Water vapor integration methods to improve the quality of Synthetic Aperture Radar observations

Presenting author
Mario Montopoli

Mario Montopoli[1,2], F.S. Marzano[2,3], E. Pichelli[2], D. Cimini[4], R. Ferretti [2], S. Bonafoni[5], D. Perissin[6], F. Rocca[7], N. Pierdicca[3]

[1]*Department of Electrical and Information Engineering – University of L'Aquila Via G. Gronchi, 18, I-67100 L'Aquila (AQ) - Italy, mario.montopoli@univaq.it*
[2] *Center of Excellence CETEMPS, University of L'Aquila, Via Vetoio, L'Aquila, Italy*
[3]*DIE, University La Sapienza of Rome ,Italy.*
[4] *Istituto di Metodologie per l'Analisi Ambientale-Consiglio Nazionale delle Ricerche, C.da S.Loja, Potenza, Italy.*
[5]*DIEI, University of Perugia, via G. Duranti 93, 06125, Italy.*
[6] *Institute of Space and Earth Information Science, Chinese University of Hong Kong.*
[7] *DIEI, Polytechnic of Milan, Italy.*

(Dated: 1 July 2010)

1. Abstract

The presence of water vapor in the atmosphere represents one of the most important problems to derive accurate products from differential Synthetic Aperture Radar (SAR) observations. In this work several indirect measurements of integrated water vapor in the Lazio region, in the center of Italy, have been considered to characterize the spatial features of water vapor. This task is of fundamental importance to address the water vapor integration procedures based on kriging algorithms. In this work several kriging procedures have been tested using ground stations GPS and satellite (eg. MERIS) retrievals to provide an integrated product which can be used to mitigate SAR impairments due to water vapor.

2. Introduction

Synthetic Aperture Radar (SAR) is a well established microwave imaging system from which measurements of surface deformations of the order of centimeters can be derived and than several useful land applications (e.g.: the analysis of progressive tectonic motions, or the improvement of a Digital Terrain Model) can be provided to the community. Among the main limitations affecting the Interferometric SAR (InSAR) measurements, especially at C and X frequency bands, the atmosphere surely plays a relevant role. When two interferometric SAR images are not simultaneously acquired, the electromagnetic wave received from the SAR sensor, mounted on a satellite platform, after interactions with the ground, may be differently affected by the atmosphere which induces an unwanted component on the received signal. In particular, the random nature of the atmospheric state (i.e.: different humidity, temperature and pressure) between two acquired SAR observations may have a visible and fatal consequences on the interferometric phase. Among others, the water vapor is an important contributor to the error budget of InSAR data and for this reason its spatial and temporal characterization plays an important role.

In this work, the spatial characterization of vertical Integrated Precipitable Water Vapor (IPWV), as seen from various satellite sensors, will be dealt with. Data acquired from Envisat- Medium Resolution Imaging Spectrometer Instrument (MERIS), and Terra- Moderate Resolution Imaging Spectroradiometer (MODIS) and Aqua-MODIS spectrometer, operating at infrared frequencies at spatial resolution of 0.3, and 1 km respectively, will be compared with simulations derived from MM5 weather forecast model at 1km resolution as well. The InSAR signal from ASAR of Envisat platform and RadarSat is also exploited to derive estimates of differential IPWV (dIPWV) at very high spatial resolutions (about 100 m). dIPWV estimates are analyzed as well and compared together with those derived from previously mentioned spectrometers in terms of correlation structures. This intense experimental activity has been dealing with the Mitigation of Electromagnetic Transmission errors induced by Atmospheric Water Vapour Effects (METAWAVE) project *(Pierdicca et al., 2009*) sponsored by the European space agency and managed by University La Sapienza of Rome which coordinates many teams from several Italian universities.

The results of the comparisons here presented put the basis for developing algorithm for merging water vapor information from different observations. In this respect, different types of kriging and co-kriging approaches are applied to MERIS and GPS water vapor fields and their performance compared each others.

