Near-Real-Time Estimation of Water Vapor Column From MSG-SEVIRI Thermal Infrared Bands: Implications for Land Surface Temperature Retrieval

Yves Julien, José A. Sobrino, Cristian Mattar, and Juan C. Jiménez-Muñoz

Abstract—The Meteosat Second Generation-Spinning Enhanced Visible and Infrared Imager (MSG-SEVIRI) instrument provides observations of half the globe every 15 min, at low spatial resolution. These data are an invaluable tool to observe daily to yearly cycle of land surface temperature (LST), as well as for various early warning systems. However, advanced algorithms for LST estimation require a previous estimation of the water vapor (WV) column above the observed pixel, for which no instantaneous retrieval methods are yet available, and therefore hinders their implementation in a near-real-time processing chain for MSG-SEVIRI data. This work analyzes three different formulations for such WV retrieval, which are compared to independent WV estimates obtained from radiosoundings. The best suited algorithm is then selected for WV estimation and compared with the results obtained with a previous noninstantaneous algorithm [23]. This comparison shows that, in spite of retrieval errors higher than the ones reported in the literature, the estimated WV compares relatively well with in situ data, while allowing for an instantaneous estimation of WV column (every 15 min). Error propagation analysis and direct comparison show that the observed increase in WV estimation error has a negligible influence on LST retrieval. Therefore, this algorithm is well suited to be implemented in a near-real-time processing chain for MSG-SEVIRI data.

Index Terms—Land surface temperature (LST), Meteosat Second Generation-Spinning Enhanced Visible and Infrared Imager (MSG-SEVIRI), water vapor (WV).

I. INTRODUCTION

THE Meteosat Second Generation (MSG) satellite platform, with its geostationary orbit, provides valuable information to the scientific community. The Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument, onboard MSG, acquires data from visible to thermal infrared (TIR) wavelengths at coarse resolution (3 km at nadir) every 15 min. Therefore, numerous applications have been carried out, such as vegetation monitoring [7], [21], dead fuel moisture estimation [15], fire monitoring [2], [9], air temperature estimation [16], land surface temperature (LST) estimation [13], [22], [24], cloud detection [6] or rainfall estimation [10]. Many of these applications are specifically dedicated to or dependent on LST estimation, which can be retrieved through split-window (SW) methods [11], among others. LST estimation depends, in turn, on the estimation of surface emissivity in TIR bands [12], and retrieval of total amount of water vapor (WV) increases the accuracy of the approach [20].

Within the framework of the CEOS-SPAIN project, the Global Change Unit of the University of Valencia is providing (among other tasks) near-real-time estimates of LST from MSG-SEVIRI data, to be uploaded onto a server from which these data will be available to the general public. To that end, near-real-time estimation of LST is required, and since advanced algorithms for LST estimation require the previous estimation of total WV column (among other parameters; for more details, see [22]), a near-real-time estimation of WV has to be carried out. Existing algorithms [19], [23] allow the estimation of WV column, although both algorithms exploit brightness temperature differences between two acquisitions to retrieve WV amount. This brightness temperature difference is set to at least 5 K for Schroedter-Homscheidt et al. [19] algorithm, while Sobrino and Romaguera [23] algorithm requires it to be above 10 K. As a result, near-real-time implementation of these algorithms is not straightforward since, depending on the time lapse chosen between acquisitions, various brightness temperature differences can be retrieved (provided they are above each method threshold), leading to different estimates of WV, assumed to be constant over the chosen time lapse. Up to 2013, MSG-SEVIRI WV was estimated at the Global Change Unit of the University of Valencia as an average of all valid estimates of WV column obtained by the application of Sobrino and Romaguera [23] algorithm for all possible 6-h-long time differences within a given day. However, this approach prevents a near-real-time estimation of LST since a whole day of data needs to be processed to retrieve the mandatory WV. Moreover, this approach assumes constant WV amount during one day, which can lead to artificial discontinuities between consecutive days. Therefore, a straightforward algorithm would be beneficial for all MSG-SEVIRI receiving stations.
The WV estimation method presented hereafter is designed to retrieve WV instantaneously, without any ancillary data (i.e., all needed information are retrieved from SEVIRI data only). Since SEVIRI sensor does not include the adequate bands in the visible wavelengths to estimate WV, this estimation can only be carried out from the available TIR SEVIRI bands. The need for instantaneous WV estimates also prevents the use of methods based on the WV daily cycle.

