RMetS
Royal Meteorological Society

Expected improvements in the atmospheric humidity profile retrieval using the Megha-Tropiques microwave payload

Hélène Brogniez,[a]* Pierre-Emmanuel Kirstetter[a†] and Laurence Eymard[b]
[a]*Laboratoire Atmosphère, Milieux, Observations Spatiales, IPSL/UVSQ/UPMC/CNRS, Guyancourt, France*
[b]*Laboratoire d'Océanographie et du Climat: Expérimentation et Approches Numériques, IPSL/UPMC/IRD/MNHN/CNRS, Paris, France*

*Correspondence to: H. Brogniez, LATMOS, 11 bvd d'Alembert, 78280 Guyancourt, France.
E-mail: helene.brogniez@latmos.ipsl.fr
[†]Now at National Severe Storms Laboratory/Cooperative Institute for Mesoscale Meteorological Studies, NOAA/University of Oklahoma, Norman, OK, USA.

The microwave payload of the Megha-Tropiques mission is explored to quantify the expected improvements in the retrieval of relative humidity profiles. Estimations of the profiles are performed using a generalized additive model that uses cubic smoothing splines to address the nonlinear dependencies between the brightness temperatures (T_B) in the 183.31 GHz band and the relative humidity of specified tropospheric layers. Under clear-sky and oceanic situations, the six-channel configuration of the SAPHIR radiometer clearly improves the retrieval and reduces by a factor of two the variance of the residuals with respect to the current space-borne humidity sounders that have three channels in this band (AMSU-B, MHS). Additional information from the MADRAS radiometer (at 23.8 and 157 GHz) further improves the restitution with correlation coefficient higher than 0.89 throughout the troposphere. Copyright © 2011 Royal Meteorological Society

Key Words: water vapour; retrieval method; microwave; Megha-Tropiques

Received 15 February 2011; Revised 23 October 2011; Accepted 23 November 2011; Published online in Wiley Online Library 21 December 2011

Citation: Brogniez H, Kirstetter P-E, Eymard L. 2013. Expected improvements in the atmospheric humidity profile retrieval using the Megha-Tropiques microwave payload. *Q. J. R. Meteorol. Soc.* **139**: 842–851. DOI:10.1002/qj.1869

1. Introduction

The study of the tropical water cycle and of its evolution under climate modifications strongly relies on a documentation of the horizontal and vertical distribution of the water vapour field and on the understanding of the different processes that interact with it at all scales (Roca *et al.*, 2010). The observation from space of the free tropospheric humidity began in the early 1970s with the launch of the Nimbus-4 satellite, which carried the Temperature–Humidity Infrared Radiometer (THIR), with a specific channel in the 6.3 μm strong absorption band (Allison *et al.*, 1972). A continuous monitoring of the atmospheric water vapour has been provided since 1983 by the series of METEOSAT geostationary satellites and analyses of this database have revealed specific features of the tropical tropospheric water vapour, be it at the scale of the convective event (e.g. Roca *et al.*, 2005) or at the climatological scale (e.g. Brogniez *et al.*, 2009). Moreover, a global view of the atmospheric humidity content is brought by the succeeding NOAA polar orbiting satellites carrying the HIRS radiometer since 1979, allowing trend studies of the upper tropospheric humidity (e.g. Bates and Jackson, 2001; McCarthy and Toumi, 2004). However, such measurements are limited to clear-sky areas and scenes of low cloud cover (e.g. Brogniez *et al.*, 2006), which limits the comprehension of the processes at play in the atmospheric water cycle.

Microwave techniques give the opportunity to observe the Earth's surface and its atmosphere from space even in the presence of clouds, which are largely transparent at frequencies below 100 GHz. Indeed, considering the mean size of water droplets and ice crystals (∼10–100 μm) with respect to the wavelength at microwave frequencies (∼1 cm, for $\nu >$ 60 GHz), the effect of scattering by non precipitating particles

can be reasonably well represented by the Rayleigh approximation in the radiative transfer. Recent efforts towards the exploitation of 183.31 GHz observations have been made, dedicated to retrieval algorithms (Sohn *et al.*, 2003; Houshangpour *et al.*, 2005), to process studies (e.g. Brogniez and Pierrehumbert, 2006; Eymard *et al.*, 2010) or to the elaboration of intercalibrated long-term datasets (Buehler *et al.*, 2008). Indeed, the use of microwave radiometer allows for the study of atmospheric humidity close to the deep convective core, the spatial resolution still being a challenge for the processes at play in between the rainy convective cells: the current space-borne microwave sounders dedicated to water vapour observations have a footprint of 16 km at nadir (Microwave Humidity Sounder (MHS) onboard the MetOp platforms and Advanced Microwave Sounding Unit-B (AMSU-B) onboard the NOAA satellites), while the infrared radiometers offer a view of the water vapour field with a sub-satellite point of up to 3 km (example of the SEVIRI imager onboard the METEOSAT Second Generation series).

The Indian–French Megha-Tropiques mission (Roca *et al.*, 2012, this issue) will sample the tropical belt within $\pm 30^\circ$ in latitude. The payload includes two microwave radiometers dedicated to the atmospheric hydrological cycle: MADRAS (Microwave Analysis and Detection of Rain and Atmospheric Structures) is a conical scanning imaging radiometer with channels ranging from 18 to 157 GHz; SAPHIR (Sondeur Atmosphérique du Profil d'Humidité Intertropicale par Radiométrie) is a cross-track scanning sounder operating in the 183.31 GHz water vapour band, with a footprint at nadir of 10 km. The channels' central frequencies and bandwidths of SAPHIR and MADRAS are listed in Table 1, along with those from AMSU-B and MHS.

In this paper we present the expected improvements in the estimation of the relative humidity profile brought by the combination of those two instruments with respect to the current operational platforms carrying similar radiometers. For this purpose, the method of restitution of the profiles is based on a generalized additive model that addresses the strong nonlinearity of the problem with additive cubic smoothing splines, and the training of the model is accomplished with a 10-year set of radiosoundings limited to oceanic and purely non-diffusive profiles. This exercise finally discusses the interest of estimating the whole distribution of the possible profiles given a set of explanatory variables, such as the brightness temperatures (T_B), and not only its expectation. Such a retrieval could be a guide for a new type of operational product.

The rest of this paper is organized as followed. Section 2 briefly recalls the information given by observations in the 183.31 GHz water vapour absorption band, section 3 is dedicated to a description of the data and tools used for the study, while section 4 provides a short discussion on the choice of training dataset for the specific problem of relative humidity profile retrieval. Section 5 discusses the statistical model used to perform the retrieval and its results in a comparative study on the number of inputs. Finally, a conclusion is drawn in section 6.

Table 1. Channel central frequencies and bandwidths for AMSU-B, MHS, SAPHIR (all channels are vertically polarized at nadir) and for MADRAS. The '–' symbol indicates that the same channel is present.

	Central frequency (GHz)		Bandwidth (MHz)	
AMSU-B & MHS	89.0	–	±1000	±1400
	150.0	157.0	±1000	±1400
	183.31 ± 1.0	–	±500	±250
	183.31 ± 3.0	–	±1000	±500
	183.31 ± 7.0	190.31	±2000	±1100
SAPHIR	183.31 ± 0.2		±200	
	183.31 ± 1.1		±350	
	183.31 ± 2.8		±500	
	183.31 ± 4.2		±700	
	183.31 ± 6.6		±1200	
	183.31 ± 11.0		±2000	
MADRAS	18.7 (H & V)		±100	
	23.8 (V)		±200	
	36.5 (H and V)		±500	
	89.0 (H and V)		±1350	
	157.0 (H and V)		±1350	

2. Information content of the 183.31 GHz band

2.1. Background

The use of microwave radiometry to estimate the Earth's tropospheric water vapour has spread since the launch of the Nimbus-E microwave spectrometer in the 1970s (Staelin *et al.*, 1976). This satellite was designed for the study of atmospheric water vapour and liquid abundances from the 22.235 GHz water vapour rotational transition, using the open ocean as a cold background. Although this line is very well adapted for the measurement of the total water vapour burden, measurements at several frequencies around the stronger rotational transition line at 183.31 GHz (see Figure 1) are better suited to estimate the vertical relative humidity profile over land and oceans (Schaerer and Wilheit, 1979; Rosenkranz *et al.*, 1982; Kakar, 1983; Wang *et al.*, 1983). One also sees in Figure 1 the absorption due to the water vapour continuum, which increases significantly with frequency.