3. Data set description

During September and October 2008 an experimental campaign has been set up within the METAWAVE project. The selected target area was that of the center of Italy in the urban area of Rome. Among other observations, the collected data include: GPS time series, MERIS and MODIS from Terra and Aqua sensors and Advanced SAR (ASAR) acquisitions. The Numerical simulations of IPWV have been considered as well from MM5 weather numerical model in the same days of the experimental campaign (*Pichelli et al., 2009*). **Table 1** lists the measurements that were available for this study, together with some of its features. In order to strengthen the subsequent spatial analysis, additional observations and simulations have been considered with respect to those performed during the experimental campaign. They are listed in Table 1 as well.

	Year	Month and Day	Temporal Res. [h]	Spatial Res. [km]	minLon [DD]	maxLon [DD]	minLat [DD]	maxLat [DD]
MM5	94	1, 16 Gen.; 6 Feb.; 26 Mar.; 7,10 Apr.	1	1	11.6404	13.2889	41.3445	42.5701
	08	17, 20, 23,24, 26,27,29,30 Sep; 02-03 Oct.	1	1				
Terra	08	23-30 Sept.; 1-7 Oct.	-	1	7.5300	18.5811	36.6670	46.8830
Acqua	08	23-30 Sept.; 1-7 Oct.	-	1	7.5300	18.5803	36.6670	46.8830
MERIS	08	20-30 Sept.; 2-3 Oct.	3 day	0.25	12.0000	13.0000	41.5000	42.3000
RSAT	03, 07	2003-2007	-	0.1	12.3465	12.6102	41.7603	42.0040
ASAR	02, 08	2002-2008		0.1	11.8120	13.1127	41.6905	42.2288
ERS1	94	Jan. Apr.	-	0.1	12.0533	13.4281	41.2689	42.2470
GPS	08	20-30 Sep. 01-04 Oct.	0.5	-				

Table 1: Available data set .

4. Water vapor spatial features

All data, listed in Table 1, have been used in the tentative to characterize the spatial behaviour of water vapour at different spatial scales ranging from few meters to dozen of kilometres. To this aim acquisitions from remote sensors, usually provided in terms of zenithal total delay from GPS, atmospheric phase screen from SAR and radiance from MERIS and MODIS, have been previously converted in columnar integrated perceptible water vapour expressed for this study in centimetres. After these conversions, the semivariogram has been extracted. Semivariogram, under the spatial stationarity hypothesis, is defined as follows (*Wackernagel, 1995*):

$$\gamma_V(d) = \frac{1}{2} \cdot \left\langle \left| \tilde{V}(\mathbf{s}+d) - \tilde{V}(\mathbf{s}) \right|^2 \right\rangle_s \tag{1}$$

where γ_V indicates the semivariogram of integrated precipitable water vapour (here indicated with the symbol V and expressed in *cm*) for a given source of information and d (km) is the lag distance between two points at position **s** and **s**+d whereas the symbolism <·> stands for average operator in the spatial domain. To avoid to misrepresent the spatial features of water vapour and to obey to the stationarity hypothesis as much as possible, the trend of V as a function of the terrain height h (m) has been removed from each source of observation leading to the de-trended integrated precipitable water vapour indicated with symbol $\tilde{V}$ and obtained as follows:

$$\tilde{V}(\mathbf{s}) = V(\mathbf{s}) - \overline{V}(\mathbf{s}) \tag{2}$$

where $\overline{V}$ is the drift function modelled with an exponential law:

$$\overline{V}(s) = k_0 \cdot \exp\left(-k_1 \cdot h(\mathbf{s})\right) \tag{3}$$

