GEWEX/ESA DUE GlobVapour workshop on long term water vapour data sets and their quality assessment

8 – 10 March 2011
ESA/ESRIN, Frascati, Italy

Working Group Summaries

The GEWEX/ESA DUE GlobVapour workshop on long term water vapour data sets and their quality assessment was hosted by the European Space Agency ESRIN centre in Frascati (Rome), Italy from March 8-10, 2011. Thirty-five representatives from space agencies (ESA, EUMETSAT, NOAA), universities and research organizations (CLS, CNRS, CSU, DMI, FUB UCAR, UKMO), the ground-based community (JPL, GRUAN, U. Cologne), and the satellite retrieval (DLR, MPI, etc.) communities participated in the meeting organized by the GEWEX Radiation Panel and ESA GlobVapour project. The list of participants is given in Annex A and the workshop presentations are available online at http://due.esrin.esa.int/meetings/meetings247.php

The workshop consisted of one and one half days of presentations that described various activities in water vapour data set generation and validation, and emphasized the need to undertake a water vapour assessment to provide consistent and well-characterized data records for use across GEWEX. This was followed by one day of working group discussions on Total Column Water Vapour (TCWV) and Water Vapour Profiles (WVP) also considering data records providing deep layer humidity information often called upper or free tropospheric humidity. In light of the fact that many data sets of water vapour are currently being constructed with ECV/CDR characteristics, Jörg Schulz, from the GEWEX Radiation Panel, provided a summary of what questions and issues were relevant in on-going GRP assessments (clouds, aerosols and precipitation). A discussion framework towards that end was provided to the two working groups.

At the core of the assessment is GEWEX's need to gain insight into a number of water vapour products now being constructed for climate applications. Because each product can have slightly different users and objectives, it was thought important to clearly describe the data set objectives in the introduction to the assessment in order to place each product in the proper perspective relative to climate needs.

Because of the distinct objectives, it was also deemed counterproductive to rank algorithms in terms of quality alone, as the usefulness of a product ultimately depends upon the application. The assessment instead should focus not only on the quality of each product, but on overall characteristics of each product as determined from comparisons against various in-situ observations, comparisons with other products, analyses of other properties such as the representation of the diurnal cycle, dependence on a-priori information, etc.

Breakout Leads

TCWV
Lead: Marc Schröder (DWD), **Rapporteur:** Janice Bytheway (CSU)

WVP
Lead: Antonia Gambacorta (NOAA)/Thierry Leblanc(JPL), **Rapporteur:** Martin Stengel (DWD)

Dataset name/identifier

General comments: With regards to "which products" to assess, the panel thought it important to define precisely what should be part of the assessment and what not. In the end, there was consensus that total column water vapour plus water vapour and temperature profiles should be assessed. The temperature, however, was seen as the temperature related to the vapour profile. As such, the assessment does not include stand-alone temperature profiles or deep layer temperature data sets such as derived from MSU/AMSU. It also does not include SST/LST unless these are integral parts of the Vapour profile. Nor should the assessment include 2m temperature or humidity unless, again, these levels are part of an overall water Vapour profile. However, it does include UTH/FTH data sets.

Datasets: A list of potential products that may be assessed is given in the following table. The first activity of the assessment chairs will be to contact potential partners for interest in these activities. The assessment will include products from all willing participants in the current activity. The assessment may include products from passive individuals or groups if the product is generally used, fully published, and freely available without restrictions. This, in particular, applies to existing data records coming from various NWP model-based re-analysis activities. Table 1 provides a first preliminary list of data sets that can be provided by the workshop participants.

Table 1: Preliminary list of water vapour data sets as identified by the workshop participants.

Dataset name or identifier	TCWV	WVP	UTH/FTH	Producer
NASA MEaSURES	x	x	-	Tom VonderHaar (Colorado St. U.)
AIRS (cloud cleared, clear-sky)	x?	(x) 3 layers	-	
IASI (cloud cleared, clear-sky)	x?	x	-	Many, ref. to GlobVapour assessment
MLS 316Pa and above		x		Bill Read (JPL)
TES	x	x		
SEVIRI/MVIRI		(x)	x	EUMETSAT CM-SAF, LMD
GOME/SCIAMACHY	x	-	-	Thomas Wagner

Table 1: Preliminary list of water vapour data sets as identified by the workshop participants.