The reference study by Schaerer and Wilheit (1979) on the use of selected frequencies in a 18 GHz bandwidth around the 183.31 GHz water vapour line clearly showed the theoretical feasibility of atmospheric profiling of relative humidity

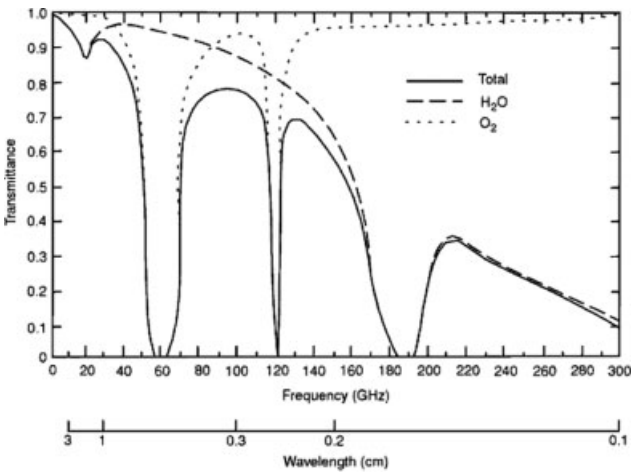

Figure 1. Atmospheric transmittances (total, H_2O and O_2) as a function of frequency and wavelength in the microwave region. Reprinted with permission from Liou (2002).

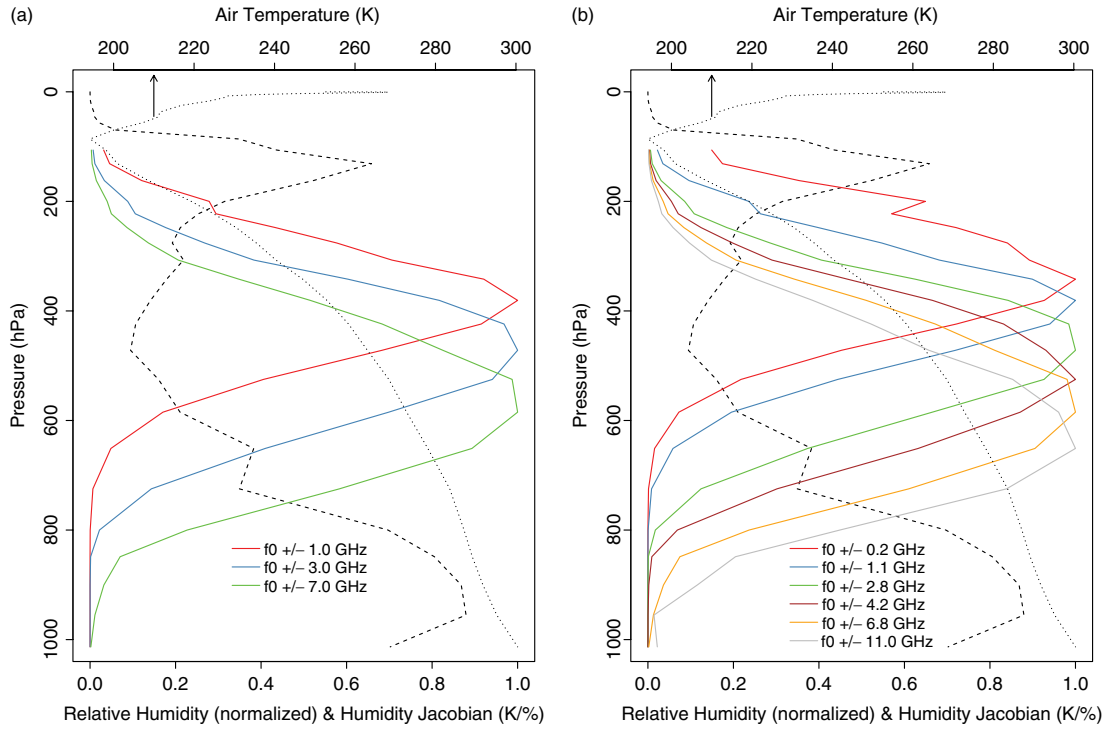

Figure 2. Normalized sensitivity function (i.e. relative humidity Jacobians $\partial T_B/\partial RH$, coloured lines, in K/%) for a tropical relative humidity profile (dashed line) and temperature profile (dotted line), estimated (a) for AMSU-B-like sounders and (b) for the SAPHIR sounder. Radiative transfer computations were performed with the RTTOV model, for a nadir viewing angle.

from a passive radiometer over both land and oceanic surfaces. This was confirmed with airborne measurements by the Advanced Microwave Moisture Sounder (AMMS) during the CCOPE campaign (Cooperative Convective Precipitation Experiment; Wang *et al.*, 1983), with three channels in the 183.31 GHz band (±2.25, ±5.0 and ±8.75 GHz; use of the double sideband mode for a better radiometric sensitivity). This experiment helped in designing the current configuration of the space-borne water vapour sounders (e.g. AMSU-B; Saunders *et al.*, 1995; Table 1) with channels at 183.31 ± 1.0, 183.31 ± 3.0 and 183.31 ± 7.0 GHz, which were shown to be a more favourable distribution of the weighting functions for a retrieval at relatively high altitudes (∼12 km; e.g. Wang and Chang, 1990).

Several recent humidity sounding radiometers share the same spectral characteristics, such as the Special Sensor Microwave/Temperature-2 (SSM/T-2) instrument onboard the DMSP satellites or the recent Microwave Humidity Sounder (MHS) onboard the MetOp platform. SAPHIR slightly differs from these sensors since it has six channels ranging from 183.31 ± 0.2 to 183.31 ± 11.0 GHz (see Table 1 for the characteristics of both AMSU-B and SAPHIR). A detailed description regarding the instrument is found in Roca *et al.*(2012, this issue) and in a preliminary study by Eymard *et al.*(2002). The first channel, close to the centre of the 183.31 GHz band, is aimed at reaching higher layers of the atmosphere compared to AMSU-B, and the use of a large bandwidth of 11.0 GHz around the central frequency locates the sixth channel on the wings of the absorbing line for a deeper sounding of the atmosphere. The usual observations in the 150 GHz window channel generally used to remove the surface contribution (present on AMSU-B; a 157 GHz channel is on MHS; see Table 1) are provided by the MADRAS instrument, with two 157 GHz channels measuring the two polarizations.

Figure 2 shows the sensitivity functions of the different observing channels ($\partial T_B/\partial RH$, i.e. the relative humidity Jacobians) for a standard tropical relative humidity profile for AMSU-B-like sounders and for the SAPHIR sounder. Sensitivity functions of channels located near the centre of the 183.31 GHz water vapour line peak higher in the atmosphere than those of channels that are farther from the line centre. The shape and behaviour of the sensitivity functions reveal the nonlinearity of the retrieval problem: as an absorbing component is added to the atmosphere, brightness temperatures decrease and the peaks of all channels shift upward, whereas for a drier atmosphere the functions shift downward (e.g. Schaerer and Wilheit, 1979; Blankenship *et al.*, 2000). In that spectral area the Rayleigh–Jeans approximation directly links the measured brightness temperature to the vertical profile of the atmospheric opacity (i.e. the absorber amount) and to the thermal structure. Without any *a priori* information on the temperature profile, the distinction between the two contributions is delicate, yielding to estimate directly the relative humidity profile, although in the Tropics the variability of the temperature field plays a minor role in the relative humidity variability (Peixoto and Oort, 1996).

