AGU Fall Meeting 2012, Poster IN11A-1461

DISCOVER
www.discover-earth.org

Contribution of Data Set Construction Methodology to Data Set Uncertainties

Deborah K. Smith (smith@remss.com), Carl A. Mears, Kyle A. Hilburn, Lucrezia Ricciardulli, Frank J. Wentz, Remote Sensing Systems
Rick Schulte, student at Santa Clara University (summer intern at Remote Sensing Systems)

Remote Sensing Systems
www.remss.com

Summary

This poster summarizes the analysis of similar, yet distinctly different, water vapor products created using DISCOVER microwave radiometer data. The purpose of the analysis was to construct a 1-deg, monthly, merged water vapor data set best suited for climate study. Different approaches were applied and the impact of each approach studied. The new water vapor data product is now available to Users in netCDF format and contains water vapor values, a climatology, vapor anomalies, and a trend map. The construction process is described.

Background

Remote Sensing Systems (RSS) processes satellite microwave radiometer data with the help of NASA MEaSUREs funding. We carefully intercalibrate the data at the brightness temperature level and apply a consistent processing methodology to all sensors to produce Version-7 ocean products which include wind speeds, water vapor, cloud liquid water, rain rates and sea surface temperatures (for some instruments). A radiative transfer model is used to relate the brightness temperatures to ocean parameters. Water vapor is the most robust of the ocean products we retrieve due to the strong spectral signature in the radiometer frequency measurements, so it is the first merged-product we chose to create. Individual satellite ocean products are available, however, here we determine the best approach for merging the satellite data to create a water vapor product best suited to climate research.

Four scientists at RSS, each very familiar with the data we produce, independently created a 1-deg, monthly water vapor data set (see Table 1 below). We expected the outcomes of the independent approaches to be very similar. Instead, it was found that initial differences varied by over half a millimeter, showed regional differences, and had different trends despite using the same high quality input data. The differences are in Table 1. We next explored how each product construction decision affected the outcome. The decisions are summarized to the right. Most decisions were then automated except in a few cases for which data were excluded/included by the processing scientist.

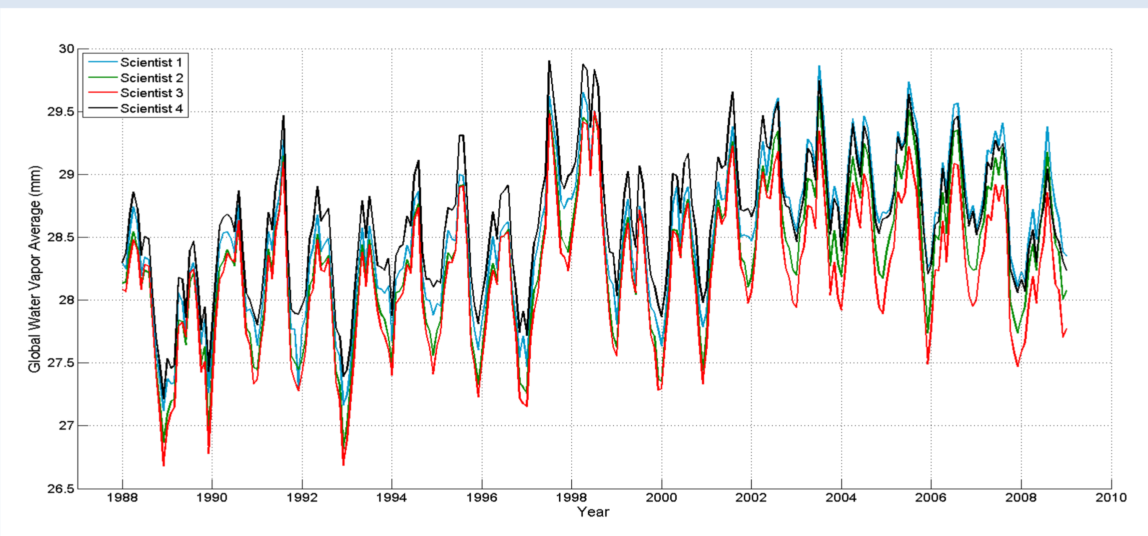


Figure 1. The global time series of water vapor mean from the four original data sets. The difference is as much as 0.5 mm. These differences decreased (see Figure 2) as we unified the approaches, gaining a better understanding of the contribution of each decision to the final product uncertainty. See the next section on construction decisions for more information.