II. DATA

Three different data sets have been used in this study: 1) real-time MSG-SEVIRI data received at the Global Change Unit of the University of Valencia; 2) atmospheric profile data sets for simulation purposes; and 3) radiosoundings launched at different permanent stations extracted from the University of Wyoming data set and used for validation purposes. These three data sets are described briefly hereafter.

A. MSG-SEVIRI Data

The MSG-SEVIRI data used in this study have been acquired using a direct-broadcast High Resolution Picture Transmission (HRPT) system implemented at the Global Change Unit of the University of Valencia since mid-2007. This system consists of a parabolic dish, a PC with the hardware and software to decode L-band data, and a set of storage devices to save all the received SEVIRI data. Since SEVIRI temporal resolution is 15 min, the received data amount to 96 images per day for each channel. Therefore, the storage space needed is 1 TB of data per month. SEVIRI data. Since SEVIRI sensor does not include the adequate bands in the visible wavelengths to estimate WV, this estimation can only be carried out from the available TIR SEVIRI bands. The need for instantaneous WV estimates also prevents the use of methods based on the WV daily cycle.

B. Atmospheric Profile Data Set

The atmospheric profiles used in the simulation procedure carried out in the framework of this study were extracted from the Thermodynamic Initial Guess Retrieval (TIGR) database [1]. In particular, a selection of 61 atmospheres from the TIGR database was performed, as proposed in Sobrino et al. [20]. This data set, referred hereafter as TIGR-61, includes global atmospheric conditions with 28 atmospheres assigned to the tropical model, 12 to the midlatitude summer model, 12 to the subarctic winter model, and 9 to the U.S. standard model). The mean WV value for all the atmospheric profiles included in TIGR-61 is (2.9 ± 1.7) g · cm⁻², with minimum and maximum values of 0.2 and 6 g · cm⁻², respectively.

C. Radiosonde Database

The radiosonde database used in this work consists of several atmospheric profiles freely available for download at the University of Wyoming server (http://weather.uwyo.edu/upperair/sounding.html). These vertical profiles were retrieved from different radiosonde types and models launched in several atmospheric conditions, such as temperate, tropical, desert, and subtropical, in order to span the whole amplitude of WV to be used in the LST algorithm retrievals. This database has been used in previous works for remote sensing applications and precipitable water long-term analysis [14], [18].

We selected all radiosonde stations located in South America, Europe, Africa, and Middle East, in order to cover the highest possible proportion of MSG-SEVIRI disk. The data acquired at these stations for year 2010 (first week of January, April, July, and October as for SEVIRI data) were used to retrieve alternative total WV column for validation purposes, as well as to build a rudimentary cloud mask associated to the WV estimates. Validation was carried out by comparing available radiosonde data to the closest (in space and time) SEVIRI estimation. This implies that we assume constant WV over SEVIRI pixel size for a 15-min duration.

III. METHODS

A. Simulated Data

WV algorithms were developed from analysis of a complete simulated data set constructed under different surface and atmospheric conditions. At-sensor brightness temperatures for the SEVIRI TIR bands were reproduced from forward simulations based on the radiative transfer code MODTRAN-4 [5]. The surface was characterized from surface emissivities of 108 natural samples (soils, rocks, vegetation, water, snow/ice) extracted from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) spectral library [4]. The atmospheric conditions were characterized from the atmospheric profiles included in the TIGR-61 data set described earlier.
B. WV Algorithms and Coefficients

In order to estimate instantaneously the total amount of WV from MSG-SEVIRI data, different mathematical structures for the combination of the SEVIRI TIR bands were considered. After a preliminary analysis of results obtained with these combinations using the simulated data, a final selection of three algorithms was performed as follows:

\[ \text{WV1} = a_0 + a_1 \cdot \text{WV062} + a_2 \cdot \text{WV073} + a_3 \cdot \text{IR087} + a_4 \cdot \text{IR097} + a_5 \cdot \text{IR108} + a_6 \cdot \text{IR120} + a_7 \cdot \text{IR134} \]  
\[ \text{WV2} = a_0 + a_1 \cdot \text{WV062} \cdot (\text{IR108} - \text{IR120}) \]  
\[ \text{WV3} = a_0 + a_1 \cdot (\text{IR108} - \text{IR120}) \]  

