1]} - \boldsymbol{\Theta}^{[t]}|| < b$ should hold for small enough and positive value of b. Following Katzfuss and Cressie (2009), b is assigned a value of $10^{-6} r^2$. The starting choice of $\mathbf{K}$ and σ_ζ^2 should be valid; strictly speaking, $\mathbf{K}^{[0]}$ must be symmetric and positive-definite and $\sigma_\zeta^{2[0]}$ must be positive, i.e., $\mathbf{K}^{[0]} = 0.9 \cdot \text{var}(\tilde{\boldsymbol{Z}}) \mathbf{I}_r$ and $\sigma_\zeta^{2[0]} = 0.1 \cdot \text{var}(\tilde{\boldsymbol{Z}})$.

The measurement error variance σ_ϵ^2 is estimated separately from the empirical semivariogram of the data. Estimating both σ_ϵ^2 and σ_ζ^2 from the data is not a trivial task. That is because the nugget effect in the semivariogram reflects not only the error variance but may be affected by the fine-scale variance. Therefore, having information about the error distribution and variance is worthwhile. In our case we estimate σ_ϵ^2 using the method of robust semivariogram (Cressie, 1993),

$$2\gamma(h) = \frac{\left(\frac{1}{|N(h)|} \sum_{N(h)} \left| Z(\boldsymbol{u}_i) - Z(\boldsymbol{u}_j) \right|^{1/2} \right)^4}{\left(0.457 + \frac{0.494}{|N(h)|} \right)} \tag{A6}$$

where h is separation distance, assuming the signal is spatially isotropic. To obtain an estimate of σ_ϵ^2, a straight line is fitted to the estimated semivariogram at short h. Since the slope of the structure function (variogram) describing atmospheric turbulence is expected to vary with h, we made the line fitting based on the estimates of the first 3 km (empirically defined). Let the line fit be $\hat{\gamma}(h) = \hat{\gamma}(0+) + bh$, then the estimate of σ_ϵ^2 is

$$\hat{\sigma}_\epsilon^2 = \hat{\gamma}(0+) \tag{A7}$$

Should $\hat{\gamma}(0+)$ have a negative value, $\hat{\sigma}_\epsilon^2$ is set to zero.

The estimate of $\mathbf{K}$ using the detrended PWV maps estimated from the PSI + GNSS and model data on 5 September 2005 is shown in Fig. 14. The corresponding covariance function is also shown. The matrix $\mathbf{S}$ is constructed as described in Sect. 4.3 using the nodes setup in Fig. 8. The $\mathbf{K}_{EM}$ has a maximum value for the element (29,29), which is equivalent to estimate at the node in the lower right corner at the location (8.524° E, 48.69° N), see Fig. 8. This can be explained by the sparseness of PWV estimates close to this node and the PWV values from PSI and GNSS are significantly higher than those from the model. The covariance matrix is computed for the observations binned into 7 km × 7 km blocks to demonstrate covariance structure. We observe from the covariance matrices that the variances, on the main diagonal, increase in areas of sparse observations. The reader should note that the observations do not exist on a regular grid (due to the spatial distribution of PS points); hence, the covariance values in the off-diagonal cells can be negative and then again positive.

Acknowledgements. The authors would like to thank the GNSS data providers: RENAG, RGP, Teria, and Orpheon (France), SA*POS*®-Baden-Württemberg and Rheinland-Pfalz (Germany), European Permanent Network, and IGS. We thank the ESA for the ASAR and MERIS data. We also acknowledge the ECWMF for providing the ERA-INTERIM reanalysis data.

The service charges for this open access publication have been covered by a Research Centre of the Helmholtz Association.

References

Alshawaf, F., Fersch, B., Hinz, S., Kunstmann, H., Mayer, M., Thiele, A., Westerhaus, M., and Meyer, F.: Analysis of atmospheric signals in spaceborne InSAR – toward water vapor mapping based on multiple sources, in: Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, 1960–1963, 2012.

Alshawaf, F., Fuhrmann, T., Knopfler, A., Luo, X., Mayer, M., Hinz, S., and Heck, B.: Accurate Estimation of Atmospheric Water Vapor Using GNSS Observations and Surface Me-

teorological Data, Geoscience and Remote Sensing, IEEE Transactions on, 53, 3764–3771, doi:10.1109/TGRS.2014.2382713, 2015a.

Alshawaf, F., Hinz, S., Mayer, M., and Meyer, F. J.: Constructing accurate maps of atmospheric water vapor by combining interferometric synthetic aperture radar and GNSS observations, Journal of Geophysical Research: Atmospheres, 120, 1391–1403, 2015b.