Instrument	Details	Scientist 1	Scientist 2	Scientist 3	Scientist 4
F08	Start/Stop	Jul 1987 - Dec 1991	Jul 1987 - Dec 1991	Jul 1987 - Dec 1991	Jul 1987 - Dec 1991
	Omitted	Dec 87	2 cases of bad regional data in 87	none	none
F10	Start/Stop	Mar 1991 - Oct 1997	Dec 1990 – Nov 1997	Dec 1990 – Dec 1996	Dec 1990 – Nov 1997
	Omitted	Dec 90, Jan 91, Apr 91, Aug 91, Sep 91, Nov 97	4 cases of bad regional data in 93,94,96	none	none
F11	Start/Stop	Jan 1992 - Apr 2000	Dec 1991 – May 2000	Dec 1992- Apr 2000	Dec 1991 – May 2000
	Omitted	Dec 91, Apr 96, Mar 97, Apr 97	4 cases of bad regional data in 94	none	none
F13	Start/Stop	May 1995 - Oct 2009	May 1995 – Nov 2009	May 1995 – Nov 2009	May 1995 – Oct 2009
	Omitted	none	none	none	none
F14	Start/Stop	May 1997 - Jul 2008	May 1997 – Sep 2008	May 1997- Aug 2008	May 1997 – Jul 2008
	Omitted	none	none	none	none
F15	Start/Stop	Jan 2000 - Jun 2006	Dec 1999 - Aug 2006	Dec 1999- Sep 2006	Dec 1999 – May 2006
	Omitted	Dec 99	Aug 05 d1-7,Jul 05 d29,region specific	none	none
F17	Start/Stop	Dec 2006 - Jun 2012	Dec 2006 – Jun 2012	Dec 2006 - Jun 2012	Not Used
	Omitted	none	none	none	
AMSR-E	Start/Stop	Jun 2002 – Sep 2011	Jun 2002 – Oct 2011	Not Used	Not Used
	Omitted	none			
WindSat	Start/Stop	Feb 2003 – Jun 2012	Feb 2003 – Jun 2012	Not Used	Not Used
	Omitted	Feb-Jun 05, Jun/Jul 07, Jun 08	none		
Min Obs Req		per satellite	112	5	Variable capability
Extra Rain Flag	use +1 cell check	no	yes	no	no
Ice Removal	remove ice present	yes	yes, extended 1 cell	no	yes, > 1%
AverageMethod	bucket vs individual	Avg by instrument first, then bucket	bucket	bucket	bucket
Mean Vapor	global (1 std dev)	28.5 (0.6)	28.37 (0.6)	28.18 (0.6)	28.62 (0.5)
Vapor Trend	percent/decade	1.89	1.66	1.24	1.33

Table 1. The four approaches used different microwave radiometers, set distinct boundaries on each instrument, and used different construction decisions. Despite using high-quality , consistently processed, intercalibrated ocean data, the mean vapor and global trends varied more than expected.

Data Product Construction Decisions

Major Impacts

Geographic Sampling : An important contribution to the differences between the four approaches involved construction decisions that altered what data were used or omitted. Affected cells included coastal areas, areas of ice and regions of rain.

Minimal Number of Data per Grid Cell: Radiometers suffer from side lobe interference that prevents obtaining vapor values near land. Due to variations in instrument observation footprints, some grid cells have more observations than others. This can result in poorer quality vapor averages. Each approach listed in Table 1 used a different minimal number of obs requirement. See Figure 3 lower left plot as an example of the variability along coasts.

Ice Flagging: Ice is likely to exist more at one end of a month than another (with the exception of floating icebergs). Three approaches removed grid cells with ice present during the month, but the thresholds varied. The mean vapor was lower for Scientist 3 who did not use any ice removal scheme, thereby including data along the growing or receding ice edge. We determined that ice removal was necessary, and developed a mean-day-of-month quality calculation to remove ice edge grid cells. To handle icebergs, we used the number of ice observations within a grid cell during the month to exclude when too much ice existed. Figure 3 top right shows an example of ice counts for a southern ocean region.