Dataset name or identifier	TCWV	WVP	UTH/FTH	Producer
				(MPI/GlobVapour)
GOME-2	x	-	-	Diego Loyola (DLR, EUMETSAT-O3M SAF)
MODIS	x	x	-	NASA
MERIS	x	-	-	Rasmus Lindstrot (FUB/GlobVapour)
RO (COSMIC, GRAS, CHAMP, GRACE)	x	x	-	EUMETSAT GRAS-SAF
ATOVS	(x)	x		EUMETSAT CM-SAF
HIRS	(x)	x	x	NOAA NCDC
AMSU-B/MHS	-	-	x	Viju John (UKMO)
AIRS	x	x	-	NASA
AATSR	x	-	-	René Preusker (FUB/GlobVapour)
SSM\I, SSMIS	x	-	-	Marc Schröder (DWD/CM SAF)

Validation data

The discussion groups felt that an important service to the community done by the water vapour assessment activity would be to build and maintain a total column water vapour and a water vapour profile database that includes data thought to be of sufficient quality and long-term stability to be of use to the assessment and future producers. GRUAN could be an important part of this activity but until such databases are fully available, the following table represents our best estimate at this time.

Table 2: Preliminary list of validation data sets as identified by the workshop participants.

Dataset	Contact
NOAA frost point launched at Boulder since 1980.	Sam Oltmans (NOAA GRL)
CFH sondes (since 1996/1998?)	Holger Vömel (DWD)
ARSA 1989 radiosondes. 20-30 stations.	Claudia Stubenrauch (LMD)
Microwave radiometers	Mario Mech (U Köln) + ?
Surface based GPS	EUMETSAT GRAS-SAF, US ?
NDACC Raman LIDAR	Thierry Leblanc (JPL)
Radio occultation data (COSMIC/Gras)	Hans Gleisner (DMI)
FTIR (Fourier Transform Infrared) spectrometers	Matthias Schneider (KIT)
BSRN	Gert Koenig-Langlo (AWI) http://www.bsrn.awi.de/en/home/
CERES	http://ceres.larc.nasa.gov/dpc_current.php

Products may incorporate one or more of the above products into their global products (e.g. GPCP incorporates rain gauges into the rainfall product when these are available). It is merely necessary that the validation data be independent of the product.

Validation Strategy

There is a need to start with a documentation of the various Level 1 datasets used to produce water vapour datasets. While this is not generally viewed as part of the assessment itself, the L1 stability and usefulness for climate data records should be documented via the data providers and verified as best as possible.

Level 2 data should be compared to the validation data described above. These comparisons should allow for the characterization of product uncertainties as well as a product's strengths and weaknesses. Suitability for climate trends and product differences are difficult to assess from Level 2 comparisons alone. An important additional validation step at this level is the use of radiative fluxes, which can be, computed both in the up- and downwelling directions. These, in turn can be compared directly to BSRN and CERES flux observations.

Level 3, monthly global and regional intercomparisons are most useful to elucidate differences among products and their root causes. Level three comparisons should focus on parameters such as pdfs in addition to means. Radiation, that is, reacts to the pdf of water vapour rather than the mean. Level 3 validation data may be added also if deemed useful at these space/time scales.

There was an agreement that the assessment should be performed in two phases. Phase I begins with three years of recent data – where more satellite data sets and more validation data sets are available. This should shed some light on the properties of the longer term data sets as well that will be useful when longer time series with less validation data are assessed in a Phase II.

Specific types of errors and bias

The working groups also discussed the type of expected errors, in particular, systematic errors that lead to differences of water vapour Level 3 data sets. The following list based on work of the GEWEX aerosol assessment was established:
- **Radiometric Bias:** Biases due to uncharacterized or ill-applied sensor calibration.
- **Retrieval Bias:** Biases related to shortcomings in a retrieval itself.
- **Sampling/Contextual Bias:** Biases related to where a retrieval is/is not performed or contextually related uncertainty in a scene. This leads to a skewed data population relative to what is thought to have been collected.
- **Aggregation/Data Reduction Bias:** Loss of required information during conversion to a gridded product or during analysis.
- **Cognitive Bias:** We, the investigators, misinterpret, withhold, or frame data/results contrary to full nature of the data.

- **Other Considerations:** a) Correlated error-"Independent" products that share similar biases; b) Tautology –Circular reasoning or treating non-independent data as independent during tuning.
- **Miscellaneous Bias:** a) Biases from uneven topography adding vapour below profile sensitivity; b) Diurnal biases; c) A-priori biases, d) Spectroscopy errors, e) *etc.*

Experiment design

There was some discussion that the details could be left for later. General thought related to starting with general product overviews to ensure overall matching of expectations. Other assessments have found errors in this step already.

Technical Issues

While this was not the focus of the workshop, experiences from other assessments show that it is useful to define common terms of reference regarding datasets in order to facilitate exchanges of data and analysis software. Some dataset attributes that would make analyses easier in the current community consist of:
- NetCDF with CF standards w. Global Attributes.
- Consistent Versioning.
- Readily available Level 2 readily deliverable.
- Appropriate flags.
- Uncertainty information.
- Available a-priori input information and relative weight in output.
- Weighting functions for profiles.