(WV algorithms presented in (1)–(3) only use at-sensor brightness temperatures as input data. The WV algorithm given by (1) is a linear combination of all SEVIRI TIR bands. In this way, we take advantage of all the TIR spectral information, including also water absorption bands (WV062 and WV073). However, WV1 algorithm is expected to be more sensitive to accuracy of TIR measurements [e.g., Noise Equivalent Differential Temperature (NE\text{Δ}T)], since the algorithm uses absolute values of at-sensor brightness temperatures. On the contrary, the algorithm given by (3) uses information extracted from only two TIR bands in the spectral region of 10–12 \text{μm} (known as the SW technique). Since this algorithm does not use absolute values of brightness temperatures but the difference between two bands, it is expected to be less sensitive to errors in TIR measurements. The algorithm given by (2) is based on the SW algorithm (3) but also includes information from one WV absorption band (WV062).

Table I presents the coefficients and estimation errors for (1)–(3) using the simulated data set. Coefficients were obtained from linear least squares minimization, whereas the total error (\(e_{\text{WV}}\)) were obtained from total sum of squares of the different errors according to the classical rules for error propagation, i.e.,

\[ e_{\text{WV}} = \sqrt{\sigma^2 + \sum_{i=1}^{N} \left( \frac{\partial W}{\partial T_i} \cdot e(T_i) \right)^2} \]  

where \(\sigma\) is the standard error of estimation, \(T_i\) are the at-sensor brightness temperatures for the different SEVIRI bands (\(i = 1, N\)), and \(e(T_i)\) is the NE\text{Δ}T, assumed to be 0.2, 0.1, 0.1, 0.3, 0.1, 0.15, and 0.4 K for SEVIRI bands WV6.2, WV7.3, IR8.7, IR9.7, IR10.8, IR12.0, and IR13.4, respectively (corresponding to the European Organization for the Exploitation of Meteorological Satellites’ (EUMETSAT) technical specifications).

Results show that WV1 has the highest correlation coefficient (\(R^2\)) and the lowest error (0.6 g · cm\(^{-2}\)), whereas WV2 and WV3 provide similar results, with \(R^2\) values of 0.78 and errors of 0.9 g · cm\(^{-2}\) in both cases.

C. Estimation of Radiosonde WV

WV estimates were derived from cloud-free atmospheric soundings following the method described by Ross and Elliot (2001). To this end, for each atmospheric profile, a simple cloud filter was implemented. This filter is based on high content of relative humidity (RH) to reveal cloud presence [25]. For this work, we discarded radiosonde, which contains any atmospheric level with RH values equal to or higher than 80%. Then, the saturated vapor pressure (\(e_s\)) was estimated by using

\[ e_s = 611 \cdot 10^{(17.27T/(237.7 + T))} \]  

where \(T\) is the air temperature (in degrees celsius). This parameter is used to estimate the specific humidity

\[ q_v = 0.622 \cdot \frac{e_s \cdot RH}{100 \cdot p} \]  

where \(q_v\) is the specific humidity, \(e_s\) is the saturated vapor pressure estimated in (5), \(p\) and \(RH\) are atmospheric pressure and RH provided for each atmospheric pressure level, respectively. Once \(q_v\) is estimated, WV values were retrieved by integration of the whole atmospheric vertical profile, as follows:

\[ \text{WV} = 0.01 \int_{p}^{p_0} a_v dp \approx 0.01 \sum q_v \Delta p \]  

where \(\Delta p\) corresponds to a finite difference for atmospheric pressure, 0 and \(z\) are the limits for integration along the whole atmospheric profile from 1000 hPa to the last atmospheric pressure level captured by the radiosonde (around 10 hPa or lower).

