IMPROVEMENTS TO THE NEURAL NETWORK RETRIEVAL OF LAYER PRECIPITABLE WATER INCLUDING AN IR SEVIRI LOCAL RADIANCE-BIAS CORRECTION

Miguel A. Martinez, Mercedes Velazquez, Gabriela Cuevas, Juan Ruiz
Instituto Nacional de Meteorología, Spain

Abstract
A nonlinear method for the retrieval of LPW (Layer Precipitable Water) has been developed from infrared (IR) MSG SEVIRI Imagery in the framework of the EUMETSAT SAF in support to Nowadays and Very Short Range Forecasting (SAFNWC). LPW provides information on the water vapour contained in a vertical column of unit cross-section area in four layers (low, middle, high and total) in the troposphere. The implemented LPW algorithm consists of three stages. First, the pre-processing stage includes the IR radiances cloud clearing using SAFNWC Cloud Mask and a local infrared radiance-bias correction. Second, a feed-forward neural network (NN) is used to estimate each desired parameter. Third, a local parameter-bias correction is applied. The NN method requires significantly less computation than traditional variational retrieval methods while achieving acceptable performance. Thus, this method is particularly suitable for SAFNWC LPW which runs on a PC in near real time together with other eleven products (at pixel by pixel scale every 15 minutes over MSG-N).

This paper has two major components:
First, the local bias correction, which has been recently included in the LPW algorithm, is discussed. The algorithm is both trained and tested using simulated radiances. These radiances are obtained using the 60L-SD database sampled from ECMWF as input to RTTOV-7 radiative transfer model. Since the IR SEVIRI radiance measurements are used as inputs to the LPW algorithm, a radiance-bias correction is required, as it is done in the direct assimilation of satellite radiance measurements in numerical weather-prediction (NWP) systems. The scheme for radiance-bias correction at the NWP model applies a latitudinal dependent scan correction and air-mass predictors computed from the background field. The use of model predictors is a good philosophy in the direct assimilation of satellite radiances in NWP systems. However, the LPW algorithm tries to be as independent from the NWP models as possible. Hence, no information from NWP is used in its code; this is the reason why different adjustments in each pixel between simulated and observed radiances have been implemented. The adjustment is performed using both collocated simulated and measured radiances over MSG North area from July 2004 to June 2005. The simulated radiances are obtained using ECMWF analyses (generated on a 0.5° grid) as input to the RTTOV-7. The resulting offset and scale parameters for each IR band are stored in binary matrices on SEVIRI projection.

Second, the performance of the algorithms including local bias correction is compared to the performance of the previous version without local bias correction. The statistical parameters, resulting from the comparison for one year (from July 2005 to December 2006) of radiosounding data and ECMWF analysis (generated on a 0.5° grid) demonstrate the positive impact of the local bias correction in each layer of the Layer Precipitable Water product. The total precipitable water parameter is also compared with the IWV (Integrated water vapour) GPS (Global Positioning System) of two processing centres (GFZ and METO) for one year.

INTRODUCTION
The EUMETSAT SAFNWC aims to support Nowadays and Very Short Range Forecasting and provides software packages in order to retrieve atmospheric parameters at high spatial and temporal resolution from SEVIRI (Spinning Enhanced Visible and Infrared Imager) data. The SEVIRI instrument is a radiometer with 12 spectral bands, between 0.635 and 13.4 μm. Seven of these spectral bands (6.2, 7.3, 8.7, 9.7, 10.8, 12.0 and 13.4 μm) are used to obtain the Layer Precipitable Water. LPW is derived over clear sky areas and produced at pixel by pixel scale, providing information on the water
vapour contained in a vertical column of unit cross-section area in four layers across the troposphere: LPW_BL (low level: 1013-840hPa), LPW_ML (middle level: 840-437hPa) and LPW_HL (high level: 437hPa-top) and LPW_TPW (total). Due to the high resolution of the LPW measurements, the temporal and spatial atmospheric moisture trends may be understood through the LPW variations. Examples of LPW outputs are available at the SAFNWC website (http://nwcsaf.inm.es).

