Towards an Optimal Merging of Satellite Data Sets

[1]Jörg Schulz and [2]Ralf Lindau

[1]Deutsche Wetterdienst, germany, e-mail: joerg.schulz@dwd.de
[2]Meteorological Institute, University of Bonn, Germany

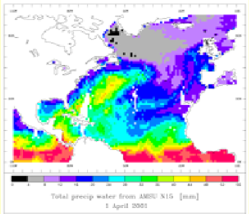
► Objectives

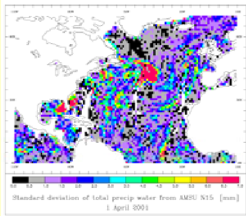
The need for a comprehensive and accurate global water vapor data set as an assisting tool for scientific studies in the atmospheric sciences has been acknowledged during the last 10 years. Such a data set is extremely useful for all aspects of climate science being dependent on accurate water budget data, e.g. general circulation model verification or regional climate studies. The Humidity Composite Product (HCP) of EUMETSAT's Satellite Application Facility on Climate Monitoring (CM-SAF) will integrate data from several existing and upcoming satellites including the SEVIRI instrument on EUMETSAT's Meteosat Second Generation (MSG) satellite and the water vapor profiling instruments Infrared Atmospheric Sounding Interferometer (IASI), High Resolution Infrared Sounder (HIRS), Microwave Humidity Sounder (MHS), Global Navigation Satellite System Receiver for Atmospheric Sounding (GRAS) on the Meteorological Operational polar platform (MetOp).

Aiming at an optimal composite, the information from different sources has to be merged in a reasonable way. In this pilot study water vapor observations from AMSU-A (Advanced Microwave Sounding Unit) onboard the NOAA (National Oceanic and Atmospheric Administration) satellites and from SSM/I (Special Sensor Microwave/Imager) on DMSP (Defense Meteorological Satellite Program) satellites are used to construct a merged product. The merger is performed by Kriging, an optimal interpolation technique that provides not only fully covered fields but also a corresponding map of errors. The technique has been chosen because its potential to merge data from several completely different sources and will be implemented within the Version 3 of CM-SAF products.

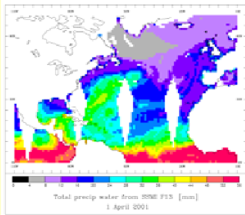
► Data

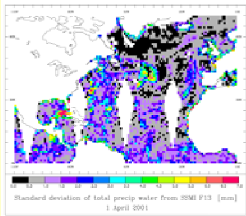
AMSU





SSM/I





The data used are swath-oriented total precipitable water estimates from two AMSU-A instruments on NOAA -15 and 16 as well as three SSM/I instruments on the DMSP platforms F13, 14, and 15. This pilot study is spatiotemporally restricted to the North Atlantic and data of four months (April, July, October 2001, and January 2002.

TPW (upper) and standard deviation (lower) in 1° x 1° grid boxes from AMSU-A on NOAA-15 (left) and the SSM/I on DMSP-F13 (right).
Note that SSM/I fields exhibit larger data gaps. Both fields show a spatial dependent standard deviation, and AMSU obeys larger errors.

► Kriging

Kriging can be regarded as a prediction of a value x_0 at a location P_0 where no measurement is available employing information from measurements at the surrounding locations P_i. A solution is not possible for one single case. However, if a time series of m measurements at each location P_i is available it is reasonable to minimise the expression:

$$\sum_{t=1}^{m}\left(x_0 = \sum_{i=1}^{n}\lambda_i(x_i+\Delta x_i)\right)^2 = \min$$

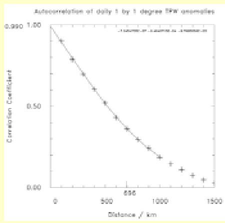
where x_i denote the available measurements, Δx_i their errors, and λ_i the weights. The minimised expression of this equation is equal to the error of the predicted value and reads :

$$[x_0x_0]-2\sum_{i=1}^{n}\lambda_i[x_0x_i]+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j[x_ix_j]+\sum_{i=1}^{n}\lambda_i\lambda_i[\Delta x_i\Delta x_i]$$

where [ ] denotes temporal averaging. The first term the variance at P_0 and is equal to unity if anomalies are considered. The second term contains the covariance between data points, the so called **information**. The third term contains the so called **redundance** because information from points P_i may not be independent. The last term describes the individual errors at the points P_i.

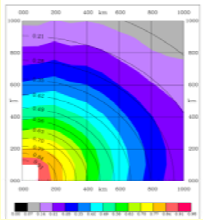
To determine the weights λ_i information on the spatial covariance $[x_ix_j]$ and the error of the individual observations $[\Delta x_i\Delta x_i]$ is needed.

► Spatial Covariance Function



The correlation function of daily means is calculated as a function of distance using anomalies. A quadratic exponential function is fitted to the correlation as shown in the left figure.

Correlation function for total precipitable water as derived from NOAA-15 AMSU data in April 2001.



Directional dependence of the correlation function for daily means.

Correlation length in km for different months and platforms.