Awan, N. K., Truhetz, H., and Gobiet, A.: Parametrization-Induced Error Characteristics of MM5 and WRF Operated in Climate Mode over the Alpine Region: An Ensemble-Based Analysis., Journal of Climate, 24, 3107–3123, 2011.

Bender, M., Dick, G., Wickert, J., Schmidt, T., Song, S., Gendt, G., Ge, M., and Rothacher, M.: Validation of GPS slant delays using water vapor radiometers and weather models, Meteorol. Z., 17, 807–812, 2008.

Bennitt, G., and Rupp, A.: Operational assimilation of GPS zenith total delay observations into the Met Office numerical weather prediction models, Monthly Weather Review, 140, 2706–2719, 2012.

Berg, P., Wagner, S., Kunstmann, H., and Schädler, G.: High resolution regional climate model simulations for Germany: part I – validation, Clim. Dynam., 40, 401–414, 2013.

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H.: GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system, J. Geophys. Res.-Atmos., 97, 15787–15801, 1992.

Braverman, A., Nguyen, H., and Olsen, E.: Space-Time Data Fusion, Earth Science, Earth Science Technology Forum, Pasadena, California, 2011.

Braverman, A. J., Cressie, N., Katzfuss, M., Michalak, A. M., Miller, C. E., Nguyen, H., Olsen, E. T., and Wang, R.: Geostatistical data fusion for remote sensing applications, AGU Fall Meeting Abstracts, p. C1014, 2009.

Chen, F. and Dudhia, J.: Coupling an advanced land surface-hydrology model with the Penn State–NCAR MM5 Modeling System. Part I: Model implementation and sensitivity, Mon. Weather Rev., 129, 569–585, 2001.

Cimini, D., Pierdicca, N., Pichelli, E., Ferretti, R., Mattioli, V., Bonafoni, S., Montopoli, M., and Perissin, D.: On the accuracy of integrated water vapor observations and the potential for mitigating electromagnetic path delay error in InSAR, Atmos. Meas. Tech., 5, 1015–1030, doi:10.5194/amt-5-1015-2012, 2012.

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Kiehl, J. T., Briegleb, B., Bitz, C., Lin, S., Zhang, M., and Dai, Y.: Description of the NCAR Community Atmo-

sphere Model (CAM 3.0), Tech. rep., Climate And Global Dynamics Division National Center For Atmospheric Research Boulder, Colorado, 2004.

Cressie, N.: Statistics for Spatial Data, Wiley, New York, 1993.

Cressie, N. and Johannesson, G.: Fixed rank kriging for very large spatial data sets, J. Roy. Stat. Soc. B Met., 70, 209–226, 2008.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Dempster, A. P., Laird, N. M., and Rubin, D. B.: Maximum likelihood from incomplete data via the EM algorithm, J. Roy. Stat. Soc. B, 39, 1–38, 1977.

Ferretti, A., Prati, C., and Rocca, F.: Permanent scatterers in SAR interferometry, IEEE T. Geosci. Remote, 39, 8–20, 2001.

Fersch, B. and Kunstmann, H.: Atmospheric and terrestrial water budgets: sensitivity and performance of configurations and global driving data for long term continental scale WRF simulations, Clim. Dynam., 42, 2367–2396, 2014.

Fersch, B., Kunstmann, H., Devaraju, B., and Sneeuw, N.: Large scale water storage variations from regional atmospheric water budgets and comparison to the GRACE spaceborne gravimetry, J. Hydrometeorol., 13, 1589–1603, 2012.

Fischer, J. and Bennartz, R.: Retrieval of Total Water Vapour Content from MERIS Measurements: Algorithm Theoretical Basis Document ; ATBD 2.4, Freie Universität Berlin, Berlin, Germany, 1997.

Gelfand, A. E., Zhu, L., and Carlin, B. P.: On the change of support problem for spatio-temporal data, Biostatistics, 2, 31–45, 2001.

Gong, W., Meyer, F., Webley, P., Morton, D., and Liu, S.: Performance analysis of atmospheric correction in InSAR data based on the Weather Research and Forecasting Model (WRF), in: Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International, 25–30 July 2010, Honolulu, Hawaii, USA, 2900–2903, 2010.

Hanssen, R.: Radar Interferometry: Data Interpretation and Error Analysis, Kluwer Academic Publishers, Dordrecht, the Netherlands, 2001.

Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, Mon. Weather Rev., 132, 103–120, 2004.

Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev., 134, 2318–2341, 2006.

Hooper, A., Segall, P., and Zebker, H.: Persistent scatterer interferometric synthetic aperture radar for crustal deformation analysis, with application to Volcán Alcedo, Galápagos, J. Geophys. Res.-Solid Earth, 112, B07407, doi:10.1029/2006JB004763, 2007.