2.2. *Algorithms overview*

Various inversion methods are used to retrieve the relative humidity profile from the measurements in the 183.31 GHz band: neural networking (e.g. Cabrera-Mercadier and Staelin, 1995; Karbou *et al.*, 2005), iterative schemes (e.g. Wang *et al.*, 1983; Wang and Chang, 1990; Wilheit and Al-Khalaf, 1994; Blankenship *et al.*, 2000; Liu and Weng, 2005) and multivariate regression methods (e.g. Rosenkranz *et al.*,

1982; Kakar and Lambrigsten, 1984). In order to solve the inverse problem, all of the methods need to deal with the strong nonlinearities of the exercise (from the variability of the sensitivity function; see section 2.1) and with the singularities (such as temperature inversions and isothermal areas) that may arise and thus combine statistical tools and physical constraints.

Both multivariate regression methods and neural network algorithms are built on a training dataset that is assumed to be representative of the atmosphere. The choice of training dataset is thus essential since it holds all the physics of the problem. Iterative schemes adjust the estimated relative humidity profile using successive perturbations on an *a priori* state of the atmosphere (selected, for example, using a Bayesian approach; Rieder and Kirchengast, 1999; Rosenkranz, 2001) and the physical constraint comes from radiative transfer computations, the best profile minimizing a cost function (e.g. Wilheit and Al-Khalaf, 1994).

For the Megha-Tropiques mission, the operational algorithm is a neural network algorithm (a multilayered perceptron) that combines the SAPHIR and MADRAS measurements to retrieve a six-layer relative humidity profile (for details of the method see Aires *et al.*, 2012, this issue). To physically constrain the retrieval, *a priori* information from ECMWF analyses (collocated temperature profile and surface temperature over land), and a climatology of microwave land surface emissivities derived from Special Sensor Microwave/Imager (SSM/I) observations complete the input vector. In the present study, we decided to use an alternative approach (see details in section 5.2) to this operational algorithm in order to suggest new scientific possibilities, which will be discussed.

3. Datasets

As stated earlier, this study aims at showing the expected gain of the SAPHIR and MADRAS combination and therefore focuses on the Tropics, limited to $\pm 30^\circ$ around the Equator. For this purpose, we consider only the oceanic situations (emissivity around 0.4, corresponding to a cold background) in order to avoid the problem of a strong contribution of the land surfaces to the upwelling radiation at these frequencies (emissivity typically between 0.8 and 1) (e.g. Bennartz and Bauer, 2003; Karbou *et al.*, 2010).

Two types of training database are generally used for the design of a retrieval scheme: the atmospheric profiles are either extracted from quality-controlled radiosoundings (see Wang and Chang, 1990; Cabrera-Mercadier and Staelin, 1995; among others) or from analyses or reanalyses (NCEP-NCAR or ECMWF models, e.g. Eymard *et al.*, 1993; Shi, 2001; Zhang *et al.*, 2008; Aires *et al.*, 2012, this issue). Both sources of information are studied, section 4 being dedicated to the selection of the most adapted training dataset for the present problem.

3.1. Clear-sky or all-sky atmospheres?

Seeing through clouds is the big advantage of microwave soundings over infrared soundings. Contamination by clouds is largely negligible in the 183.31 GHz band, although a significant depression occurs from scattering of the upwelling radiation back to the surface by precipitating particles of deep convective clouds and cirrus anvil clouds (e.g. Rosenkranz *et al.*, 1982; Burns *et al.*, 1997; Greenwald and Christopher, 2002). Methods that filter out the diffusive cases exist and allow interpretation of the 183.31 GHz T_B in terms of relative humidity only (e.g. Hong *et al.*, 2005; Buehler *et al.*, 2007). However, in the present case where no observed T_Bs are yet available, the development of a realistic training database requires association of a synthetic T_B to each thermodynamic profile, which is done using a radiative transfer model (see section 3.4).

To our knowledge, there is no database of thermodynamic profiles that also includes *in situ* measurements of the cloud profile (i.e. the liquid and ice water contents). To overcome this problem, cloud profiles could be extracted from the ECMWF archive, but a recent evaluation study (Delanoe *et al.*, 2011) shows that, with respect to the ice water path (IWP) derived from joint observations by CloudSat and CALIPSO, the ECMWF model overestimates very thin clouds and underestimates the global IWP. The depression in the T_B induced by ice particles being a strong function of microphysics (Liu and Curry, 1996), it would be unwise to use the ECMWF cloud profiles to produce all-sky T_Bs.

For this reason, we decided to concentrate on the case of non-diffusive atmospheres in order to avoid any bias in the computation of synthetic T_Bs that could inevitably translate into a bias of retrieval, not due to the retrieval algorithm itself but to the inputs of the algorithm.

3.2. Radiosounding measurements

Radiosondes are extracted from the ECMWF operational sounding archive used during the assimilation process in the ECMWF reanalyses. The data (pressure, temperature and moisture) have been analysed and underwent quality control in order to discard incomplete profiles (threshold of 30 hPa for temperature and 350 hPa for humidity), and a vertical extrapolation was applied up to the top of the atmosphere (2.10^{-3} hPa) using climatology. Finally, the measurements are interpolated on a fixed-pressure grid (the ARSA database, developed by the Laboratoire de Météorologie Dynamique; see http://ara.abct.lmd.polytechnique.fr/index.php?page=arsa). The selected clear and oceanic cases cover the 1996–2006 period and form a base of about 1400 profiles, limited for the present exercise to 21 levels in the troposphere (from the surface up to 100 hPa).

3.3. ECMWF reanalyses

ERA-Interim is the latest global atmospheric reanalysis produced by the ECMWF, covering the period 1989 onward (Berrisford *et al.*, 2009) and provides a bridge between the former reanalysis ERA-40 (1957–2002) and the next-generation reanalyses planned at ECMWF. The assimilating model represents the basic dynamical fields on 60 vertical hybrid levels from the surface up to 0.1 hPa. Its main improvements concern many aspects, among them the use of variational bias correction for satellite data and a revised humidity analysis. In the present study, we selected clear-sky and oceanic temperature and humidity profiles from the reduced resolution grid ($1.125^\circ \times 1.125^\circ$) for the months of July and December 2001. This sample represents about 40 000 profiles, with equal weight for both months. Finally, in order to easily compare these profiles to the radiosounding

