VALIDATING TIME SERIES OF A COMBINED GPS AND MERIS INTEGRATED WATER VAPOR PRODUCT

R. Lindenbergh[1], **H. van der Marel**[1], **M. Keshin**[2], **and S. de Haan**[3]

[1]*Delft Institute of Earth Observation and Space Systems, Delft University of Technology, P.O. Box 5058, 2600 GB, The Netherlands*
[2]*Finnish Geodetic Institute, Finland*
[3]*KNMI Royal Netherlands Meteorological Institute, The Netherlands*

ABSTRACT

Increased knowledge of atmospheric water vapor can improve weather predictions and is expected to reduce errors in products derived from GPS and (In)SAR data. At GPS ground stations Integrated Water Vapor (IWV) is estimated from the GPS signal delay with a high temporal resolution. The Envisat MERIS spectrometer obtains spatially dense IWV observations but at limited moments in time. In this research the additional value of MERIS IWV is evaluated when added to GPS IWV for the purpose of obtaining a high quality spatial-temporal water vapor product. At each of 39 stations, first GPS IWV from surrounding stations is used to produce a two months time series of IWV with a temporal resolution of one hour. Then both GPS and MERIS IWV are used together. The two resulting time series are validated against direct GPS IWV as measured at the station.

Key words: Water vapor; GPS; MERIS; data fusion; geostatistics; spatio-temporal; validation.

1. INTRODUCTION

Water vapor is the atmosphere's dominant greenhouse gas, but it is still a challenge to determine its spatial-temporal distribution at sufficient resolution from one single type of contemporary meteorological instrument. It is possible however to retrieve and consecutively combine water vapor estimates from complementary satellite systems. At ground stations from the world wide Global Positioning System (GPS), the zenith Integrated Water Vapor (IWV) is derived from estimates of GPS signal travel time delay in troposphere. This derivation results in relative good measurements with high temporal, e.g. 1 hour, but low spatial resolution, e.g. tenths of kilometers over Western Europe. The Medium Resolution Imaging Spectrometer (MERIS) on the Envisat satellite retrieves IWV by comparing radiances in two spectral bands in the near infrared, with a maximum spatial resolution of 300 m. Its temporal resolution is only in the order of days. Moreover, MERIS only provides useful IWV observations under clear sky conditions.

The topic of this research is the validation of a method that creates time series of, say, hourly IWV prediction maps by combining complementary MERIS and GPS IWV observations, (Lindenbergh et al., 2008). Fusion of complementary data sets representing the same attribute in some spatial-temporal domain is expected to result in a data product of better quality. Here we want to quantify this potential improvement in quality. A suitable quantification can be used to assess whether the proposed data fusion is worthwhile to be implemented in an operational setting, for example for the purpose of numeric weather prediction, (De Haan, 2008) or for atmospheric corrections on other satellite products, (Li et al., 2006).

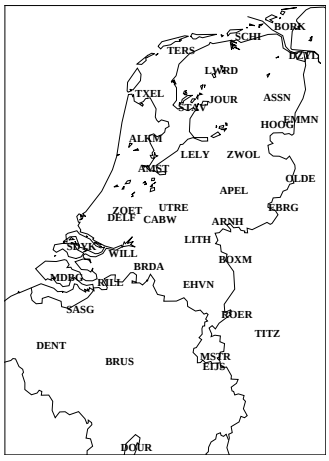

Figure 1. Region of Interest with GPS ground stations

In our particular setting we aim to quantify the additional value of the MERIS IWV data when used in combination with the GPS IWV data for creating a spatial-temporal water vapor product. The approach that we take is a kind of double cross-validation: the available data is used to