Jade, S. and Vijayan, M.: GPS-based atmospheric precipitable water vapor estimation using meteorological parameters interpolated from NCEP global reanalysis data, J. Geophys. Res.-Atmos., 113, D03106, doi:10.1029/2007JD008758, 2008.

Kain, J. S.: The Kain–Fritsch convective parameterization: an update, J. Appl. Meteorol., 43, 170–181, 2004.

Kang, E. L. and Cressie, N.: Bayesian inference for the spatial random effects model, J. Am. Stat. Assoc., 106, 972–983, 2011.

Karabatić, A., Weber, R., and Haiden, T.: Near real-time estimation of tropospheric water vapor content from ground based GNSS data and its potential contribution to weather now-casting in Austria, Adv. Space Res., 47, 1691–1703, 2011.

Katzfuss, M. and Cressie, N.: Maximum likelihood estimation of covariance parameters in the spatial-random-effects model, in: Proceedings of Joint Statistical Meetings Alexandria, American Statistical Association, 1–6 August 2009, 3378–3390, 2009.

Luo, X., Mayer, M., and Heck, B.: Extended neutrospheric modelling for the GNSS-based determination of high-resolution atmospheric water vapor fields, Bol. Cien. Geod., 14, 149–170, 2008.

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., and Rabaute, T.: The displacement field of the Landers earthquake mapped by radar interferometry, Nature, 364, 138–142, 1993.

Mateus, P., Nico, G., Tomé, R., Catalão, J., and Miranda, P.: Comparison of precipitable water vapor (PWV) maps derived by GPS, SAR interferometry, and numerical forecasting models, SPIE, 2008, 782714–782714–7, 2010.

Meyer, F., Bamler, R., Leinweber, R., and Fischer, J.: A comparative analysis of tropospheric water vapor measurements from MERIS and SAR, in: Geoscience and Remote Sensing Symposium, IGARSS 2008, IEEE International, 6–11 July, 2008, Boston, Massachusetts, Vol. 4, IV–228–IV–231, 2008.

Nguyen, H.: Spatial Statistical Data Fusion for Remote Sensing Applications, PhD thesis, University of California, Los Angeles, USA, 2009.

Onn, F. and Zebker, H. A.: Correction for interferometric synthetic aperture radar atmospheric phase artifacts using time series of zenith wet delay observations from a GPS network, J. Geophys. Res.-Solid Earth, 111, B09102, doi:10.1029/2005JB004012, 2006.

Pichelli, E., Ferretti, R., Cimini, D., Perissin, D., Montopoli, M., Marzano, F. S., and Pierdicca, N.: Water vapor distribution at urban scale using high-resolution numerical weather model and spaceborne SAR interferometric data, Nat. Hazards Earth Syst. Sci., 10, 121–132, doi:10.5194/nhess-10-121-2010, 2010.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N. P. M., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53, 323–361, doi:10.1002/2014RG000475, 2015.

Rocken, C., Hove, T. V., Johnson, J., Solheim, F., Ware, R., Bevis, M., Chiswell, S., and Businger, S.: GPS/STORM-GPS sensing of atmospheric water vapor for meteorology, J. Atmos. Ocean. Tech., 12, 468–478, 1995.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227, 3465–3485, doi:10.1016/j.jcp.2007.01.037, 2008.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A Description of the Advance Research WRF Version 3, NCAR Technical Note NCAR/TN-468+STR, National Center for Atmospheric Research, Boulder, Colorado, USA, 2008.

Uppala, S., Dee, D., Kobayashi, S., Berrisford, P., and Simmons, A.: Towards a climate data assimilation system: status update of ERA-Interim, ECMWF Newsl., 115, 12–18, 2008.

Zebker, H. A. and Goldstein, R. M.: Topographic mapping from interferometric synthetic aperture radar observations, J. Geophys. Res.-Solid Earth, 91, 4993–4999, 1986.

Wang, W. and Seaman, N. L.: A Comparison Study of Convective Parametrization Schemes in a Mesoscale Model, Month. Weather Rev., 125, 252–278, 1997.

Zebker, H. A., Rosen, P. A., Goldstein, R. M., Gabriel, A., and Werner, C. L.: On the derivation of coseismic displacement fields using differential radar interferometry: the Landers earthquake, J. Geophys. Res.-Solid Earth, 99, 19617–19634, 1994.

Zebker, H. A., Rosen, P. A., and Hensley, S.: Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps, J. Geophys. Res.-Solid Earth, 102, 7547–7563, 1997.

Table 1. Model components from point-level and areal-level data.