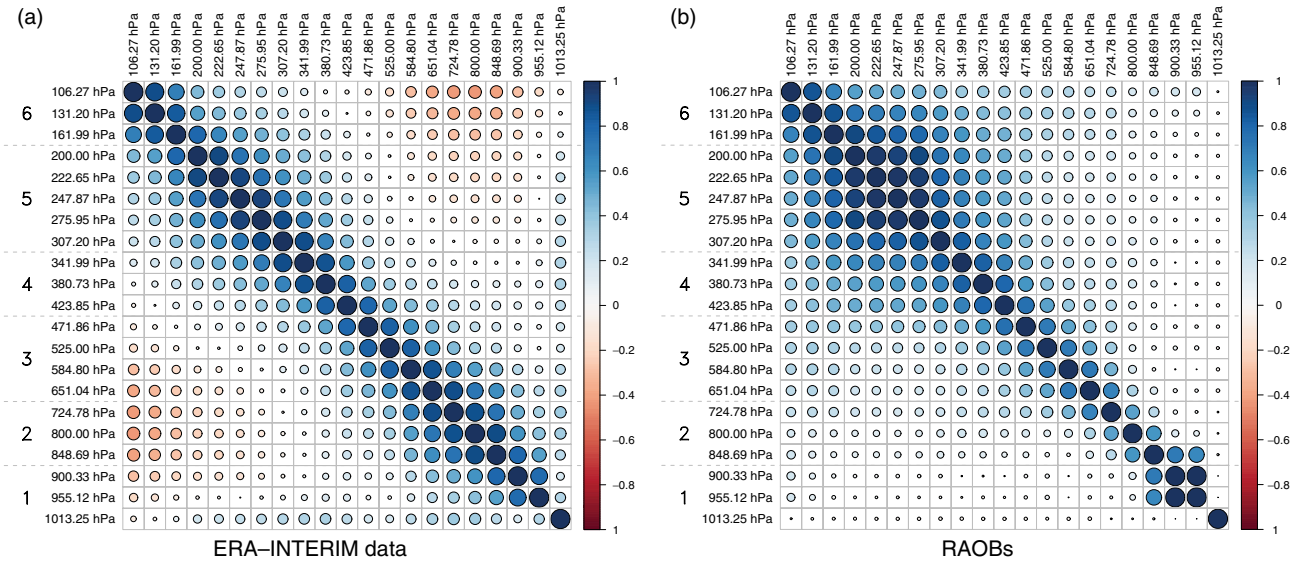

Figure 3. Representation of the vertical correlation between pressure levels of relative humidity profiles provided by (a) the ERA-Interim model for July and December 2001 (∼40 000 profiles), and (b) by an analysed radiosoundings archive covering 1996–2006 (>1400 profiles). For the purpose of comparison, the ECMWF profiles have been projected on to the radiosounding levels.

profiles, they have been interpolated from the hybrid model levels to the fixed-pressure grid of the latter.

3.4. The RTTOV radiative transfer model

The Radiative Transfer for Television and Infrared Observation Satellite Operational Vertical Sounder (TIROS-OVS) model (RTTOV version 9.3, supported by ECMWF NWP SAF Saunders *et al.*, 1999; Matricardi *et al.*, 2004) is used to simulate AMSU-B, SAPHIR and MADRAS instruments. This model is based on a regression scheme of the transmittances and thus allows for rapid simulations of brightness temperatures for satellite infrared and microwave radiometers for a given atmospheric state. The oceanic emissivities are computed by the FASTEM-3 surface model (Deblonde and English, 2001). For the present study, AMSU-B and SAPHIR brightness temperatures are simulated at nadir viewing angle, while for MADRAS the simulations are performed at the viewing angle of 53° (conical scan).

4. Discussion on the training dataset

The use of reanalysis profiles for training purposes in a retrieval problem is more and more widespread since they offer a larger spatial and temporal coverage of the system, thus forming statistically representative training databases of various atmospheric situations. However, while such databases are very well suited for the study of integrated contents of water vapour (Brogniez *et al.*, 2009; Obligis *et al.*, 2009), their use for the purpose of the vertical distribution of relative humidity is more delicate. Indeed, recent comparisons between satellite observations of single-layer upper tropospheric humidity (∼250 hPa, estimated from HIRS and AIRS) and reanalyses (ERA-40 and NCEP) have highlighted discrepancies at the interannual scale (Huang *et al.*, 2005; Chuang *et al.*, 2010) that are reduced in the ERA-Interim model (Dee and Uppala, 2009). However, on the vertical the ERA-Interim humidity field still exhibits noticeable inconsistencies in the tropical water and energy budget (Chiodo and Haimberger, 2010; Eymard *et al.*, 2010), indicating the need for special care when using profiles of moisture from reanalyses.

The vertical differences between profiles from reanalyses and profiles from *in situ* observations are illustrated in Figure 3, which shows the vertical correlations between the various levels of relative humidity provided by ERA-Interim and by the radiosounding record. The structures of the correlation matrices reveal noticeable differences between the two sets of profiles: the relative humidity profiles of the radiosoundings show a strong correlation between adjacent levels of the upper troposphere (∼400–100 hPa), which is not present in the ERA-Interim set, while the latter show a weak pattern of correlation throughout the free troposphere. The differences of patterns in the correlation matrices suggest that some information on the mechanisms that regulate the upper tropospheric humidity (roughly 350–100 hPa) is missing in the ERA-Interim profiles. Indeed, at these levels, the upper tropospheric moisture is mainly influenced by the level of maximum outflow of saturated air by deep convective clouds (see Folkins *et al.*, 2002; among others), together with drying effects induced by mixing or extratropical air intrusion (e.g. Pierrehumbert and Roca, 1998). Thus, although in ERA-Interim the moist physics, in particular, has been improved (Chuang *et al.*, 2010) the differences on the vertical can bias training and thus affect retrieval of the relative humidity profiles. Based on these results, we decided to use radiosoundings instead of re-analysed profiles for elaboration of the training database. However, this paper is not aimed at studying thoroughly these differences and this is left for a future work in the framework of profile retrieval.

5. Comparative study: three *versus* six channels around 183.31 GHz

5.1. Characteristics of the training dataset

Figure 4 summarizes the characteristics of the sampled database. The total column water vapour (TCWV) evolves with the sea-surface temperature following a well-known

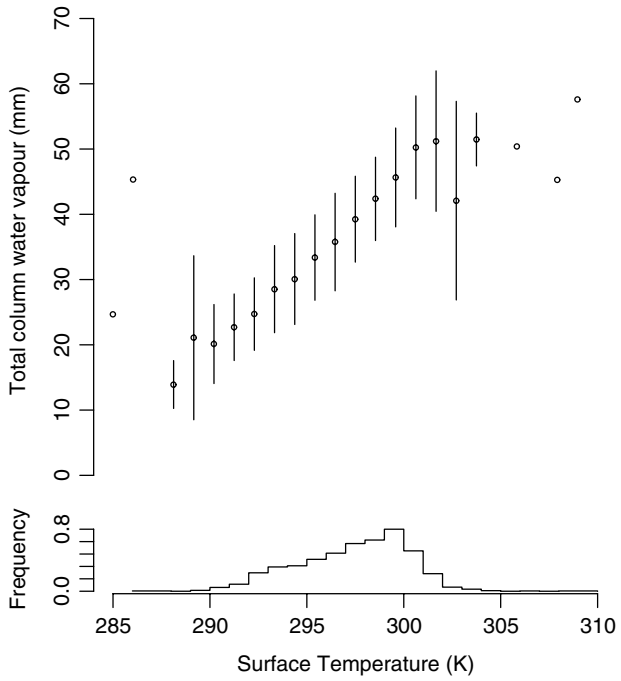

Figure 4. Distribution of total column water vapour (mm) from the radiosounding database according to surface temperature (K). The diagram represents the average plus or minus one standard deviation. The number of observations is represented by the histogram in the lower panel.

pattern of the tropical atmosphere, with an increase in the TCWV as the surface temperature increases (see Stephens, 1990; among others). Synthetic brightness temperatures are associated with each thermodynamic profile with the RTTOV model (see section 3.4). Then, for the present exercise, the 21 levels of relative humidity are averaged into the six layers, delimited in Figure 3: 100–200, 200–325, 325–450, 450–700, 700–900 and 900 hPa–surface. These limits have been defined to frame the structure of the correlation matrix. This database of six-layer relative humidity profiles is finally randomly split in two: a training base (50%) and a validation base (50%).