In the equation (3), k_0 (*cm*) and k_1 (m^{-1}) are regression coefficients, which are in principle different for each data source. Evidence of gently exponential behaviour of IPWV with terrain height h has been already investigated in the past using GPS data alone (e.g. *Basili et. al., 2004*; *Morland, 2007*) and here confirmed using heterogeneous data (not here shown for brevity). Typical average values of k_0 and k_1 have been found, for the 2008 campaign, to be of the order 1.67 cm and $5 \cdot 10^{-4}$ cm·m^{-1}. After the removal of the spatial trend on water vapour data by means of Equation (2), the equation (1) has been applied to all source of data listed in Table 1. The result of this analysis is shown on the left panel of **Figure 1**. All curves in this figure seem to follow a Power Law (PL) relation of the form:

$$\gamma_{PL}(d) = \alpha \cdot d^{\beta} \tag{4}$$

as already consolidated in past studies (e.g. *Williams et al., 1998*). In general, α and β respectively represent the characteristic scale of the process and the power index which expresses the rate at which the water vapour, supposed to be a

random function, decorrelates with increasing lag distance. For the curves depicted in the left panel of Figure 1**,** at least at large scales, β approaches 2/3 (see the dotted grey trends overimposed on the figure grid) and this seems to agree with the Kolmogorov theory of turbulence (*Tatarskii, 1971*). It should be mentioned that, since the DInSaR observations are differences between two acquisitions, supposed independent and made at different instants, the corresponding semivariogram have to be divided by a factor 2 to be compared together with the semivariograms derived from other types of observations. From the left panel of Figure 1, all variograms are biased towards different levels and this reflects the fact that different source of information are associated to different sensors and then different way to observe the same scene. For example, MODIS acquisitions (i.e.: cyan and orange curves respectively showing varigromas from AQUA and TERRA sensors) are quite flat around values higher than those of the other semivariograms retrieved from observations made at the same days (e.g. from MERIS red curve).

Further analysis of semivariograms, shown on the left panel of Figure 1, reveals a good agreement together with MM5 numerical weather simulations and dInSAR observations during the 2008 experimental campaign (compare blue curves of ASAR passages and black curve with the star markers for MM5 simulations). Semivariograms estimated from other observations and numerical simulations are shown in the same figure as well and they reveal a fairly good agreement in terms of semivariogram shapes with respect to that extracted from 90-m Digital Elevation Model (DEM) (freely downloadable from Shuttle Radar Topography Mission (SRTM) web site). The latter has been normalized by subtraction of the value of 10^-7 to deal with a better comparison with the other curves. In a whole left panel of Figure 1 indicates that spatial features from different source are difficult to interpret due to the different process and noise associated to the instrument used to acquire the observed scene. In other words, the IPWV can be seen as the result of a filtering process representing the impact of the instrument used to observe IPWV field itself. Thus, if not properly accounted for, the resulting semivariograms will be affected by the instrumental error noise which, in general, depends from the used instrument.

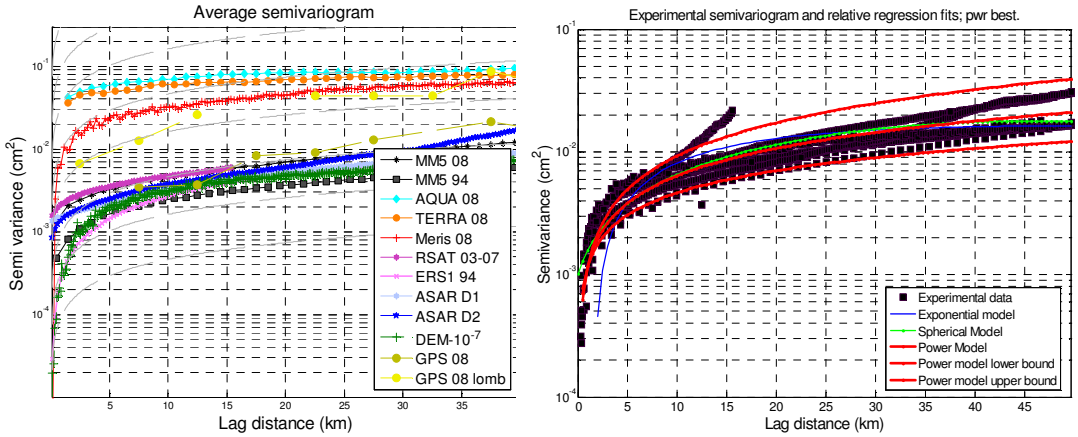

Figure 1: On the left panel: semivariograms from different source of data where 2/3 power law is shown with grey dotted lines. On the right panel: the normalized version of semivariograms shown on the left. Red curves on the right panel represent the power law model of the type $\gamma = \alpha \cdot d^{\beta}$ *with* α *and* β *regression coefficients equals to* $\beta = 0.74$, $\beta = 0.90$, $\beta = 0.60$ *respectively for middle upper and lower bounds red curves and* $\alpha = 0.0012$*, for all of them.*