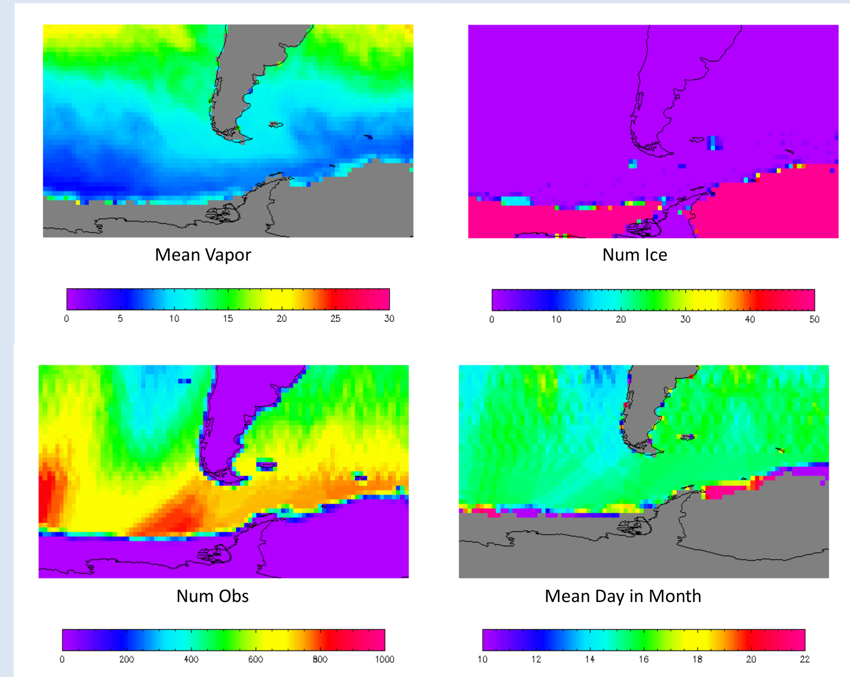


Figure 3. These four plots show the values retained from stage 1 in processing individual satellite water vapor data into a merged product. The num obs, num ice obs, and mean day of month are grid cell level q/c used in developing the best product for users. We used grid cells for which num ice <30, num obs > 160 and mean-day-in-month < 6 days from then center of the month.

Rain Flag Necessity: Scientist 2 included a +/- 1 cell extended rain flag and his vapor mean and trend were different from the other approaches (see Figure 4 difference map). The inclusion of a rain flag creates a geographic sampling problem as it removes data from primarily rainy, high vapor, tropical areas resulting in lower mean vapor and a higher global trend. Figure 4 shows a comparison of SSM/I F13 vapor to GPS vapor in rain-free and rainy conditions. Since little difference can be attributed to rain, we determined that an extended rain flag is not needed.