5.2. *Description of the retrieval model*

To illustrate the interest of the combination of SAPHIR and MADRAS observations, a generalized additive model (GAM; Hastie and Tibshirani, 1987) is developed for estimation of the tropospheric relative humidity at a given level k, namely the response (or explained) variable Y^k, from brightness temperatures only, namely the explanatory variables X_j. Such a model allows one to infer nonlinear relationships between the response variable and the different chosen explanatory variables. This model has the form

$$\mathbb{E}(Y^k|\mathbf{X}) = \sum_{j=1}^{p} f_j(X_j) + \epsilon^k, \qquad (1)$$

and expresses the expectation of Y^k (here the relative humidity of level k) conditionally on a p-dimensional covariate $\mathbf{X}$ $(X_1, \ldots, X_p)$ (here the vector of T_Bs), as a sum of $f_j(X_j)$. $f_j(\cdot)$ are smooth non-parametric functions of the explanatory variables, and ϵ^k is the residual that follows a Gaussian distribution. In the present case, the basis of regression cubic smoothing splines ($s_j(\cdot)$; third-order polynomial functions) are used to represent the nonlinear variations $f_j(\cdot)$ of the T_Bs. Such functions offer a nice trade-off between over- and under-smoothing. In fact, a GAM is a conditional density function. It consists of two components: a parametric conditional normal density function, and a functional relationship between the normal parameters and X.

Figure 5 shows the estimated spline functions $s_j(\cdot)$ for determination of the 325–450 hPa relative humidity layer with the six T_Bs of SAPHIR as explanatory variables, using the training dataset described above. In this figure, the x-axes correspond to each explanatory variable X_j and the y-axes show the contribution of X_j, i.e. $s_j(X_j)$, to explain $Y^{325-450}$. Hence, according to the weighting functions presented in Figure 2, the 325–450 hPa layer is mainly documented by channels ± 2.8 GHz (X_4), ± 1.1 GHz (X_5) and ± 0.2 GHz (X_6), with a major contribution of the ± 1.1 GHz channel (X_5). Figure 5 shows indeed that this channel has the larger dynamic of variation ($[-1; 1]$), with an increase of T_B as the 325–450 hPa relative humidity decreases, while the X_4 and X_6 T_Bs have a smaller contribution to the signal ($[-0.3; 0.8]$ and $[-0.7; 0.5]$ respectively). Finally, the estimated contributions from each of the six T_Bs to the variation of the 325–450 hPa relative humidity are nonlinear, the nonlinearity being higher for the 183.31 ± 4.2 GHz T_B (X_3) and the 183.31 ± 2.8 GHz T_B (X_4) than for the other T_Bs. This highlights the flexibility of a GAM to represent more or less strong nonlinear relationships.

5.3. *Discussion*

A GAM was developed for each of the six layers of relative humidity. In order to quantify the gain induced by the six channels of SAPHIR with respect to the three usual channels of a AMSU-B-like radiometer (one near the centre and two more), a GAM has been applied for these two configurations. The statistical characteristics of each model are summarized in Table 2, with the percentage deviance of the data explained by the model, the mean and the variance of the error for each layer (estimated − observed relative humidity), and the Pearson correlation coefficient in order to quantify the scatter induced by the model. The deviance measures how much fit is lost (in terms of likelihood) by the modelling compared to a perfect (or *saturated*) model: $D = -2[\ln L_c - \ln L_s]$, with L_c the likelihood of the current model and L_s the likelihood of the saturated model.

The addition of three channels clearly improves the restitution of the relative humidity throughout the entire troposphere, with a more pronounced improvement for the upper layer, and an increase of the explained deviance in the adjustment step, from 55% with a AMSU-B radiometer to 87.1% with a SAPHIR radiometer, and a noticeable reduction of the scatter, with the correlation coefficient reaching 0.91 with SAPHIR data (0.76 with AMSU-B). Similar observations are made for the other layers, and more specifically for the lower layers. Figures 6 and 7 support these numbers, with the use of box-and-whiskers diagrams to represent the variability (median, first and third quartiles, and upper and lower limits of the distribution) of the observed and estimated data at each layer. While the mean profile does not show strong differences when the GAM is elaborated with three channels or with six channels (Table 2), the addition of three channels in the 183.31 GHz band

Table 2. Statistics of the different GAM-based estimations of the six-layer relative humidity profile (radiosoundings from the 1996–2006 period) using a SAPHIR-only configuration compared to an AMSU-B-like configuration (values in parentheses). For the model itself the deviance explained (%) is provided. For the validation step: the mean (%) and variance (%^{2}) of the residuals (estimated − observed relative humidity) as well as the Pearson correlation coefficient are given.

Layer	SAPHIR-only (AMSU-B like)			
	Model adjustment:	Reproduction:		
	deviance explained (%)	Mean of residuals (%)	Variance of residuals (%^{2})	Correlation (Pearson)
100–200 hPa	87.1 (55)	0.39 (0.42)	104.8 (253)	0.91 (0.76)
200–325 hPa	91.2 (81.9)	−0.07 (−0.29)	46.9 (61.9)	0.93 (0.91)
325–450 hPa	91.3 (83.1)	0.34 (0.60)	29.9 (53.9)	0.94 (0.89)
450–700 hPa	88.4 (81.4)	1.08 (0.90)	46.4 (51.7)	0.92 (0.91)
700–900 hPa	82.9 (54.3)	0.15 (0.50)	120.9 (233.6)	0.86 (0.72)
900 hPa–surface	75.7 (34.1)	−0.43 (0.18)	120. (275.1)	0.80 (0.49)

Table 3. The same as Table 2 but for the Megha-Tropiques payload, i.e. including the 23.8 GHz and 157 V and H GHz channels of MADRAS.

Layer	SAPHIR and MADRAS			
	Model adjustment:	Reproduction:		
	deviance explained (%)	Mean of residuals (%)	Variance of residuals (%^{2})	Correlation (Pearson)
100–200 hPa	89	0.26	94.1	0.92
200–325 hPa	91.7	0.01	42.9	0.94
325–450 hPa	92.9	0.14	25.1	0.95
450–700 hPa	94.3	0.66	28.1	0.96
700–900 hPa	90.7	−0.52	95.3	0.89
900 hPa–surface	82.8	0.29	85.2	0.89

allows for a better restitution of the distribution of the data. Indeed, from the AMSU-B configuration (Figure 6(a)) to the SAPHIR configuration (Figure 6(b)) the boxes representing the majority of the data (between 25% and 75% percentiles) are closer to each other, as well as the whiskers. This is particularly convincing for the two upper layers and for the lowest layer.

5.4. *Contribution of the MADRAS channels*

Three channels of the MADRAS radiometer are sensitive to the water vapour content of the atmosphere: the 23.8 GHz channel and the two 157 GHz channels (V and H polarizations). While the 23.8 GHz channel is often used to retrieve the integrated water vapour content of the atmosphere, measurements at 157 GHz are useful for the estimation of the water vapour continuum (e.g. English *et al.*, 1994). Thus, although the prime role of this channel is to improve the description of high-level ice clouds associated with the convective systems, it can also be used as an additional channel for observation of the water vapour in the low-level atmosphere. For this purpose, a 150 GHz channel is present on the AMSU-B/NOAA radiometers, whereas there is a 157 GHz channel on the MHS/MetOp radiometers. To test their contributions, a GAM was developed adding these three channels of MADRAS to the six channels of SAPHIR. The statistics of this model are provided in Table 3 and a quick comparison with Table 2 clearly shows the interest in these channels for the restitution of the atmospheric water vapour: the model adjustment is improved for the whole profile, and for each layer the scatter is reduced, with a greater reduction for the lower layers. Again, the box-and-whiskers diagrams of Figures 6(c) and 7(c) support the interest of the combination of SAPHIR and MADRAS for the restitution of relative humidity profiles, since the distributions of the estimated values are much closer to the observed ones when information on the 23.8 GHz channel and of the two 157 GHz channels are added to the model.

6. Conclusion

The purpose of this article is to present the expected gain in the retrieval of the relative humidity profile by the microwave payload of the Megha-Tropiques platform. Synthetic brightness temperatures of the two microwave radiometers are computed using the RTTOV-v9.3 radiative code and a 10-year radiosounding dataset covering the tropical belt. Our results have revealed noticeable differences of patterns in the correlation matrices of the relative humidity profiles between this source of profiles and ERA-Interim profiles. Note that the impact of these differences in the retrieval of ∼200 hPa wide layers of relative humidity has not been evaluated and should be performed in a dedicated work.