4.1 Variogram normalization

To be able to compare variograms derived from different source of information, a semivariogram normalization procedure has been applied and the normalized variograms are shown on the right panel of Figure 1. After normalization, it appears that the power law model is very suitable for describing the spatial features of IPWV even at small spatial scales. The 2/3 (β=0.67) slope has been found after normalization (see middle red fitting curve having a slope of 0.74) even thought variations from this value are possible (see upper and lower red curves respectively with slope parameter 0.90 and 0.60). Semivariogram normalization procedure just mentioned is devoted to characterize the spatial filter that ideally describes the IPWV measurement process transforming the true, but unknown IPWV, into the measured IPWV. To accomplish this step, the MM5 model is used by exploiting their characteristics of being numerical simulations instead of observations subjected to a measurement processes.

5. Water vapor data integration

5.1 Interpolation Methods

A macroscopic classification of interpolation techniques can be made into deterministic and geostatistical methods. All methods rely on the similarity of nearby sparse sample points to create, in a bi-dimensional framework, a surface of regular

gridded points. Deterministic techniques use mathematical functions for interpolation. Geostatistical techniques assume that a spatiotemporal process is composed of deterministic and stochastic components. The deterministic components, sometimes called drifts can be global or local trends. The stochastic component is formed by a purely random and an autocorrelated part. An autocorrelated component implies that on average, closer observations are more similar than distant observations. This behaviour is described by the variogram where squared differences between observations are plotted against their separation distances. The fundamental idea of Dr. Krige was to use the variogram for interpolation as means to determine the magnitude of influence of neighbouring observations when predicting values at unobserved locations. Due to the statistical nature of IPWV, geostatistical methods are here preferred with respect to deterministic ones.

5.2 Kriging approaches

There exists different formulation of kriging. For example, Ordinary Kriging (OK) is able to make interpolations with constant mean. In other words, the interpolated field has a constant trend. In the case of IPWV, there is a trend with the terrain height that can be assumed to be exponential. To include this trend in the resulting interpolated field, the Kriging with External Drift (KED) has to be used. In principle, there exist two types of procedure to include external drift within kriging. The simplest approach foresees the use of OK on de-trended values $\tilde{V}$ and then to sum the trend, previously subtracted, on the resulting interpolated values. This way to proceed, sometimes referred as regression kriging, has been here labelled as KEDP where letter "P" indicates the *a posteriori* operation of summing the trend before removed. The second way to proceed based on KED is more rigorous and directly includes the trend within the kriging framework. When two source of punctual information are available, Co Ordinary Kriging (COK) approaches can be followed. More details on kriging can be found on (*Cressie, 1991*).

In this work several interpolation approaches have been tested: i) the Ordinary kriging ,OK, on GPS; ii) Kriging with External Drift ,KED, on GPS; iii) KED on GPS where de trending and trending procedure are respectively performed before and after running the kriging algorithm (KEDP); iv) the case where de-trending and trending operations are performed when CO- Ordinary Kriging (COKP) is used merging MERIS and GPS; v) the case where the MERIS samples, employed also for COKP, are added to the those from GPS stations and used for running KED (labeled as KED+). The use of Meris together with GPS is justified by the fact that Meris IPWV estimates have been found to be the most consistent with those derived from GPS with respect the other available source of information.

5.3 Multy-Source integration strategy

The approach used to integrate IPWV data foresees the use of MERIS, GPS observations and numerical weather forecasts by MM5 model. The choice to include MERIS data, instead of MODIS acquisitions, relies on the fact that the first one are available together with ASAR acquisitions on ENVISAT platform and they additionally show a more suitable correlation structure which is a desirable feature within the kriging framework. The flow chart scheme, representing the operations for water vapour integrations, is shown on **Figure 2**.