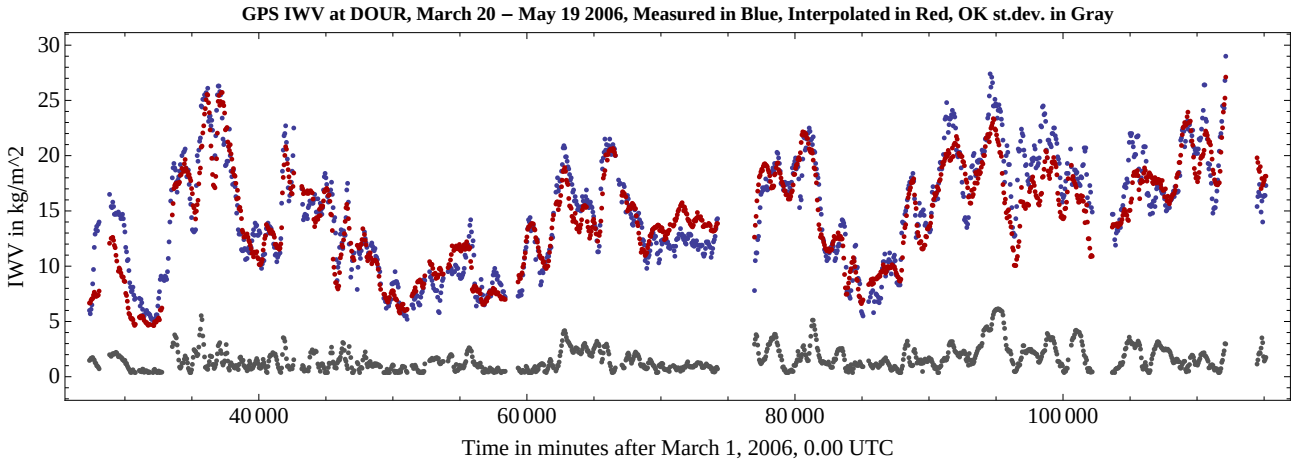

Figure 2. **In red:** *Reference time series of GPS IWV at station location DOUR.* **In blue:** *IWV time series at DOUR obtained by Ordinary Kriging interpolation from GPS IWV observations at other ground stations.* **In gray:** *Ordinary Kriging st.dev. at DOUR obtained by Ordinary Kriging interpolation from GPS IWV observations at other ground stations.*

estimate the IWV value at the locations and times where we know it best, i.e. at the GPS ground station locations, in two different ways. Firstly, only GPS IWV from other stations is used in estimating the IWV value at a given station. Secondly GPS and MERIS IWV are used together for the purpose of IWV estimation. By comparing the two outcomes to the reference data, i.e. the direct GPS IWV observations, insight in the additional value of the MERIS IWV is obtained.

If we use all 39 station locations for validation, we have a spatially representative set of test locations, some close to other stations, some relatively remote. The location of the stations is shown in Fig. 1. In order to minimize the effect of random fluctuations in the GPS IWV measurements, we consider time series of two months at a temporal resolution of one hour. In practice this means that each of the three considered data sets consists of approximately 50 000 data points (i.e. $39 \times 24 \times 60$ rounded down, to allow for some gaps in the available GPS IWV observations.) Note that a benefit of this validation approach is that only a limited amount of processing is required. Rather than to produce series of maps we only have to produce time series at 39 designated location which makes the processing much more efficient.

This paper is organized as follows. In Section 2 the available GPS and MERIS IWV observations are presented and the construction of the two sets of time series to be validated is described in detail. In Section 3 the created time series are compared to the reference data, i.e. the direct GPS IWV, before the paper finishes with conclusion.

2. TIME SERIES CONSTRUCTION

In this section the set of reference time series and the two sets of interpolated GPS IWV and fused GPS + MERIS IWV are described.

2.1. GPS IWV time series

In the nineties of the last century a method was developed to map the tropospheric delay time of GPS signals directly to the amount of water vapor in the atmosphere, (Bevis et al., 1994). Since then this method became operational at networks of GPS ground stations, compare e.g. (Elgered, 2001). In practice this means that processing centers convert the delay time near real-time into Integrated Water Vapor (IWV) content at temporal intervals typically between 15 minutes to 1 hour at each of the ground stations they manage. In recent years the amount of ground stations from which at least hourly GPS IWV observations can be obtained is increasing. Disadvantage of this approach is that GPS IWV is only monitored at locations where hard infrastructure is available (i.e. the GPS ground station), which implies that areas with less infrastructure are typically undersampled. Recently however it is considered to obtain GPS IWV observations from ships as well, (Fujita et al., 2008). In this research we consider time series of GPS IWV from 39 ground stations in and around The Netherlands as processed by the Royal Dutch Meteorological Institute, KNMI. We consider all GPS IWV observations available at the whole hour for the period from March 20, 2006 to May 19, 2006. For the purpose of the fusion with MERIS data, we consider time t in minutes with minute 0 starting at March 1, 0:00 UTC time. As an example, the time series of GPS IWV at the ground station of DOUR is shown in Fig. 2 in red. As can be seen, some gaps in the observations exist, for example near $t = 75000$. This is typical for all GPS IWV time series considered.