	Point data	Block data
True process	$Y(\boldsymbol{s})$	$Y(B_i) = \frac{1}{\lvert B_i \rvert} \sum_{\boldsymbol{s} \subset B_i} Y(\boldsymbol{s})$
Trend	$T(\boldsymbol{s})\boldsymbol{\alpha}$	$\left(\frac{1}{\lvert B_i \rvert} \sum_{\boldsymbol{s} \subset B_i} T(\boldsymbol{s}) \right) \boldsymbol{\alpha}$
Weighting matrix	$\mathbf{S}(\boldsymbol{s})$	$\tilde{\mathbf{S}}(B_i) = \frac{1}{\lvert B_i \rvert} \sum_{\boldsymbol{s} \subset B_i} \mathbf{S}(\boldsymbol{s})$
Medium-scale signal	$\mathbf{S}(\boldsymbol{s})\boldsymbol{\eta}$	$\tilde{\mathbf{S}}(B_i)\boldsymbol{\eta}$
Fine-scale signal	$\zeta(\boldsymbol{s})$	$\zeta(B_i) = \frac{1}{\lvert B_i \rvert} \sum_{\boldsymbol{s} \subset B_i} \zeta(\boldsymbol{s})$
Error	$\epsilon(\boldsymbol{s})$	$\epsilon(B_i)$

Table 2. Spatial correlation coefficients (CC) and RMS values when comparing the prediction maps with MERIS PWV maps.

Method	5 September 2005		27 June 2005	
(lr)2-3 (lr)4-5	Spatial CC	RMS [mm]	Spatial CC	RMS [mm]
WRF data	0.70	1.33	0.85	0.87
Remote sensing data	0.87	0.90	0.72	1.13
Data fusion	0.91	0.82	0.86	0.92

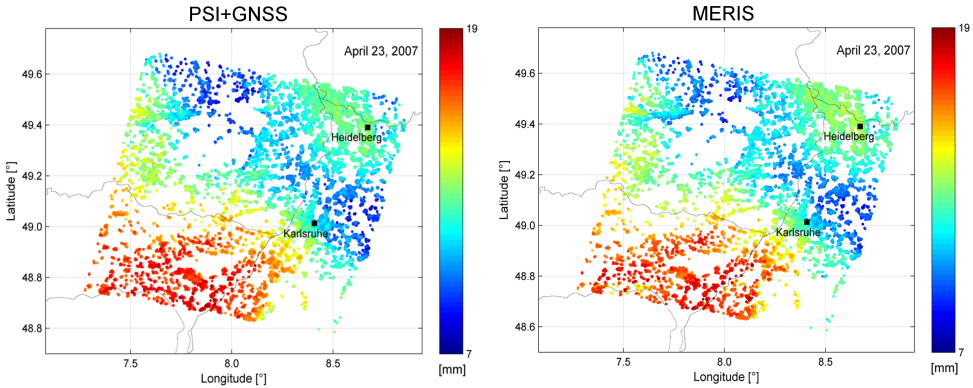

Figure 1. Maps of the absolute atmospheric PWV derived by combining PSI and GNSS data and the corresponding map from MERIS. The spatial correlation is 95 % and the RMS value of the differences is 0.68 mm.

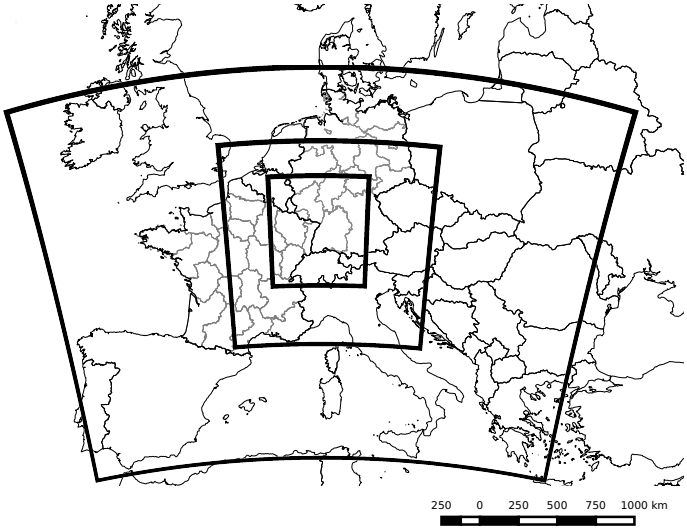

Figure 2. WRF model set up with a parent domain of resolution 27 km × 27 km and two nests of 9 km × 9 km and 3 km × 3 km, respectively.