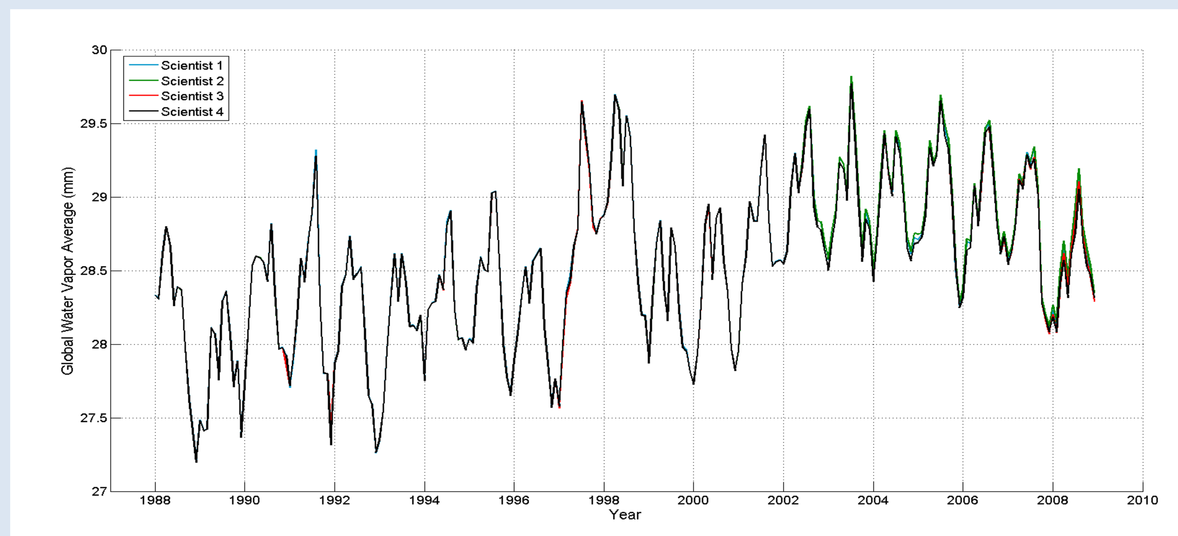


Figure 2. Time series of the same four data sets shown in Figure 1, only that data for a given grid cell was included in the calculations when it was present in all four data sets. This process essentially unified all decisions that affected geographical sampling (such as rain flagging, ice flagging, and what data were used/excluded). The remaining small difference visible after 2002 is due to small AMSR-E and WindSat offsets.

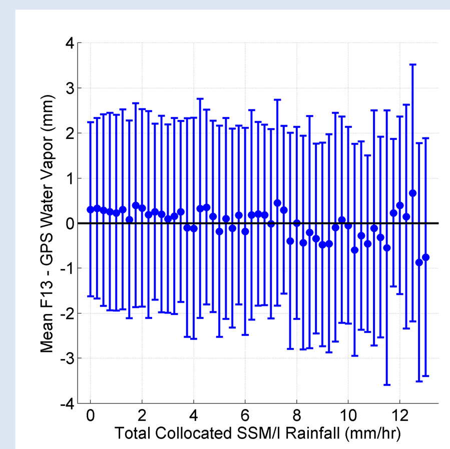


Figure 4. This comparison of SSM/I water vapor to ocean GPS island station water vapor contains both rain-free and raining collocations. The rainy comparisons show no greater difference in vapor .

Minor Impacts (if any)

Averaging Method: Three scientists used a drop-in-bucket averaging method. The fourth first determined monthly vapor averages and then created the merged vapor data set. There was virtually no difference between the methods. However, by calculating individual instrument monthly vapor values and then merging, we had better quality control over what data were added. We determined this approach to be best for creating the public product.

Remaining Instrument Offsets: We all used only RSS Version-7 water vapor values. At the time of this study, that included all SSM/I, F17 SSMIS, AMSR-E and WindSat. As shown in Table 1, some approaches used only SSM/I while others included AMSR-E and WindSat. Figure 2 shows increased differences between the four data sets after 2002 when AMSR-E data started. Some of the differences may be due to the later sampling time of AMSR-E (1am/pm as opposed to 6-9 am/pm of other radiometers), however some may be due to the fact that these two instruments were intercalibrated with rain-free data as opposed to using all data (rain-free and rainy). We believe these factors cause a slight difference requiring correction. We calculated offsets of AMSR-E (-0.2 mm) and WindSat (-0.05 mm) and applied these to the data.

Type of correction applied: As listed above, it was agreed that a correction was needed for both AMSR-E and WindSat data. We examined the differences created by using a simple additive offset or a more complicated multiplicative correction. The difference was negligible so we used the simpler additive correction.

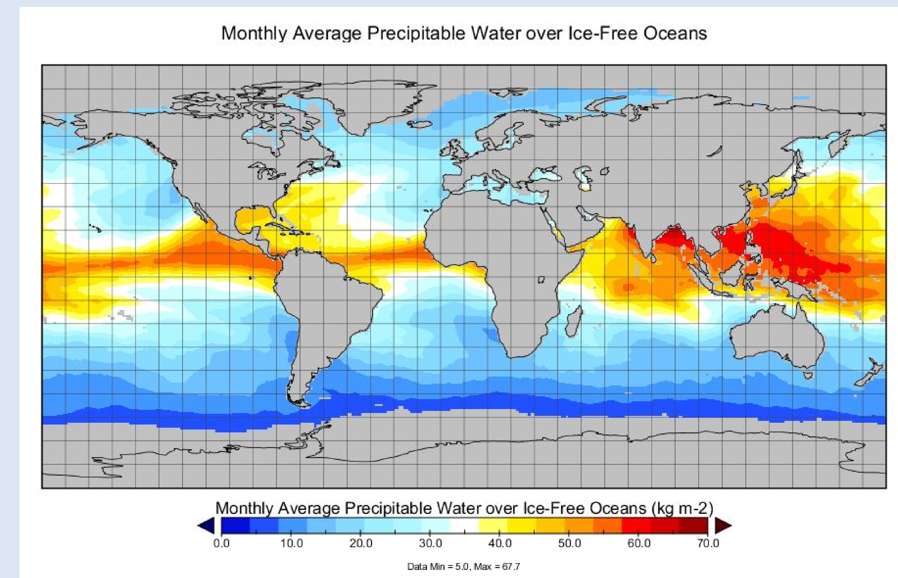
Guidelines for Merging Data from Individual Satellites into a Climate-Quality Product