2.2. Cross-validation of GPS-IWV

At each station location a second time series is constructed using the GPS IWV observations from the other

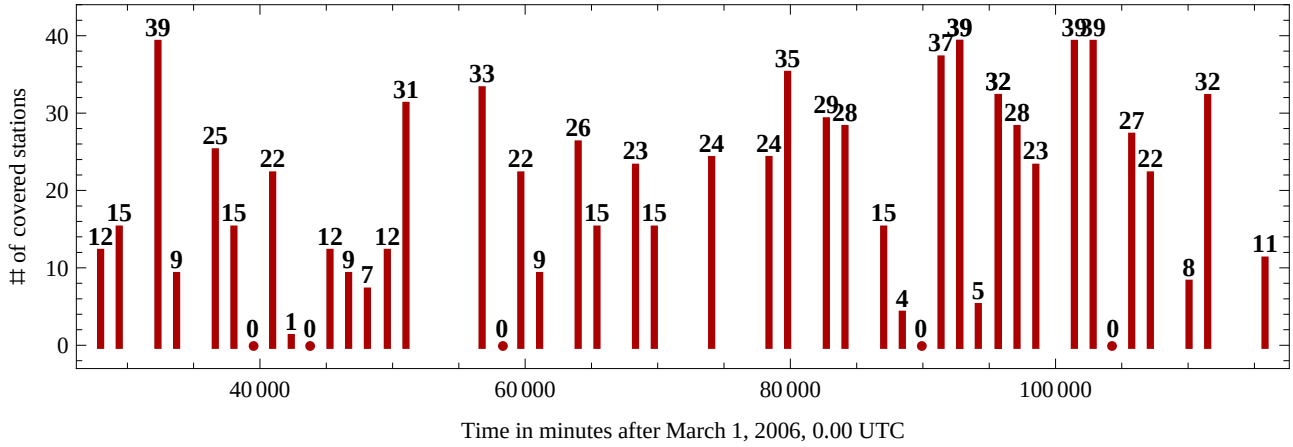

Figure 3. Temporal spreading of the MERIS scenes and number of station locations covered by at least one MERIS pixel per scene.

stations. For this purpose spatial interpolation by means of Ordinary Kriging, (Goovaerts, 1997), is applied on the GPS IWV observations of the other stations that are available at the moment of interpolation. Here some assumptions have to be mentioned. First of all, the height of the ground stations is not taken into account: in contrast to for example Switzerland, (Morland and Mätzler, 2007), the topography in and around The Netherlands is flat; still height differences in the order of 100 m between station locations may exist. It is also assumed that no trend exists in the water vapor fields. Statistically it means that it is assumed that the water vapor signal is observed from a stationary random field with covariance function, such that Ordinary Kriging can be used for interpolation. In practice, sometimes a water vapor front may occur, implying a strong trend in the data, which may result in non-optimal interpolation results.

To take the temporal variability in the water vapor field into account, covariogram parameters are reestimated at each hour. For this purpose, first an experimental covariogram, (Goovaerts, 1997), is determined for each epoch from all observations of that epoch at the different station locations. To this experimental covariogram a spherical covariance model is fitted, resulting in two parameter values, corresponding to an optimal fit in the least squares sense. The first parameter is the correlation range, the maximal distance at which still spatial correlation between observations exists, the second is the sill, which parameterizes the maximal average variability in the signal. An additional fixed parameter is the nugget that expresses the uncertainty in the direct GPS IWV observations (1 kg/m^2)^2). This covariance function is consecutively applied to determine weights for the available observations at the other stations, resulting in a spatial interpolation value and a spatial interpolation variance. This variance value expresses how close the observations are to the estimation location relative to the correlation between the observations. Therefore a high Ordinary Kriging variance indicates that no highly correlated observations could be used for the interpolation.