The question of all-sky profiles should be also addressed: a combination of thermodynamic profiles measured by radiosoundings and of profiles of cloud liquid and ice-water content derived from CloudSat-CALIPSO observations could be an interesting approach that should help in simulating realistic cloudy T_Bs. Such an approach needs to be fully evaluated against the use of ECMWF profiles, which offer the advantage of consistent all-sky thermodynamics.

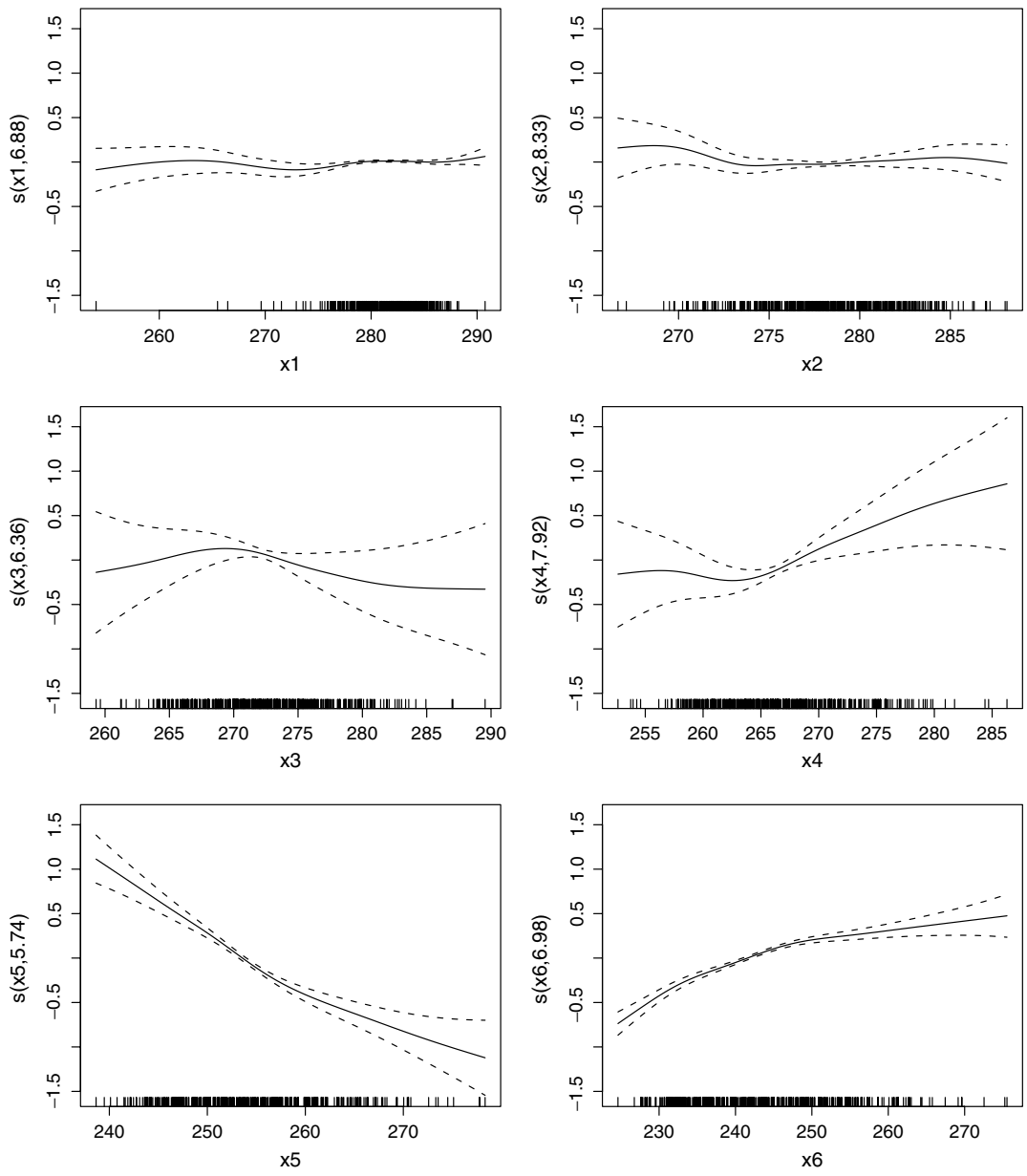

Figure 5. Spline functions s_j estimated for regression of the 325–450 hPa relative humidity layer data, with the six T_Bs of SAPHIR as explanatory variables ($X_1 \ldots X_6 = 183.31 \pm 11$ GHz $\ldots$ 183.31 ± 0.2 GHz, in K). The dashed lines correspond to two standard errors above and below the estimated splines, and the ticks on the x-axis represent the distribution of the observed predictor values of the database. Also indicated in parentheses are the estimated degrees of freedom for each spline.

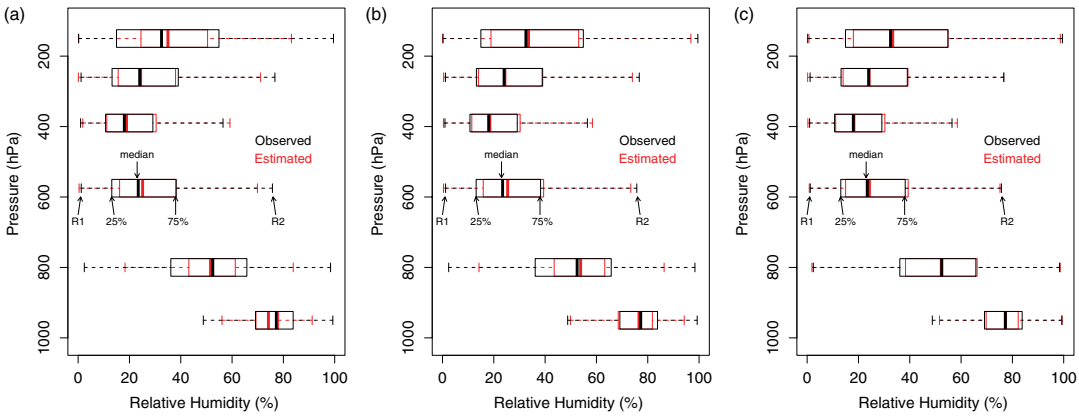

Figure 6. Observed (black) and estimated (red) six-layer relative humidity profiles using a GAM on the 1996–2006 radiosoundings with (a) the three channels of AMSU-B like sounders, (b) the six channels of SAPHIR and (c) the six channels of SAPHIR plus the 23.8 GHz and 157 GHz (V and H) channels of MADRAS as the explanatory variables. For each layer, the box-and-whiskers diagram indicates the median (the central vertical line) and the lower and upper quartiles (left and right edges of the box). R1 and R2 (the whiskers) indicate the lower and upper limits of the distribution within 1.5 times the interquartile range from the lower and upper quartiles, respectively.

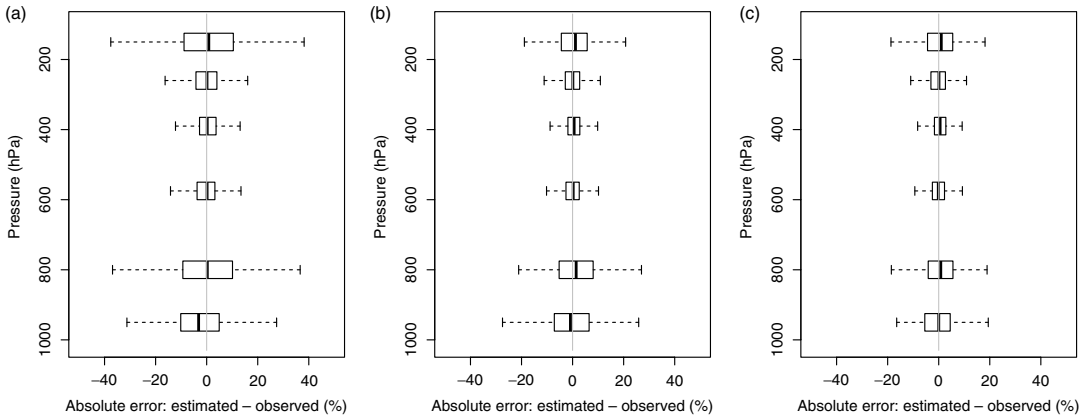

Figure 7. The same as Figure 6 but the box-and-whiskers diagrams are defined upon the error between the estimated and observed relative humidity.