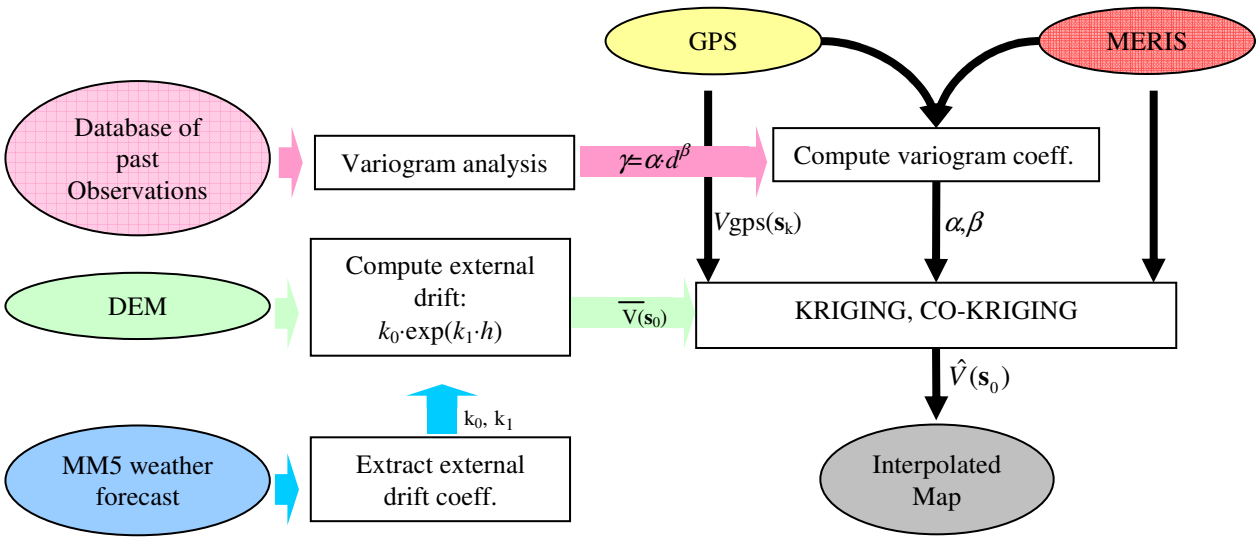

Figure 2:Flow chart of the Integrated Precipitable Water Vapor (IPWV) integration scheme.

Following figure 3, the main steps can be summarized as follows: 1) Variogram models are defined from a preliminary analysis based on a representative data set of heterogeneous source of water vapour information. In our case the power law model seems to be adequate to describe the spatial variability at different scales. 2) Run MM5 simulations of IPWV for the desired days with a time and spatial resolutions at least of 1 hour and 1 km, respectively. 3) For each observation instant (for

example SAR passages): 3.1) Compute semivariogram γ and relative regression coefficients α and β (see Eq. (1) and (4)); 3.2) Compute external drift coefficients k_0 and k_1 from MM5 simulations; 3.3) Compute external drift function $\overline{V}$ at the desired spatial resolution using the available Digital Elevation Model (DEM); 3.4) Perform the kriging using the available sources (MERIS or GPS of both) 3.5) Repeat steps from 3.1 to 3.4 for each instant for which the MM5 simulations are available. It is worth mentioning that step 3.2 could appear superfluous since also GPS and MERIS potentially can provide information of IPWV against the height. However, GPS stations not always cover the whole target area so that a representative behaviour of IPWV with height can be extracted whereas MERIS, apart from being not always available as a LEO platform, can be strongly affected by cloud contamination. For these reasons, the use of weather forecast models like MM5 has been included in the interpolation procedure.