In Fig. 2 the time series resulting from this spatial interpolation scheme is shown in blue for the station of DOUR. No interpolation was performed for epochs with not enough observations available at the other stations. In gray also the Ordinary Kriging st.dev., the square root of the OK variance, is shown. As the correlation range is determined anew in every epoch, the st.dev. is varying, mainly depending on the correlation range of the epoch considered and to a lesser extend on the available observations in a certain epoch.

2.3. Spatio-temporal fusion of GPS-IWV and MERIS-IWV

Except for the time series of GPS IWV also Integrated Water Vapor observations from the Medium Resolution Imaging Spectrometer (MERIS) on board of the Envisat satellite are available, (Bennartz and Fischer, 2001). Basically MERIS retrieves IWV by comparing radiances in two spectral bands in the near infrared at a maximal resolution of 300 m. For this research 47 MERIS reduced resolution products were available for the period March 20 - May 19, 2006 that had non-empty overlap with the region of interest shown in Fig. 1. At reduced resolution one MERIS IWV pixel is available for every 1.2×1.2 km^2, (ESA, 2006). First cloudy pixels were removed in a standard procedure based on MERIS product quality flags as described in (Lindenbergh et al., 2008). It has been indicated by several people that this procedure does not remove all cloudy pixels, and new cloud detection algorithms are under development, (Gomez-Chova et al., 2007; Preusker et al., 2006). These new methods were not yet available to the authors though.

For the fusion with the GPS IWV observations, the Co-Kriging approach, as described in (Lindenbergh et al., 2008) was implemented. In this approach, the IWV is only estimated at fixed epochs, notably those at which GPS IWV observations are available. Here these fixed

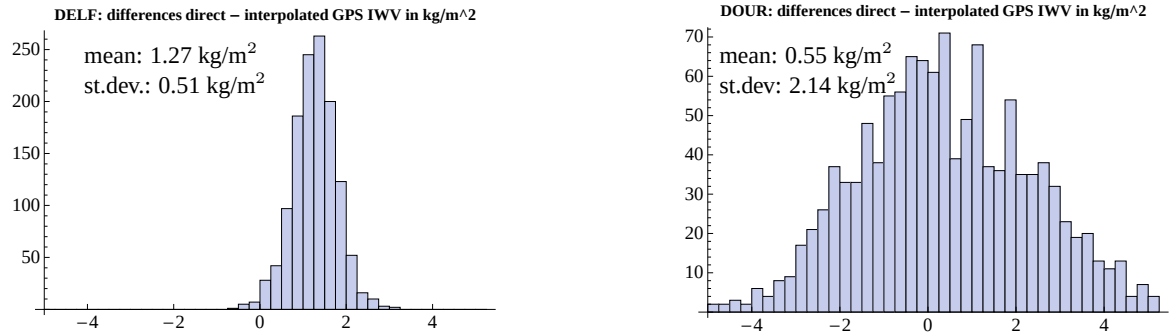

Figure 4. Histograms of differences between direct GPS IWV values and cross-validated GPS IWV values at two stations, one at the middle right of the region of interest and one at the bottom, compare Fig. 5. **Left.** *DELF* **Right.** *DOUR.*

epochs are the one hour intervals that coincide with KNMI GPS IWV acquisition times. An IWV estimate at a certain location at such epoch time is constructed from two ingredients: a spatial interpolation contribution of the available GPS IWV observations for that epoch, and a temporal interpolation contribution consisting of a MERIS IWV observation, available at the same location but in general observed at a different moment in time. The relative weights over the available GPS IWV observations is divided according to the description in Section 2.2.

The temporal correlation between the collocated MERIS IWV observations and the moment of observation is determined from a temporal spherical covariance function with a fixed temporal range of 10 hours. In a more elaborated, but probably more correct approach the temporal range could be determined dynamically in an experimental procedure as well. The sill in general changes every epoch and is taken equal to the spatial sill as determined from the spatial covariogram fitting procedure as described in Section 2.2 as well. Again a nugget of $(1\ \mathrm{kg/m^2})^2$ is implemented to incorporate the uncertainty in the MERIS IWV observations. The relative correlations of the measurement space-time point to the GPS IWV observations and the single MERIS observation balances the weight that each observation gets in the interpolation. As before, not only a Kriging estimated value but also a Kriging variance is obtained.