The retrieval algorithm uses a generalized additive model that allows modelling of the strong nonlinearities of the problem with additive cubic smoothing splines. The use of AMSU-B-like information *versus* the use of SAPHIR observations as inputs of the retrieval algorithm has shown a clear reduction of the error of retrieval, the highest improvement concerning the upper and lower layers of the profile, with a variance of the residuals divided by 2 for these layers and correlation coefficients reaching at least a value of 0.8. The addition of MADRAS measurements, at 23.8 GHz and 157 GHz, improves even further the retrieval: throughout the troposphere, the mean residual is less than 1% and the correlation coefficients exceed 0.89. Although the present exercise is restricted to the oceanic and clear-sky situations for simplification purposes, similar improvements should be expected when dealing with cloudy atmosphere and continental surfaces. For those cases, however, it is likely that the other channels of MADRAS (at 18.7 and 36.5 GHz) will bring valuable information to better constrain the retrieval of the lowest layer. Moreover, the use of other explanatory variables from climatologies such as the sea-surface temperature or vertical stratification of the atmosphere (e.g. Obligis *et al.*, 2009), and the correlations between the successive levels, should also provide additional constraints.

Currently, numerous methods of retrieval exist, among them the neural network approach (e.g. Cabrera-Mercadier and Staelin, 1995) or the use of iterative schemes mixing physical constraints and statistics (e.g. Kuo *et al.*, 1994): all of them provide a conditional estimate of the expectation of relative humidity given a set of inputs. A future study will consider the use of the generalized additive model presented here to estimate the whole distribution and not only its expectation. In other words, a GAM can be applied to generate ensembles of probable relative humidity profiles conditioned on given T_Bs (and other explanatory variables). Such a conditional distribution of the relative humidity profile should be more advantageous to users than deterministic products since the distribution will enable access to the whole variability of the atmospheric moisture. It will also be useful to address the issue of the possible non-Gaussian or even non-symmetric distribution of the relative humidity at each layer to infer a more appropriate model. A generalized additive model for location, scale and shape (GAMLSS; Rigby and Stasinopoulos, 2005) may be used for such a purpose.

Acknowledgements

The authors acknowledge the LMD/ABC(t)/ARA group for producing and making available to the community their radiosounding database ARSA (http://ara.abct.lmd.polytechnique.fr/index.php?page=arsa). Thanks also to Rémy Roca and to Nicolas Viltard for fruitful discussions on the radiative transfer in the microwave spectrum. The first author thanks Mathieu Vrac for his advice on the use of generalized additive models and for proofreading the manuscript. Finally, the authors thank the two anonymous reviewers for their comments, which greatly enhanced the quality of the manuscript.

References

Aires F, Bernardo F, Prigent C. 2012. Atmospheric water vapor profiling from passive microwave sounders over ocean and land. Part 1: Methodology for the Megha-Tropiques mission. *Q. J. R. Meteorol. Soc.* (submitted, this issue).

Allison L, Steranka J, Cherrix G, Hilsenrath E. 1972. Meteorological applications of the Nimbus 4 Temperature–Humidity Infrared Radiometer, 6.7 microns channel data. *Bull. Am. Meteorol. Soc.* **53**: 526–535.

Bates J, Jackson D. 2001. Trends in upper-tropospheric humidity. *Geophys. Res. Lett.* **28**: 1695–1698.

Bennartz R, Bauer P. 2003. Sensitivity of microwave radiances at 85–183 GHz to precipitating ice particles. *Radio Sci.* **38**: 8075, DOI: 10.1029/2002RS002626.

Berrisford P, Dee D, Fielding K, Fuentes M, Kallberg P, Kobayashi S, Uppala S. 2009. *The ERA-Interim archive. ERA Report series*, No. 1. ECMWF: Reading, UK. Available: www.ecmwf.int/publications.

Blankenship C, Al-Khalaf A, Wilheit T. 2000. Retrieval of water vapor profiles using SSM/T-2 and SSM/I data. *J. Atmos. Sci.* **57**: 939–955.

Brogniez H, Pierrehumbert R. 2006. Using microwave observations to assess large-scale control of free tropospheric water vapor in the mid-latitudes. *Geophys. Res. Lett.* **33**: L14801, DOI: 10.1029/2006GL026240.

Brogniez H, Roca R, Picon L. 2006. A clear-sky radiance archive from METEOSAT 'water vapor' observations. *J. Geophys. Res.* **111**: D21109, DOI: 10.1029/2006JD007238.

Brogniez H, Roca R, Picon L. 2009. A study of the free tropospheric humidity interannual variability using Meteosat data and an advection–condensation transport model. *J. Climate* **22**(404): 6773–6787.

Buehler S, Kuvatov M, Sreerekha T, John V, Rydberg B, Eriksson P. 2007. A cloud filtering method for microwave upper tropospheric humidity measurements. *Atmos. Chem. Phys.* **7**: 5531–5542.

Buehler S, Kuvatov M, John V, Milz M, Soden B, Jackson D, Notholt J. 2008. An upper tropospheric humidity data set from operational satellite microwave data. *J. Geophys. Res.* **113**: D14110, DOI: 10.1029/2007JD009314.

Burns B, Wu X, Diak G. 1997. Effects of precipitation and cloud ice on brightness temperatures in MASU moisture channels. *IEEE Trans. Geosci. Remote Sens.* **35**: 1429–1437.