1. Use intercalibrated data that are consistently processed throughout instrument operation and between instruments.
2. Determine a minimum number of observations for the given time period (month as used here) to remove poor quality data.
3. Keep track of ice and omit when present for more than a determined percentage of the time.
4. If any adjustments are needed, keep it simple.
5. Create yearly product for each satellite then merge the individual satellite products into one product using minimal data counts for each satellite. This allows for better quality control.

New Vapor Product at www.remss.com

From this exercise, we learned the lessons summarized in the Guidelines listed below. We took these lessons learned and applied them to create a new 1-deg, merged water vapor product for release to the public. We first created 1x1 deg maps of water vapor from individual satellites, keeping track of number of obs, number of ice obs and the mean day of month (as shown in Fig. 3). A mean water vapor value is only reported if the number of obs is greater than 160, number of ice obs is less than 30, and the (mean day of month - center of month) is less than 6 days. 3 satellite-months did not meet these requirements: F08 in Jan 1988, F08 in Oct 1990 and F10 in Dec 1991. In each of these cases, these month values are needed for consistency of the time series so an exception was made and the data were included. Quality checks and post-hoc corrections were applied in the second stage of processing. A set of 12 month climatology maps and monthly anomaly maps are then calculated from the monthly vapor values and added to the data set. In addition, we include a water vapor trend map and a hovmöller plot.

The RSS DISCOVER Water Vapor 1-deg Monthly Values Data Set will be available at ftp.remss.com/vapor and www.discover-earth.org/vapor in mid-December 2012. The file contains monthly mean water vapor maps for Jan 1988 to Nov 2012, 12 climatology water vapor maps (Jan - Dec), associated monthly anomaly maps (monthly water vapor minus climatology), vapor trend maps and a time-latitude (hovmöller) plot. The netCDF 4 file contains CF-compliant metadata. This water vapor product will be updated each month (for example, by Feb 15th for Jan 2013).



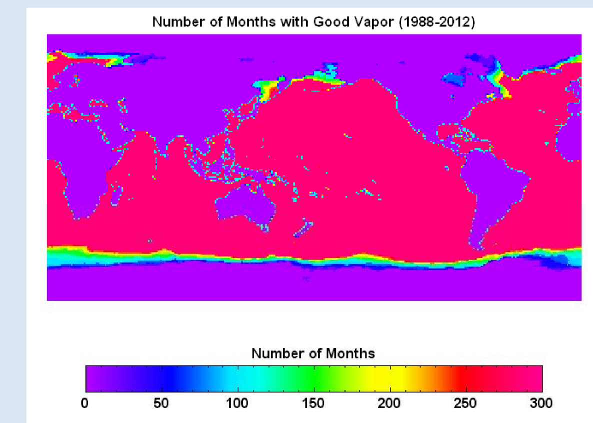


Figure 5. The number of months available in the data product. This plot shows the growth and recession of ice edges and variations in coverage near land boundaries. Quality checks removed grid cells for which less than 160 obs were available during a given month.