In this setup only the MERIS pixels covering a station location are relevant. Therefore a preprocessing step is applied that for each MERIS scene and for each station location is determining all MERIS pixels within 1.75 km of the station location. In a previous work, (Lindenbergh et al., 2008), this 1.75 km distance resulted in maximal correlation between MERIS and GPS IWV. The found MERIS pixels at a given station location at a given MERIS scene are averaged and stored as the MERIS IWV value at that station location. Possibly the amount of MERIS pixels within the 1.75 km radius circle and their spread in IWV value can be incorporated in future to obtain a more realistic variance for the final MERIS IWV value.

Due to the Envisat orbit configuration and due to the removal of cloudy IWV pixels from the MERIS scenes, the number of station locations where a MERIS IWV value is present is rather limited. An overview of the temporal spreading of the MERIS scenes and of the number of the stations covered in each scene is given in Fig. 3.

Potentially the signal available in the MERIS scenes could be exploited more completely by also adding a spatial interpolation step to the processing procedure of the MERIS scenes. For the particular setup here, this would mean that a MERIS value at a station location is obtained from more far away MERIS pixels. As such values would be less reliable in general, they also should be accompanied by a relative high variance that downweight their influence, i.e. weight, in the further fusion procedure.

2.4. Fusion wrap-up

Here a short summary of the fusion procedure as implemented for this paper is given. The fusion time series are constructed by going from the first to the last epoch one by one. At a given epoch it is first considered if there at least 12 GPS IWV values present. Otherwise the epoch is canceled. Then it is checked whether there is a MERIS scene available in a time span of 10 hours before that epoch. Only the last available MERIS scene is taken into account. Now one by one an IWV value is estimated at the 39 station locations. If for a station a MERIS IWV value is available in the last MERIS scene of less than 10 hours old, this value is incorporated, together with the available GPS IWV values from that epoch. If no MERIS IWV value is present, the estimated IWV is based on GPS IWV only.

3. COMPARISON OF RESULTING IWV TIME SERIES

In this section the set of interpolated GPS IWV and the set of fused GPS and MERIS IWV time series are compared to the reference set of directly measured GPS IWV at the 39 station locations considered.