6. Results

In this section tests on interpolation methods before introduced are carry out. Additional MM5 outputs and the drift function itself (see equation (3)) are included too as supplier of IPWV. The benchmarks, taken as reference for testing the effectiveness of the interpolated maps, are the MERIS overpasses. The convenience of this choice is threefold: i) MERIS observations are available on ENVISAT platform in conjunction with ASAR overpasses, ii) MERIS with its 250-m spatial resolution is closer with respect to the highest resolution of ASAR or other SARs iii) as above mentioned, MERIS observations, acquired during the 2008 experimental campaign and considered for the purpose of the data integration, show a good agreement with respect to the GPS retrievals of IPWV.

In order to run kriging procedures in their different formulation, semivariograms and cross-semivariograms have been computed from GPS and MERIS on the detrended water vapor field $\tilde{V}$ (see equation (2)). Thus, the kriging interpolations of IPWV have been carried out on the target domain of Lazio region. The resulting interpolated maps $\hat{V}(\mathbf{s}_0)$ on the final regular gridded domain are compared together with the MERIS acquisitions $V(\mathbf{s}_0)$ at the same position $\mathbf{s}_0$. The errors between $\hat{V}(\mathbf{s}_0)$ and $V(\mathbf{s}_0)$ have been computed and listed on **Table 2** in terms of Root Mean Square Error (RMSE), Standard Deviation (STD), bias and Correlation coefficient (Cc).

Scores	*KED (250 m)*	*KEDP (250 m)*	*KED+ (250m)*	*COKP (250m)*	*DEM (250m)*	*MM5 (1km)*	*OK (1km)*	*Event dates*
Std	0.13	0.13	0.12	0.11	0.13	0.12	0.20	***Sept. 20th 09:30 UTC***
Bias	-0.06	-0.09	-0.01	-0.07	-0.01	-0.04	-0.10	
Rmse	0.14	0.15	0.12	0.13	0.13	0.12	0.23	
C.corr	0.83	0.84	0.83	0.88	0.86	0.85	0.47	
Std	0.15	0.16	0.26	0.16	0.16	0.14	0.28	***Sept. 23th 09:30 UTC***
Bias	0.02	-0.07	-0.07	0.03	-0.12	-0.10	-0.21	
Rmse	0.15	0.17	0.26	0.16	0.20	0.17	0.35	
C.corr	0.94	0.94	0.79	0.93	0.92	0.94	0.76	
Std	0.09	0.09	0.08	0.08	0.08	0.12	0.20	***Sept. 27th 09:00 UTC***
Bias	-0.06	-0.12	-0.02	-0.12	0.02	0.04	-0.13	
Rmse	0.11	0.15	0.08	0.14	0.08	0.12	0.24	
C.corr	0.91	0.92	0.94	0.93	0.93	0.85	0.50	
Std	0.12	0.12	0.11	0.11	0.13	0.15	0.24	***Sept. 29th 10:00 UTC***
Bias	0.04	-0.01	-0.01	-0.02	-0.17	-0.17	-0.04	
Rmse	0.12	0.12	0.11	0.11	0.22	0.23	0.25	
C.corr	0.88	0.88	0.89	0.89	0.84	0.84	0.23	
Std	0.10	0.11	0.07	0.09	0.12	0.11	0.27	***Sept. 30th 09:00 UTC***
Bias	0.04	-0.03	0.01	0.03	-0.13	-0.10	-0.06	
Rmse	0.11	0.11	0.07	0.10	0.18	0.15	0.28	
C.corr	0.96	0.96	0.98	0.97	0.93	0.94	0.62	

Table 2: Error table for different interpolation algorithms: Ordinary Kriging (OK) of GPS, Kriging with External Drift (KED) of GPS, Kriging with External Drift which is subtracted and summed before and after running OK on GPS (KEDP), KED with the inclusion of some MERIS samples (KED+), Co-kriging with external drift between GPS and MERIS (COKP), Digital Elevation Model (DEM) transformed into IPWV trough equation (3) and coefficients computed for each case study. All score error numbers are expressed in cm.