Volunteers

Thomas August (EUMETSAT): review of the availability of uncertainty estimate description for IR hyper-spectral sounders.
Bojan Bojkov (ESA): will look into best ways to gather collocation information between satellites and GB validation datasets.
Antonia Gambacorta (NOAA): volunteered to start coordinating the profile assessment (needs to be confirmed).
Shu-Peng Ho (NCAR): volunteered to start coordinating the profile assessment (needs to be confirmed).
Klemens Hocke (UB): to ask in his institute if colleagues are willing to contribute.
Thierry Leblanc (JPL): organize web server to list potential validation data sets.
Remy Roca (LMD): volunteered to perform Level 1 data comparisons related to UTH/FTH and to provide general technical support to the assessment.
Matthias Schneider (KIT): validation of IASI to FTIR.
Lei Shi (NOAA): coordinating UTH/FTH activities (needs to be confirmed).
Marc Schröder (DWD): volunteered to start coordinating the total column assessment (funding needs to be confirmed).

Timetable

Try to finish phase I in 2-3 years. First deadline is GRP meeting in Tokyo (30 Aug. – 02 Sept). The participants agreed that a good start would be to have this document fully elaborated and product developers contacted regarding potential participation.

Annex I: Workshop Participants

Name	Surname	E-mail	Affiliation	Country
Olivier	Arino	oliver.arino@esa.int	ESA/ESRIN	IT
Thomas	August	thomas.august@eumetsat.int	EUMETSAT	DE
Michael	Berger	Michael.Berger@esa.int	ESA/ESRIN	IT
Bojan	Bojkov	bojan.bojkov@esa.int	ESA/ESRIN	IT
Janice	Bytheway	janice@atmos.colostate.edu	Colorado State University	USA
Stefano	Casadio	stefano.casadio@esa.int	Serco and ESA/ESRIN	IT
Antonio	Di Noia	dinoia@disp.uniroma2.it	University Tor Vergata (Rome)	IT
Mark	Doherty	mark.doherty@esa.int	ESA/ESRIN	IT
Thorsten	Fehr	Thorsten.Fehr@esa.int	ESA/ESRIN	IT
Jürgen	Fischer	fischer@zedat.fu-berlin.de	Free University of Berlin (FUB)	DE
Antonia	Gambacorta	antonia.gambacorta@noaa.gov	NOAA	USA
Hans	Gleisner	hgl@dmi.dk	Danish Meteorological Institute	DK
Mitch	Goldberg	mitch.goldberg@noaa.gov	NOAA	USA
Shu-peng Ben	Ho	spho@ucar.edu	UCAR	USA
Klemens	Hocke	klemens.hocke@iap.unibe.ch	University of Bern	CH
Chris	Kummerow	kummerow@atmos.colostate.edu	Colorado State University	USA
Thierry	Leblanc	leblanc@tmf.jpl.nasa.gov	NASA	USA
Diego	Loyola	Diego.Loyola@dlr.de	DLR	DE
Mario	Mech	mech@meteo.uni-koeln.de	University of Cologne	DE
Yasjka	Meijer	Yasjka Meijer/esrin/ESA	ESA/ESRIN	IT
Bruno	Picard	bpicard@cls.fr	CLS	FR
Simon	Pinnock	simon.pinnock@esa.int	ESA/ESRIN	IT
René	Preusker	rene.preusker@wew.fu-berlin.de	Free University of Berlin (FUB)	DE
Roca	Rémy	roca@lmd.jussieu.fr	CNRS	FR
Mark	Ringer	mark.ringer@metoffice.gov.uk	UK Met Office	UK
Matthias	Schneider	matthias.schneider@kit.edu	Karlsruhe Institute of Technology (KIT)	DE
Marc	Schröder	marc.schroeder@dwd.de	Deutscher Wetterdienst	DE
Jörg	Schulz	Joerg.Schulz@eumetsat.int	EUMETSAT	DE
Lei	Shi	Lei.Shi@noaa.gov	NOAA	USA
Sander	Slijkhuis	Sander.Slijkhuis@dlr.de	DLR	DE
Theo	Steenbergen	theo.steenbergen@dwd.de	Deutscher Wetterdienst	DE
Martin	Stengel	martin.stengel@dwd.de	Deutscher Wetterdienst	DE
Claudia	Stubenrauch	stubenrauch@lmd.polytechnique.fr	CNRS / IPSL LMD	FR
Rolf	von Kuhlmann	Rolf.Von.Kuhlmann@esa.int	ESA/ESRIN	IT
Thomas	Wagner	thomas.wagner@mpic.de	MPI for Chemistry	DE