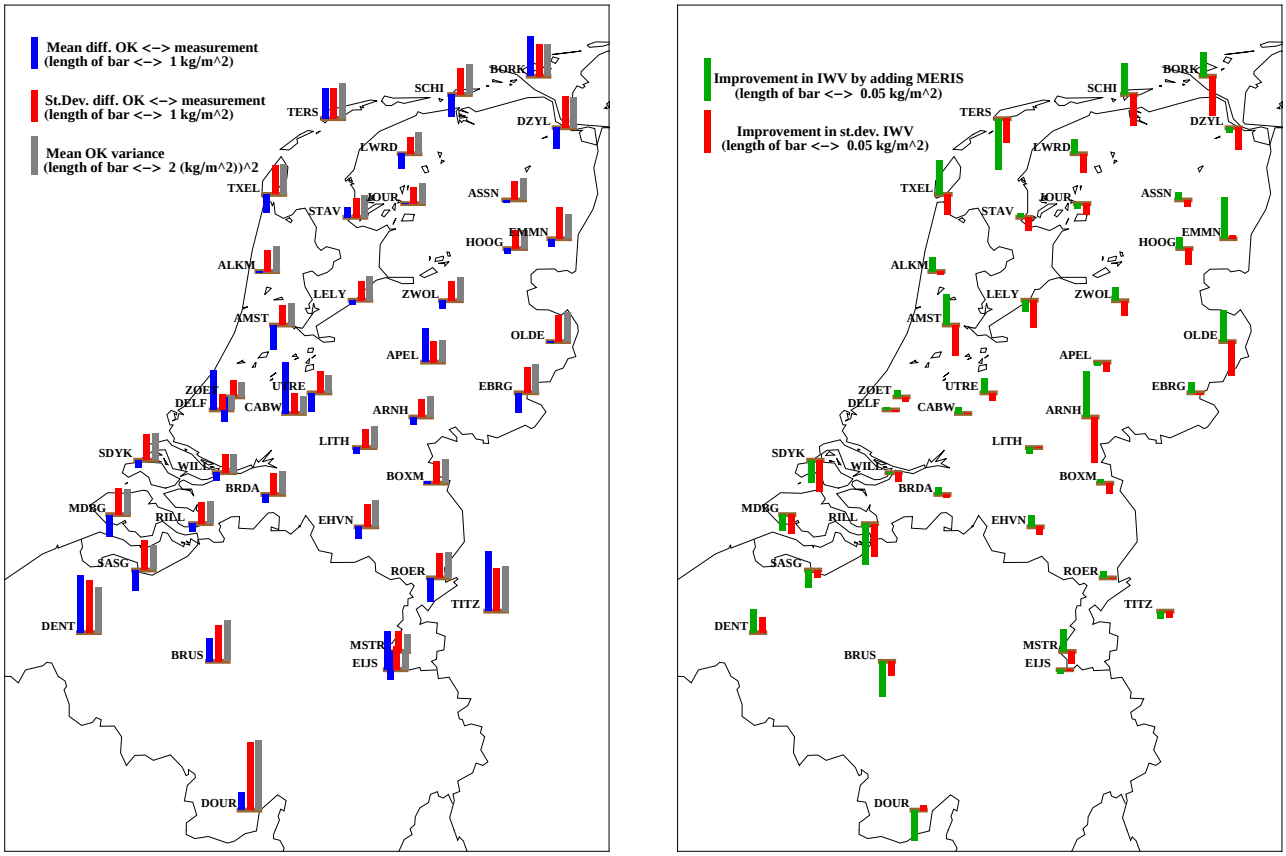

Figure 5. **Left.** *Cross validation results. The blue bars give the mean difference between the GPS IWV at each station and the values resulting from interpolating measurements at the other stations to that location. The red bars give the st.dev. of the differences, and the gray bars the mean Ordinary Kriging variance of the interpolated value at each station.* **Right.** *For each station location the average 'improvement' over two months, obtained by adding MERIS IWV to the interpolated GPS IWV, is shown in green. An upward green bar indicates a real average improvement, a downward bar indicates that on average the interpolation results degraded when compared to the original GPS IWV time series at that station. Similarly in red the improvement or degradation in the st.dev. of the differences between interpolated IWV and measured GPS IWV is indicated.*

3.1. Cross validation results GPS IWV.

In Fig. 5, left, an overview of the comparison of the direct GPS IWV observations to IWV estimates obtained by interpolating GPS IWV of the other stations is given. Two type of statistics are distinguished that are further illustrated by the histogram of differences in Fig. 4.

The first statistic is the difference in mean between observed and interpolated GPS IWV. As can be seen in Fig. 4, left, the mean difference at station DELF is high with 1.27 kg/m^2 compared to the mean difference at station DOUR of 0.55 kg/m^2. These mean differences are indicated on the map in Fig. 5, left, by the blue bars. Relative high mean differences at the interior stations of APEL, DELF and CABW may indicate systematic biases in the GPS IWV observations at these locations. Higher mean differences more at the border of the area of interest could also be caused by the smoothing effect of the interpolation method Kriging.

The second statistics is the standard deviation of the differences. At DELF the st.dev. is relatively small with 0.51 kg/m^2 compared to the st.dev. of 2.14 kg/m^2 at DOUR. In Fig. 5, left, these st.dev. values are indicated by the red bars. Clearly, larger st.dev. values are dominant at the border of the map. This is explainable by noting that the signal at more isolated ground stations is to a lesser extent represented by the other ground stations. This also explains why the Ordinary Kriging variance, indicated by the gray bars in Fig. 5, left, shows a similar pattern: near the borders of the map less correlated IWV values are available, resulting in higher OK variances.

3.2. Validation of GPS and MERIS IWV fusion.

In Fig. 5, right, the 'improvements' obtained by adding MERIS IWV to the GPS IWV interpolation is visualized per station. The green bars represent the difference between ([interpolated MERIS + GPS IWV] - [reference])

and ([interpolated GPS IWV] - [reference]), which is indicated as the improvement. The results seem to show that the additional value of the MERIS IWV observations is negligible. Indeed, the mean improvement over all 39 stations is only 0.005 kg/m^2, which cannot be considered significant. Similarly the spread of the differences between interpolated GPS + MERIS IWV and the reference set is also not significantly improved w.r.t. the differences between interpolated GPS IWV only and reference set: the mean improvement in st.dev. over all 39 stations is -.02 kg/m^2.