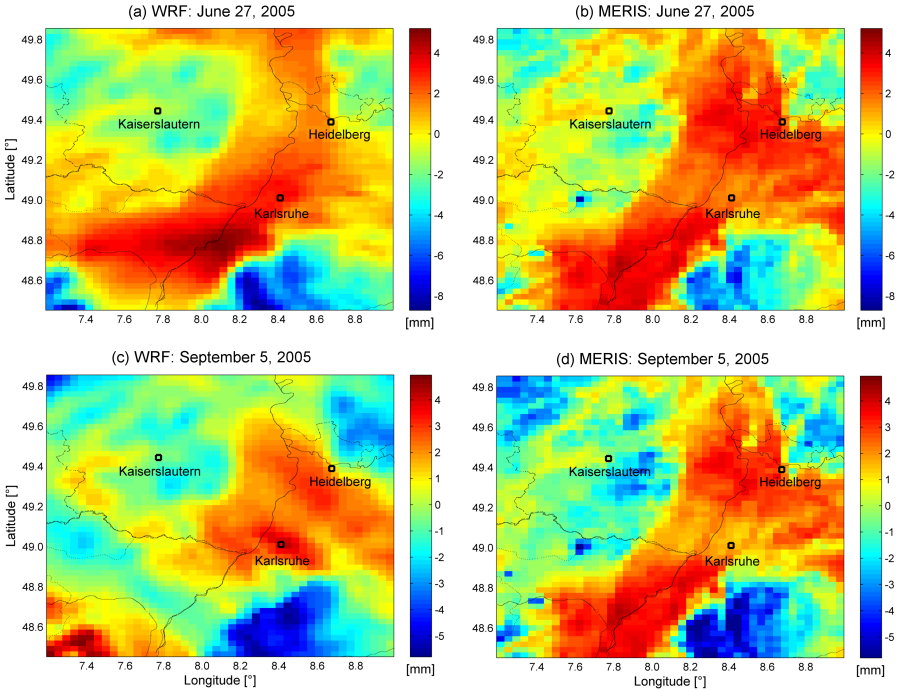

Figure 3. Maps of PWV content as received from MERIS and WRF, where a linear trend is subtracted from each map. The upper data are received on 27 June 2005 (09:51 UTC), while the lower data on 5 September 2005 (09:51 UTC). Gaussian averaging is applied to scale the MERIS data at WRF resolution, 3 km × 3 km. The spatial correlation coefficient between the upper maps is 0.8 and 0.71 for the lower.

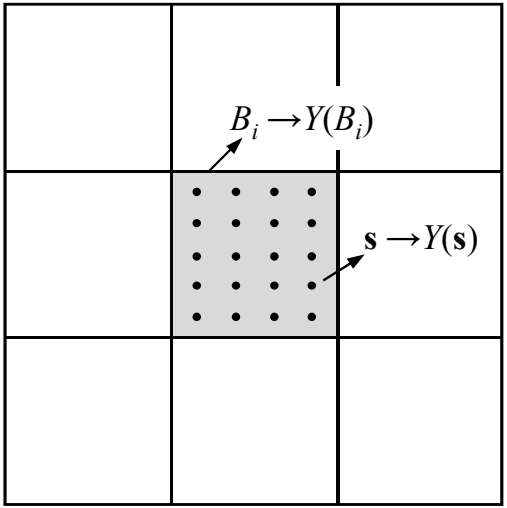

Figure 4. Point and block data, such that for spatial data, $Y(B_i)$ represents the average of the point data within the block.

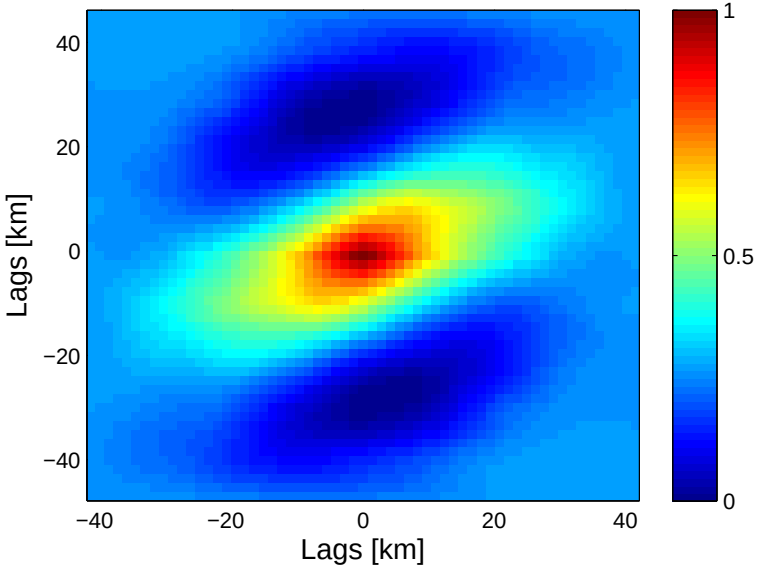

Figure 5. Spatial autocorrelation function for a PWV map, with the long-wavelength component removed, computed from remote sensing data acquired on 5 September 2005, 10:51 UTC.