From this table, as expected, it emerges that OK performs worse than the other approaches since the imposed trend is constant and this is in contrast with the IPWV field features. Interpolations on the 250 m and 1 km final grid resolution have been performed but we found that going down to finer resolutions does not improve the score errors. When KED is performed at 250m, thus exploiting a digital elevation model at that resolution, the error scores improves. This result tends to demonstrate the importance of the use of an accurate DEM within the proposed scheme. Even when DEM is used alone to

run equation (3), with proper coefficients k_0 and k_1, derived from MM5, the error scores are comparable and in some cases even better than other methods (see for example the case on Sept. 27, 2008). This is, to some extent, surprising since the coefficient k_0 and k_1 are estimated from an independent source of information like MM5 forecast model. When comparing the KED and KEDP strategies, KED performs slightly better or at least is comparable to KEDP. Of course, when KED+ is used with a denser number of samples (i.e. in or case taking the GPS stations and some MERIS random samples), the score errors decrease in most of case studies. Eventually co-ordinary kriging between GPS and MERIS samples seems to performs slightly worse with respect other cases. This is probably due to the complexity of co-kriging approaches which require the estimation of additional information like cross variograms. When few samples are available (as probably in this case) cross-variogram estimation can results very inaccurate leading to large interpolation errors.

7. Conclusions

In this work a heterogeneous data set of retrievals of integrated precipitable water vapor is analyzed. To be aimed at integrating water vapor sources, its spatial characterization has been carry out by variogram analysis. The latter is fundamental to run kriging algorithms in their different formulations.

Results highlight that variograms computed from different sensors are affected by an intrinsic noise leading to a different variogram trends even when IPWV retrievals are taken nearly at the same time. This leads to the concept of variogram normalization. To normalize variograms, forecast simulation of IPWV can be taken into account for example. Of course to avoid orographic contamination, variograms should be computed after removing the water vapor trend with the terrain height. Thus, a power law model as been found to be suitable for describing variograms of water vapor.

Several formulations of kriging have been tested as well using GPS and MERIS data. Kriging with external drift with the use of both GPS stations and some MERIS samples is resulted, in most of considered cases, the best choice.

Future developments will be focused on the comparisons of integrated maps of IPWV with differential SAR acquisitions to validate the proposed integration strategy

References

Basili P., S. Bonafoni, V. Mattioli, P. Ciotti, N. Pierdicca, (2004), "Mapping the atmospheric water vapour by integrating microwave radiometer and GPS measurements", *IEEE Trans. on Geosci. and Remote Sensing*, **42**, no. 8, pp. 1657-1665.

Cressie Noel A.C., (1991), "Statistics for Spatial Data", *A Wiley-Interscience publication*.

Morland J. and Matzler C., (2007), "Spatial interpolation of GPS integrated water vapour measurements made in the Swiss Alps", *Meteorol. Appl.*, **14**, 15–26

Pichelli E., R. Ferretti, D. Cimini, D. Perissin, M. Montopoli, F.S. Marzano, and N. Pierdicca, (2009), "Water vapour distribution at urban scale using high-resolution numerical weather model and spaceborne SAR interferometric data", *Nat. Hazards and Earth System Sciences*.

Pierdicca, N., F. S. Marzano, F. Rocca; D. Perissin, P. Basili, S. Bonafoni, V. Mattioli, D. Cimini, P. Ciotti, R. Ferretti, M. Montopoli, W. Foster, R. Notarpietro, S. Padmanabhan, E. Pichelli, S. Reising, S. Sahoo, G. Venuti, (2009), "Atmospheric water-vapour effects on spaceborne Interferometric SAR imaging: data synergy and comparison with ground-based measurements and meteorological model simulations at urban scale", *Proc. of 3rd European Conference on Antennas and Propagation*, Berlin, Germany.

Tatarskii, V.I., (1971),"The effects of the turbulent atmosphere on wave propagation", 472 pp., *Israel Program for Scientific Translations Ltd*.

Wackernagel H., (1995), Multivariate Geostatistics. Berlin, Germany, Springer-Verlag.

Williams, S., Y. Bock, and P. Fang, Integrated satellite interferometry (1998): "Troposphere noise, GPS estimates, and implications for synthetic aperture radar products", *Journal of Geophysical Research*, **103** (B11), 27051-27067.