The theoretically expected improvement of adding the MERIS IWV to the GPS IWV interpolation can be quantified by the decrease in Kriging variance. The mean variance of all 49 142 interpolated values before adding MERIS equals 1.364 (kg/m^2)^2, after adding MERIS the mean variance is reduced to 1.341 (kg/m^2)^2, an improvement of 1.67 %. This theoretical improvement takes only the correlation between the observations into account, not the actual values of the observations.

4. CONCLUSIONS AND OUTLOOK

In this paper it is shown how to validate the additional value of adding MERIS IWV to GPS IWV observations for the purpose of obtaining a better spatial-temporal integrated water vapor product. The results of the validation give strong indication that the additional value of MERIS IWV is limited over an area where a dense network of GPS ground stations is available. Main reason is the limited temporal coverage of the MERIS scenes, combined with the large data loss due to cloud coverage. An additional analysis is still necessary to determine the exact influence of several choices that were made in the fusion procedure.

Cross validation of the GPS IWV observations indicate that biases of 1-2 kg/m^2 exists between the KNMI processing results of ground stations of their GPS network.

Although the results presented here indicate that the benefit of a GPS and MERIS IWV fusion for e.g. numerical weather prediction, (De Haan, 2008), is limited over an area like The Netherlands, still several applications exists where a fusion could be beneficial. Examples are areas were GPS ground stations are sparse, but GPS IWV could also be used to 'repair' MERIS IWV suffering from cloud cover for applications where knowledge of water vapor can improve the processing of other data from for example the Envisat satellite.

ACKNOWLEDGMENTS

The MERIS data used in this paper are disseminated by the European Space Agency, ESA. This project is funded under number EO-085 by the Netherlands Institute for Space Research, SRON.

REFERENCES

Bennartz, R. and Fischer, J. (2001). Retrieval of columnar water vapour over land from back-scattered solar radiation using the Medium Resolution Imaging Spectrometer (MERIS). *Remote Sensing of Environment*, 78:271–280.

Bevis, M., Businger, S., Chiswell, S., Herring, T., Anthes, R., Rocken, C., and Ware, R. (1994). GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water. *Journal of Applied Meteorology*, 33:379–386.

De Haan, S. (2008). *Meteorological applications of a surface network of Global Positioning System receivers*. PhD thesis, Wageningen University.

Elgered, G. (2001). An overview of COST Action 716: Exploitation of ground-based GPS for climate and numerical weather prediction applications. *Phys. Chem. Earth.*, 26A:399–404.

ESA (2006). MERIS Product Handbook, Issue 2.1. http://envisat.esa.int/dataproducts/meris/, last visited: September 19, 2008.

Fujita, M., Kimura, F., Yoneyama, K., and Yoshizaki, M. (2008). Verification of precipitable water vapor estimated from shipborne GPS measurements. *Geophysical Research Letters*, 35(L13803). doi:10.1029/2008GL033764.

Gomez-Chova, L., Camps-Valls, G., Calpe-Maravilla, J., Guanter, L., and Moreno, J. (2007). Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 45(13):4105 – 4118.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, Oxford.

Li, Z., Muller, J.-P., Cross, P., and Fielding, E. J. (2006). Interferometric synthetic aperture radar (InSAR) atmospheric correction: GPS, moderate resolution imaging spectroradiometer (MODIS), and InSAR integration AM-CW terrestrial laser scanner system. *Journal of Geophysical Research*, 11(B03410). doi:10.1029/2004JB003446.

Lindenbergh, R., Keshin, M., van der Marel, H., and Hanssen, R. (2008). High resolution spatio-temporal water vapor mapping using GPS and MERIS observations. *International Journal of Remote Sensing*, 29(8):2393–2409.

Morland, J. and Mätzler, C. (2007). Spatial interpolation of GPS integrated water vapour measurements made in the Swiss Alps. *Meteorological Applications*, 14:15–26.

Preusker, R., Huenerbein, A., and Fischer, J. (2006). Cloud detection with MERIS using oxygen absorption measurements. *Geophysical Research Abstracts*, 8:09956.