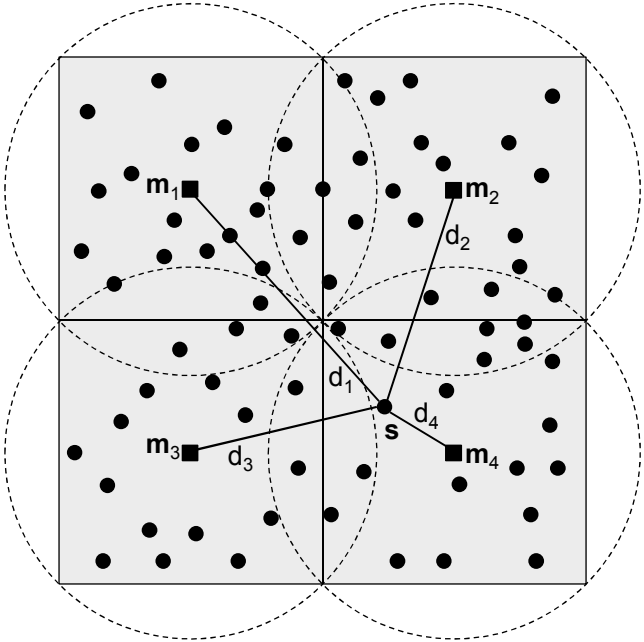

Figure 6. The observation domain with the black dots define the locations at which the data are available. The black little squares indicate the nodes. The weights for each location $\boldsymbol{s}$ are related to the distances d_i. The dashed circles define the radius for each node.

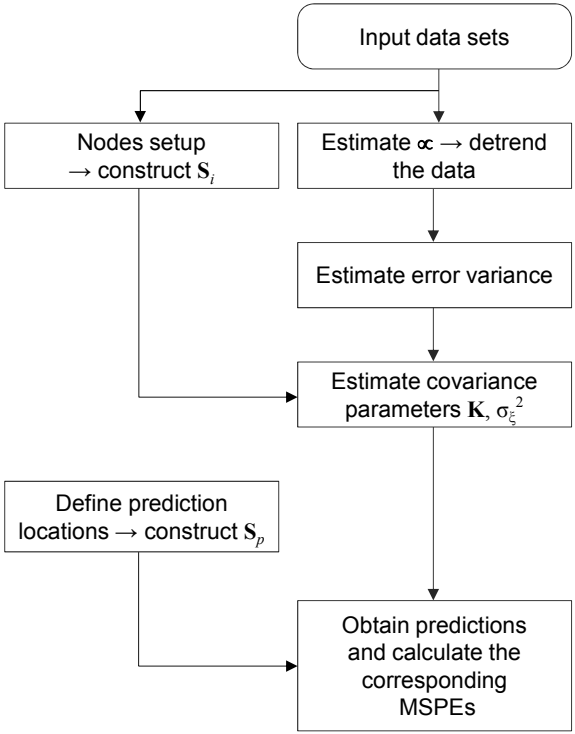

Figure 7. Obtaining predictions via the FRK method.

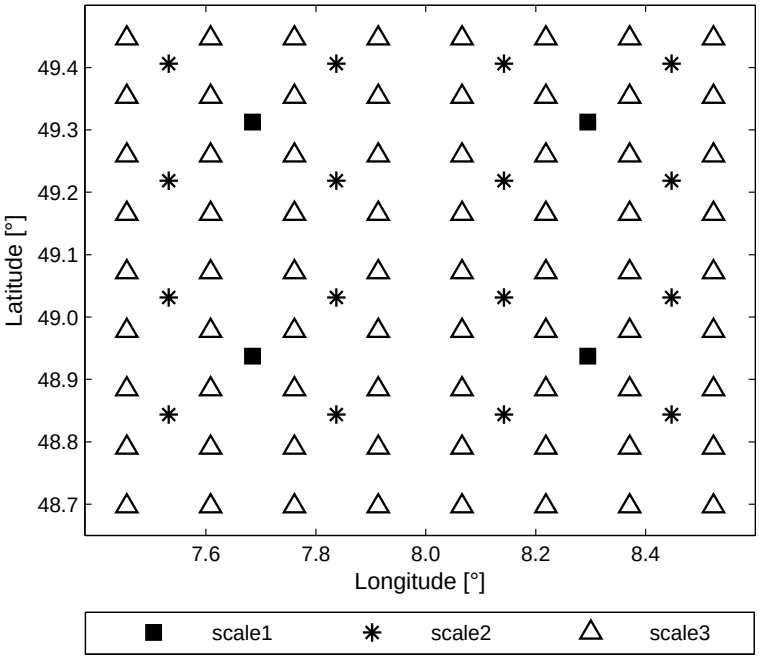

Figure 8. FRK nodes or center locations of 93 basis functions at three spatial resolutions. The first resolution is 40 km, the second resolution is 20 km, and the third resolution is 10 km.