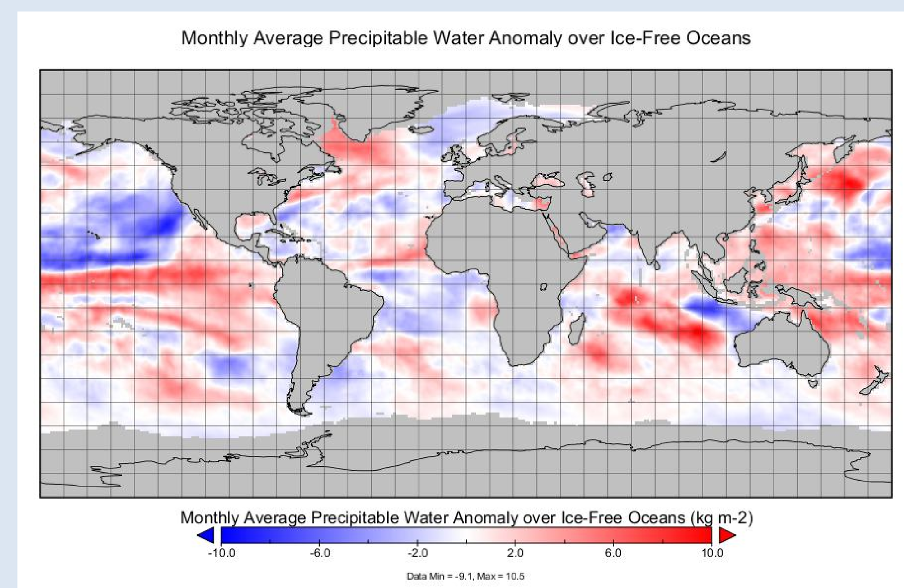


Figure 6. Vapor (left,top) and anomaly (left, bottom) plots for July 2012. Since the data product is in netCDF 4 format, a tool like Panoply can be used to easily display and access the data and metadata (http://www.giss.nasa.gov/tools/panoply**).**

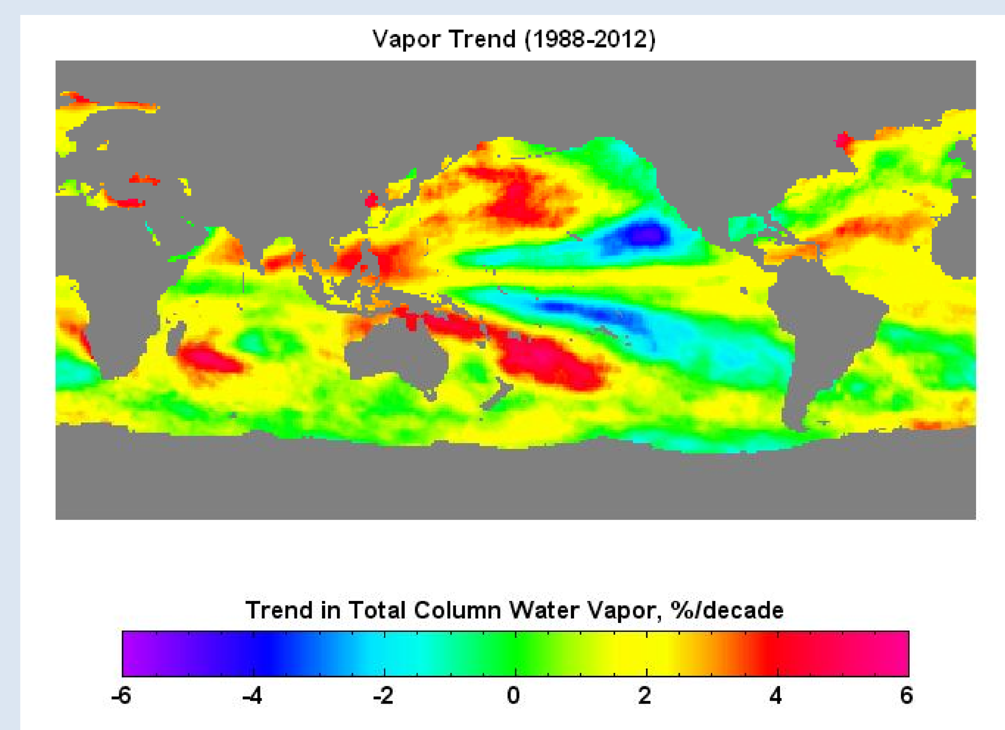


Figure 7. Water Vapor Trend Map (%/decade) available within the new data product. This trend map will be updated monthly along with all other data.