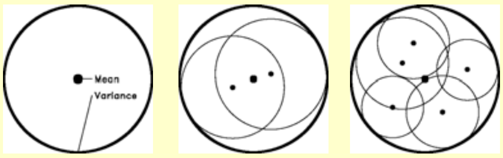
	January	April	July	October
AMSU N-15	632	696	551	590
AMSU N-16	634	700	543	574
SSM/I F13	608	695	529	590
SSM/I F14	627	707	525	590
SSM/I F15	632	710	545	602

A handy measure for this function is the correlation length which is approx. 696 km in this example. The correlation length stays rather constant for different platforms and shows a moderate variability over time as shown in the table. The effect of a directional dependence of the correlation length is indicated in the figure on the top right.

► Error Determination

The determination of the error variance is accomplished by a decomposition of the total variance into four components. These are i) the error of the monthly mean, ii) the seeming extra daily variance, iii) the mean error of daily means, and iv) the true intra daily variance. The total variance is split into an external and an internal part where ii) describes the variance between daily means which is external and iii) and iv) describe the variance within the days.

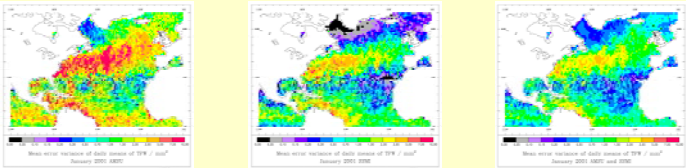
The most relevant problem in using satellite data within a Kriging approach is that satellite pixels are not independent of each other. Here we have five different platforms with only two different instruments/algorithms (AMSU and SSM/I). To find out which measurements are independent we consecutively shift the separation of variance into the internal and the external part by computing the average in 1° x 1° grid boxes i) from all pixels, ii) for each of the two instruments, and iii) for each of the five platforms. The changes in variance from step to step are equal to the variance of the added characteristic.



Schematic scheme for the computation of the error variance.

The table below shows the resulting mean error variances in mm² for the three different assumptions of independence. The assumption that pixels are independent yields a considerable lower error than the two others, i.e., pixels are not independent. Otherwise, the grid box errors for all three cases would be of comparable size, because an arbitrary averaging within a homogeneous data set would exhibit a reduction of variance which is proportional to the aggregation of individual observations. The transition from individual satellite platforms to only two different instrument types show that behaviour. Thus, each of the satellite platforms is considered as independent. It should be noted that the spatial error covariance is not considered within this study but will be included in the future.

Independency of ...	Mean number of independent observations per grid box	Internal variance	Estimated mean grid box error
Pixels	81.61	6.77	0.09
Satellites	5	2.38	0.60
Instrument types	2	0.65	0.65

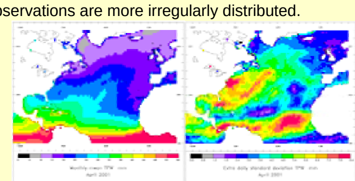


Spatial distribution of the monthly average of mean daily error of TPW as derived from AMSU (left panel) and SSM/I (middle panel) measurements as well as for the merged product (right panel) for January 2001.

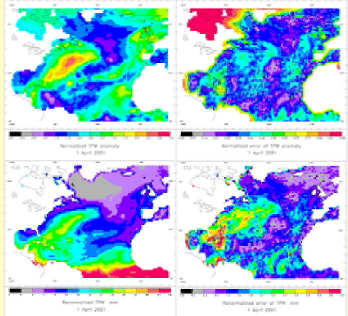
The common features of the errors for both instrument types are high errors over the Gulf Stream and in the Tropics which may be explained by the high variability of TPW in these regions. The figures clearly show that regional variations of errors exist and have to be taken into account. Even in the monthly average errors differ from about 0.1 mm² in the eastern subtropics to about 10 mm² over the Gulf Stream. The combination of both instruments diminishes the error in general.

► Products

Knowing the correlation function and the error of each observation Kriging has been applied successfully as demonstrated by the displayed products. A great benefit of the used technique is that each interpolated daily field is accompanied by a daily error map. The technique seems somewhat oversized for passive microwave data over the ocean but not for a merger of estimates from SEVIRI, IASI and GRAS data where the observations are more irregularly distributed.



Monthly mean and extra daily standard deviation of TPW for April 2001



Normalised TPW anomaly field (upper) and re-normalised TPW fields and their corresponding error fields for 1st of April 2001.

Acknowledgement
This research is part of the Visiting Scientist Program of the Satellite Application Facility on Climate Monitoring hosted by Deutscher Wetterdienst.The author thanks Dr. Ralph Ferraro(NOAA/NESDIS) for providing total precipitable water estimates from the AMSU instruments in orbit format.The SSM/I estimates were provided by the Sonderforschungsbereich 512 "Tiefdruckgebiete und Klimasystem des Nordatlantik" at the University of Hamburg, Germany. The authors are indebted to Karsten Fennig for the extraction of the relevant data from the Hamburg Ocean-Atmosphere Parameters and Fluxes from Satellite Data (HOAPS) data base.