IV. RESULTS AND DISCUSSION

We estimated WV from MSG-SEVIRI data by using all three algorithms (1)–(3) and over all selected radiosonde stations, which we compared to clear-sky (see Section III) radiosonde WV estimates, considered as ground truth for this study. When considering all stations, the RMSE obtained for each algorithm are 3.24 g · cm\(^{-2}\) for WV1, 1.12 g · cm\(^{-2}\) for WV2, and 1.25 g · cm\(^{-2}\) for WV3. These errors are quite high when compared to other existing algorithms (6.8 mm, 0.68 g · cm\(^{-2}\), for [19]; 0.5 g · cm\(^{-2}\) for [23]), particularly for WV1. We also carried out simulations including an observation angle dependence of (1)–(3), although the improvement in retrieved RMSE was not substantial (not shown). When looking at the spatial distribution of the errors presented earlier as well as
their dependence on both SEVIRI viewing angle and mean radiosonde WV for each of the WV algorithms (see Fig. 2), one can observe that the errors for WV1 have a clear angular dependence, with WV1 estimates unavailable for the outer portion of MSG-SEVIRI disk, due to the high amount of data needed by the algorithm, and high errors (above \(3 \text{ g} \cdot \text{cm}^{-2}\)) decreasing to lower errors when getting closer to nadir. In that case, including an angular dependence of WV1 coefficients decreased the RMSE of the algorithm (not shown), although not below the RMSE obtained for the other two algorithms.
The WV2 algorithm shows higher errors close to the Equator, particularly when both mean radiosonde WV and SEVIRI viewing angle are high, corresponding to areas where the cloud cover is more frequent and has probably been underestimated by the simple cloud mask we used in this study. For example, most values in Amazonia correspond to the highest observed errors for this algorithm. When we exclude these stations from the RMSE calculation, it drops to 1.04 g · cm⁻², still twice the lowest errors obtained by other existing algorithms [19], [23]. Finally, the WV3 algorithm shows slightly higher RMSE than the WV2 algorithm, mainly for stations close to SEVIRI nadir. Therefore, we selected the WV2 algorithm to estimate WV instantaneously from MSG-SEVIRI, and hereafter, we analyze more thoroughly the validity of the retrieved WV for this algorithm only.

To have a closer look at the time series of retrieved WV, and in order to compare with existing algorithms, we estimated WV from WV2, on one hand, and from the method developed by Sobrino and Romaguera [23], on the other hand, for the whole MSG-SEVIRI disk for the first week of January, April, July, and October, in 2010. Sobrino and Romaguera [23] WV estimates (WVSR) were summarized as daily averages, by estimating
WV for any possible 6-h time lapse within a given day, considering as valid only the pixels for which the temperature during these 6-h time lapses had experienced a change above 10 K. This allows a reduction of WVSR daily dispersion (not shown) associated to the uncertainty regarding the “effective hour” to which each 6-h time lapse should be affected. Fig. 3 presents the resulting WV time series at four selected radiosonde stations, using SEVIRI WV estimates closest in space to their geographical location. These stations were selected as belonging to different countries and/or atmospheric conditions among the 53 stations, which had an average of at least one clear-sky observation every two days. Their names and geographical coordinates are indicated in Table II. Compared to radiosonde data, WV2 estimates are generally in agreement, although some slight discrepancies can be observed for low (station 08430, DOY 94–95) and high values (station 08430, DOY 186–188; station 26063, DOY 182; station 83612, DOY 4–7). On the other hand, WVSR estimates tend to diverge from radiosonde data, except for station 08430 where the overall agreement is comparable to the ones reached with WV2. This finding is surprising since WVSR estimates are expected to have a better accuracy than WV2, although more investigation is needed to draw definitive conclusions. One can observe higher values for radiosonde stations, which may correspond to undetected cloud presence, as could be inferred from the vertical dispersion for simultaneous observations in SEVIRI WV2. This is clearly the case for the first four observations of the bottom graph in Fig. 3 (Campo Grande, Brazil). Finally, for all four stations, the WV2 algorithm shows a clear diurnal pattern, with seasonal variations in amplitude, although WV2 daily minimum tends to remain stable through the year.

However, these discrepancies do not lead to high differences in resulting LST. Fig. 4 shows the root-mean-square difference (RMSD) between retrieved LST using Atitar and Sobrino [4] algorithm with WV column estimated with the methods of Romaguera and Sobrino [22] versus (2).

Validation of the WV2 algorithm shows higher RMSE errors (1 g·cm⁻²) than the ones reported for other algorithms in the literature [19], [23]. However, Fig. 4 and error propagation analysis show that the resulting error increase in LST estimation is negligible, to be compared with an increased simplicity in WV estimation. As stated earlier, the two other available methods need a difference of 5 to 10 K between two images to retrieve accurate WV amount. However, in a near-real-time processing chain, these methods are tricky to implement. For instance, one can use a previous image acquired X hours earlier (X = 6 in the implementation we used earlier for the Sobrino and Romaguera [23] algorithm). However, nothing guarantees that the temperature difference between these two acquisitions will be enough for the applicability of the chosen method, since this temperature difference depends directly on the temperature daily cycle of the considered pixel location. Not to mention the fact that the implicit assumption of these methods is that WV remains constant during these X hours, in contradiction to time series of WV2 estimates in Fig. 3. The simplicity of the proposed approach allows a straightforward estimation of WV, which makes this approach well suited for the instantaneous retrieval of MSG-SEVIRI LST. On the other hand, we do not recommend this method for strict WV estimation, since already existing methods [19], [23] have been shown to provide WV estimates with better accuracy, although their formulation prevents their use in any near-real-time MSG-SEVIRI processing, in opposition with the method (WV2) presented here.