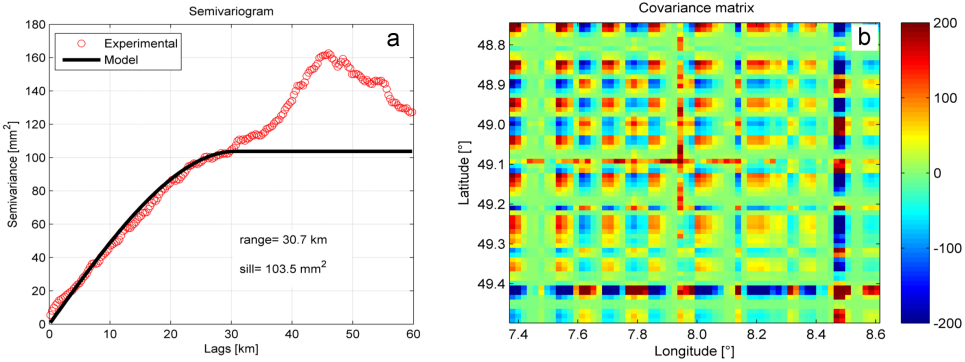

Figure 9. (a) The experimental semivariogram and the fitted spherical variogram model, (b) Covariance matrix used to predict the wet delay maps in Fig. 10

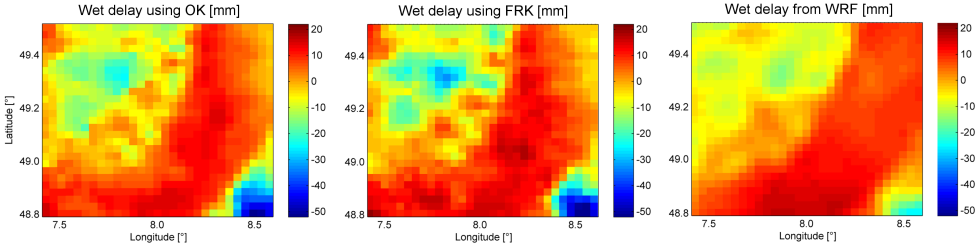

Figure 10. Wet delay prediction map using the block OK and FRK. The resolution of the grid is 3 km × 3 km. A point-level wet delay map, on 23 May 2005 at 09:51 UTC, is used as input to the algorithms.

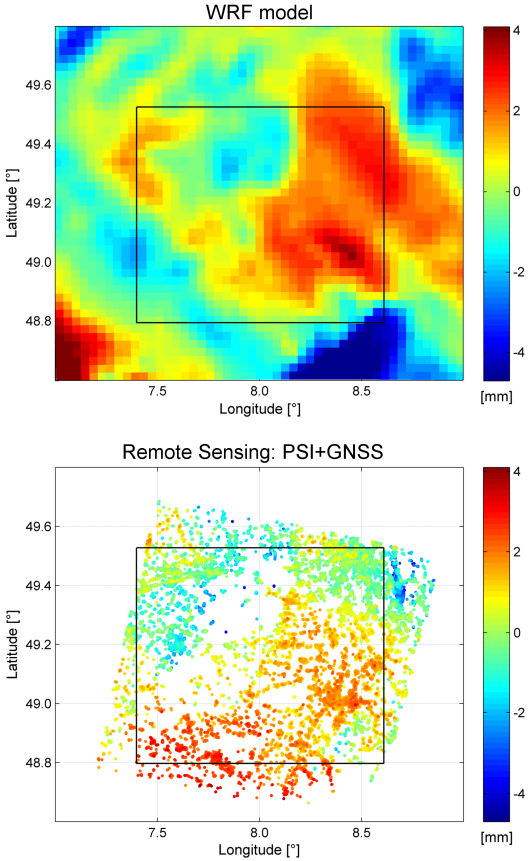

Figure 11. PWV maps from PSI + GNSS combination and WRF on 5 September 2005, with a linear trend subtracted from each map. PSI + GNSS provide point-level observations, while WRF generates block data with a block size of 3 km × 3 km. The predictions will be obtained within the area indicated by the black box.