Cabrera-Mercadier C, Staelin D. 1995. Passive microwave relative humidity retrievals using feedforward neural networks. *IEEE Trans. Geosci. Remote Sens.* **33**: 1324–1328.
Chiodo G, Haimberger L. 2010. Interannual changes in mass consistent energy budgets from ERA-Interim and satellite data. *J. Geophys. Res.* **115**: D02112, DOI: 10.1029/2009JD012049.
Chuang H, Huang X, Minschwaner K. 2010. Interannual variations of tropical upper tropospheric humidity and tropical rainy-region SST: comparisons between models, reanalyses, and observations. *J. Geophys. Res.* **115**: D21125, DOI: 10.1029/2010JD014205.
Deblonde G, English S. 2001. Evaluation of the FASTEM-2 fast microwave oceanic surface emissivity model. In *Technical Proceedings ITSC-XI*, Budapest; 20–26 September 2000; 67–78.
Dee D, Uppala S. 2009. Variational bias correction of satellite radiance data in the ERA423 Interim reanalysis. *Q. J. R. Meteorol. Soc.* **135**: 1830–1841.
Delanoe J, Hogan R, Forbes R, Bodas-Salcedo A, Stein T. 2011. Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and SALIPSO data. *Q. J. R. Meteorol. Soc.* **137**: 2064–2078.
English S, Guillou C, Prigent C, Jones D. 1994. Aircraft measurements of water vapour continuum absorption at millimetre wavelengths. *Q. J. R. Meteorol. Soc.* **120**: 603–625.
Eymard L, Bernard R, Lojou J-Y. 1993. Validation of microwave radiometer geophysical parameters, using meteorological model analyses. *Int. J. Remote Sens.* **14**: 1945–1963.
Eymard L, Gheudin M, Laborie P, Sirou F, Le Gac C, Vinson J-P, Franquet S, Desbois M, Karbou F, Roca R, Scott N and Waldteufel P. 2002. The SAPHIR humidity sounder. *Notes des activités instrumentales de l'IPSL* **24**:.
Eymard L, Karbou F, Janicot S, Chouaib N, Pinsard F. 2010. On the use of Advanced Microwave Sounding Unit-A and -B measurements for studying the monsoon variability over West Africa. *J. Geophys. Res.* **115**: D20115, DOI: 10.1029/2009JD012935.
Folkins I, Kelly K, Weinstock E. 2002. A simple explanation for the increase in relative humidity between 11 and 14 km in the tropics. *J. Geophys. Res.* **107**: 4736, DOI: 10.1029/2002JD002185.
Greenwald T, Christopher S. 2002. Effect of cold clouds on satellite measurements near 183 GHz. *J. Geophys. Res.* **107**: 4170, DOI: 10.1029/439 2000JD000258.
Hastie T, Tibshirani R. 1987. Generalized additive models: some applications. *J. Am. Statist. Assoc.* **82**: 371–386.
Hong G, Heygster G, Miao J, Kunzi K. 2005. Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements. *J. Geophys. Res.*, **110**: D05205, DOI: 10.1029/2004JD004949.
Houshangpour A, John V, Buehler S. 2005. Retrieval of upper tropospheric water vapor and upper tropospheric humidity from AMSU radiances. *Atmos. Chem. Phys.* **5**: 2019–2028.
Huang X, Soden B, Jackson D. 2005. Interannual co-variability of tropical temperature and humidity: a comparison of model, reanalysis data and satellite observation. *Geophys. Res. Lett.* **32**: L17808, DOI: 10.1029/2005GL023375.
Kakar R. 1983. Retrieval of clear sky moisture profiles using the 183 GHz water vapor line. *J. Climate Appl. Meteorol.* **22**: 1282–1289.
Kakar R, Lambrigsten B. 1984. A statistical correlation method for the retrieval of atmospheric moisture profiles by microwave radiometry. *J. Climate Appl. Meteorol.* **23**: 1110–1114.
Karbou F, Aires F, Prigent C, Eymard L. 2005. Potential of Advanced Microwaves Sounding Unit-A (AMSU-A) and AMSU-B measurements for atmospheric temperature and humidity profiling over land. *J. Geophys. Res.* **110**: D07109, DOI: 10.1029/2004JD005318.
Karbou F, Rabier F, Lafore J-P, Redelsperger J-L, Bock O. 2010. Global 4D-Var assimilation and forecast experiments using AMSU observations over land. Part II: Impact of assimilating surface sensitive channels of the African Monsoon during AMMA. *Weather Forecast.* **460**: 20–36.
Kuo C, Staelin D, Rosenkranz P. 1994. Statistical iterative scheme for estimating atmospheric relative humidity profiles. *IEEE Trans. Geosci. Remote Sens.* **32**: 254–260.
Liou KN. 2002. *An introduction to Atmospheric Radiation*, 2nd edition, Academic Press: San Diego, CA.
Liu G, Curry J. 1996. Large-scale cloud features during January 1993 in the North Atlantic Ocean as determined from SSM/I and SSM/T2 observations. *J. Geophys. Res.* **101**: 7019–7032.
Liu Q, Weng F. 2005. One-dimensional variational retrieval algorithm of temperature, water vapor and cloud water profiles from Advanced Microwave Sounding Unit (AMSU). *IEEE Trans. Geosci. Remote Sens.* **43**: 1087–1095.
Matricardi M, Chevallier F, Kelly G, Thépaut JN. 2004. An improved general fast radiative transfer model for the assimilation of radiance observations. *Q. J. R. Meteorol. Soc.* **130**: 153–173.
McCarthy M, Toumi R. 2004. Observed interannual variability of tropical troposphere relative humidity. *J. Climate* **17**: 3181–3191.
Obligis E, Rahmani A, Eymard L, Labroue S, Bronner E. 2009. An improved retrieval algorithm for water vapor retrieval: application to the Envisat microwave radiometer. *IEEE Trans. Geosci. Remote Sens.* **47**: 3057–3064.
Peixoto J, Oort A. 1996. The climatology of relative humidity in the atmosphere. *J. Climate* **9**: 3443–3463.
Pierrehumbert R, Roca R. 1998. Evidence for control of Atlantic subtropical humidity by large-scale advection. *Geophys. Res. Lett.* **25**: 4537–4540.
Rieder M, Kirchengast G. 1999. Physical-statistical retrieval of water vapor profiles using SSM/T-2 sounder data. *Geophys. Res. Lett.* **26**: 1397–1400.
Rigby R, Stasinopoulos D. 2005. Generalized additive models for location, scale and shape. *Appl. Statist.* **54**: 507–554.
Roca R, Lafore J-P, Piriou C, Redelsperger J-L. 2005. Extratropical dry-air intrusions into the West African Monsoon midtroposphere: an important factor for the convective activity over the Sahel. *J. Atmos. Sci.* **62**: 390–407.
Roca *et al.* 2012. Overview of the Megha-Tropiques mission. *Q. J. R. Meteorol. Soc.* (submitted, this issue).
Roca R, Bergès J-C, Brogniez H, Capderou M, Chambon P, Chomette O, Cloché S, Fiolleau T, Jobard I, Lémond J, Ly M, Picon L, Raberanto P, Szantai A, Viollier M. 2010. On the water and energy cycles in the Tropics. *C.R. Geosci.* **342**: 390–402.
Rosenkranz P. 2001. Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements. *IEEE Trans. Geosci. Remote Sens.* **39**: 2429–2435.
Rosenkranz P, Komichak M, Staelin D. 1982. A method for estimation of atmospheric water vapor profiles by microwave radiometry. *J. Appl. Meteorol.* **21**: 1364–1370.
Saunders R, Hewison T, Stringer S, Atkinson N. 1995. The radiometric characterization of AMSU-B. *IEEE Trans. Microwave Theory* **43**: 760–771.
Saunders R, Matricardi M, Brunel P. 1999. An improved fast radiative transfer model for assimilation of satellite radiance observations. *Q. J. R. Meteorol. Soc.* **125**: 1407–1425.
Schaerer G, Wilheit T. 1979. A passive microwave technique for profiling of atmospheric water vapor. *Radio Sci.* **14**: 371–375.
Shi L. 2001. Retrieval of atmospheric temperature profiles from AMSU-A measurement using a neural network approach. *J. Atmos. Oceanic Technol.* **18**: 340–347.
Sohn B-J, Chung E-S, Schmetz J, Smith E. 2003. Estimating upper tropospheric water vapor from SSM/T-2 satellite measurements. *J. Appl. Meteorol.* **42**: 488–505.
Staelin D, Kunzi K, Pettyjohn R, Poon R, Wilcox R. 1976. Remote sensing of atmospheric water vapor and liquid water with the Nimbus-5 microwave spectrometer. *J. Appl. Meteorol.* **15**: 1204–1214.
Stephens G. 1990. On the relationship between water vapor over the oceans and sea surface temperature. *J. Climate* **3**: 634–645.
Wang J, Chang L. 1990. Retrieval of water vapor profiles from microwave radiometric measurements near 90 and 183 GHz. *J. Appl. Meteorol.* **29**: 1005–1013.
Wang J, King JL, Wilheit TT, Szejwach G, Gesell LH, Nieman RA, Niver DS, Krupp BM, Gagliano JA. 1983. Profiling atmospheric water vapor by microwave radiometry. *J. Climate Appl. Meteorol.* **22**: 779–788.
Wilheit T, Al-Khalaf A. 1994. A simplified interpretation of the radiances from the SSM/T-2. *Meteorol. Atmos. Phys.* **54**: 203–212.
Zhang L, Qiu C, Huang J. 2008. A three-dimensional satellite retrieval method for atmospheric temperature and moisture profiles. *Adv. Atmos. Sci.* **25**: 897–904.