V. CONCLUSION

This paper has investigated three algorithms to retrieve instantaneously the total amount of WV from MSG-SEVIRI data. Among these three algorithms, one (WV2), based on only three MSG-SEVIRI bands, has been shown to produce smaller errors, although these errors are twice the errors reported in previous works. When compared with the algorithm previously used

### Table II

**Characteristics of Radiosonde Stations Used for Method Validation and Comparison (See Fig. 3)**

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>08430</td>
<td>Murcia, Spain</td>
<td>38.00, -1.16</td>
</tr>
<tr>
<td>26063</td>
<td>Voyeykovo, Russia</td>
<td>59.95, 30.70</td>
</tr>
<tr>
<td>62378</td>
<td>Helwan, Egypt</td>
<td>29.860N; 31.330E</td>
</tr>
<tr>
<td>83612</td>
<td>Campo Grande, Brazil</td>
<td>-20.46, -54.66</td>
</tr>
</tbody>
</table>

Fig. 4. RMSD between LSTs estimated for the first weeks of January, April, July, and October, in 2010, by using Atitar and Sobrino [4] LST algorithm with WV column estimated with the methods of Romaguera and Sobrino [22] versus (2).
by the authors to estimate WV, as well as independent data from radiosoundings, the novel algorithm behaves correctly. Moreover, error propagation analysis and direct comparison show that the increased error of this novel WV algorithm can be neglected when estimating LST. Therefore, this straightforward algorithm (WV2) has been implemented in the near-real-time MSG-SEVIRI processing chain of the receiving station located at the Global Change Unit of the University of Valencia, Spain, and could be implemented in other MSG-SEVIRI receiving stations as well.

ACKNOWLEDGMENT

The authors would like to thank the members of the Global Change Unit of the University of Valencia for their dedication in keeping the MSG receiving-station running over the years, and the University of Wyoming for providing free access to their radiosoundings database.

REFERENCES


Yves Julien

received the Ph.D. degree in earth physics and thermodynamics from the University of Valencia, Valencia, Spain, in 2008 and the Ph.D. degree in electronics, electrotechnics, and automatics (specialized in remote sensing) from the University of Strasbourg, Strasbourg, France.

He is a Researcher with the Global Change Unit, University of Valencia. He is the author of more than 30 international papers. His research interests include temperature and vegetation index interactions as well as time series analysis for land cover dynamic monitoring.

José A. Sobrino

is a Professor of physics and remote sensing, the President of the Spanish Association of Remote Sensing, and the Head of the Global Change Unit at the University of Valencia, Spain. He is the author of more than 150 papers and the Coordinator of the European projects WATERMED and EAGLE. His research interests include atmospheric correction in visible and infrared domains, the retrieval of emissivity and surface temperature from satellite images, and the development of remote sensing methods for land cover dynamic monitoring.

Dr. Sobrino has been a member of the Earth Science Advisory Committee of the European Space Agency since November of 2003. He is the Chairperson of the series of International Symposiums on Recent Advances in Quantitative Remote Sensing.

Cristian Mattar

received the Ph.D. degree in physics from the University of Valencia, Valencia, Spain, in 2011.

He is currently a Research Scientist with the Laboratory for Analysis of the Biosphere, Department of Environmental Sciences, University of Chile, Santiago, Chile. His main research interests include calibration and validation of remote sensing products, soil moisture estimates, and long-term trend analysis.

Juan C. Jiménez-Muñoz

received the Ph.D. degree in physics from the University of Valencia, Valencia, Spain, in 2005.

He is currently a Research Scientist in the Global Change Unit of the Image Processing Laboratory at the University of Valencia, where he has been an Assistant Professor in the Department of Earth Physics and Thermodynamics, Faculty of Physics, since 2010. His main research interests include applications of thermal remote sensing and land surface temperature and emissivity retrieval.