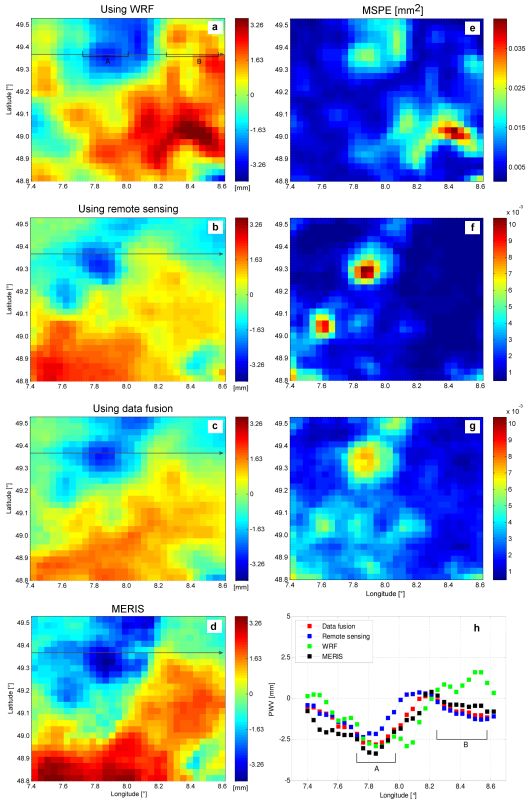

Figure 12. PWV prediction and MSPE maps obtained by data fusion of PWV estimates from PSI and GNSS and maps from WRF as well as predictions obtained by applying FRK to individual data sets. The data are available on 5 September 2005 at 09:51 UTC. The output grid has a block size of 3 km × 3 km. The label A defines a region of sparse remote sensing data and and the model data in region B are highly overestimated.

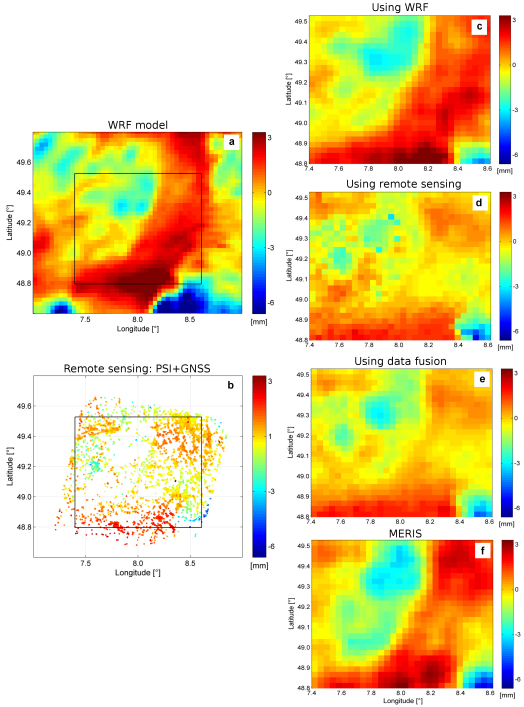

Figure 13. PWV maps from remote sensing (PSI+GNSS) and WRF model data on 27 June 2005 at 09:51 UTC as well as prediction maps obtained by data fusion and individual data sets. The output grid has a block size of 3 km $\times$ 3 km over the area indicated by the black box in a and b.

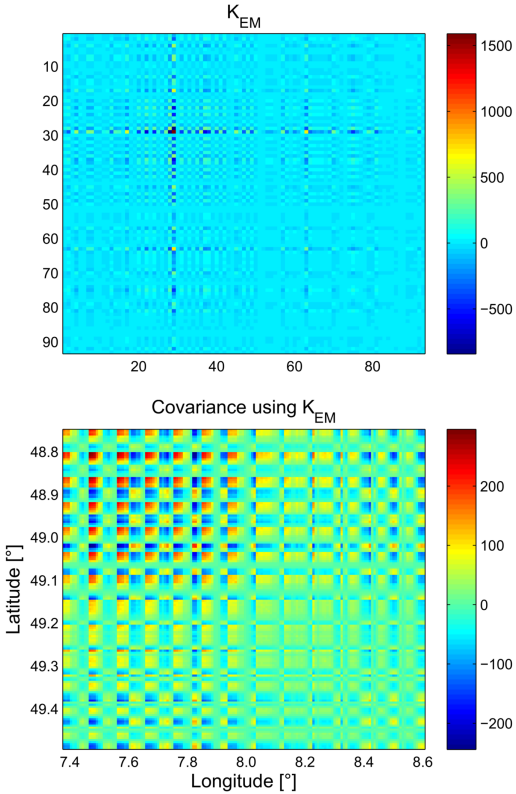

Figure 14. Estimate of the covariance matrix $\mathbf{K}$ using the E-M algorithm and the corresponding covariance matrix for the Wet delay map from PSI + GNSS. The wet delay observations are aggregated into maps of 7×7 km^2 cells before their covariance matrices are computed.