Each LPW parameter is retrieved using a statistical retrieval based on multi-layer perceptron neural networks (MLP-NN). The main advantage of these algorithms versus physical retrieval algorithms is the independence from the Numerical Weather Prediction (NWP) models as well as much lower computer load (imposed by the SAFNWC user requirements). It runs on a PC in less than 3 minutes time together with other 11 products, at pixel by pixel scale every 15 minutes over MSG-N region. A detailed description of the LPW algorithm can be found in Martínez et al. (2007).

The LPW validation is a very import task in order to estimate the accuracy of each product parameter; therefore, a great effort is being performed in this task. This paper details the comparison of the SAFNWC LPW version 1.3 and version 2.0 (included in the 2008 delivery) with independent water vapour measurements in order to improve the algorithm in future versions. A quantitative validation of the LPW SAFNWC parameters with the equivalent parameters extracted from the ECMWF analysis over the MSG North area has been performed from July 2005 to June 2006 and the results are shown in this paper.

LPW_TPW has also been tested with other independent water vapour measurement sources for at least 1-year period: radio-sounding profiles from RAOB stations and Integrated Water Vapour (IWV) from Global positioning system (GPS) ground-based receivers.

**IMPROVEMENT IN V2.0: THE LOCAL RADIANCE-BIAS CORRECTION**

The LPW algorithm was both trained and tested using simulated radiances. 60L-SD database profiles sampled from ECMWF analysis were used as input to RTTOV-7 radiative transfer model in order to obtain these radiances. This database represents diverse profiles from the ECMWF 40-year re-analysis ERA-40. Hence, a radiance-bias correction is required since the 7 IR SEVIRI radiances are used as inputs to the LPW algorithm.

Files containing the local scale and offset were obtained using the procedure described next. Simulated radiances were calculated at each 0.5°x0.5° grid point over MSG-N by means of ECMWF analyses and RTTOV-7 for 1-year period (July 2004 to June 2005, at 00 and 12 UTC). On the other hand, the maximum of SEVIRI radiances at every ECMWF 0.5° grid point surroundings was searched for the same period. For this, only clear-air SEVIRI pixels were considered (using CMa), and maxima were obtained separately over sea and over non-desert land pixels. An estimation of the scale and offset between real and simulated radiances was performed at every 0.5° grid point. A robust regression (IDL LADFIT function) was used in a 5x5 window at every point. These equally spaced estimations of scale and offset were bi-linearly interpolated to the SEVIRI MSG-N points over sea and non-desert land. Land-sea mask was used for this. Finally, they were stored in METEOSAT full-disc matrices (3712x3712) for each IR channel.

![Figure 1: LPW_TPW product for current version V1.3 over MSG-N region](image-url)
The implemented LPW algorithm consists of 3 stages. On the pre-processing step, the above mentioned local IR radiance bias correction has been added. Next, the retrieval of Layer Precipitable Water is performed by a feed-forward neural network (NN). Finally, a local parameter bias correction is applied on the post-processing step (matrices similarly obtained by a robust regression in a 1x1 window, 0.5º grid).

**LPW VALIDATION WITH ECMWF ANALYSIS PROFILES**

ECMWF analyses (00 and 12 UTC), from July 2005 to June 2006, have been downloaded from MARS/ECMWF. The region is defined by the corners (70ºN, 40ºW) and (28ºN, 40ºE), with a grid step of 0.5º. The ECMWF parameters have been remapped to SEVIRI projection (2200 elements, 1019 lines) using McIDAS. One every ten has been extracted to build the validation dataset (220 elements, 101 lines) and only zenith angles lower than 70º are considered. LPW has been reprocessed using the current version V1.3 and the version 2.0 (2008 delivery) from July 2005 to June 2006 at 00 and 12 UTC. To separate clear and cloudy pixels, the CMa (cloud mask) SAFNWC v1.3 has been used (this product uses as inputs ECMWF analysis fields instead of the NWP forecasts). All pixels classified as clear are included in the validation dataset (any other additional constraint has been used to remove data). It was observed in previous works a bad behaviour of land 2D-histogram which was associated with the non-realistic results over desert areas. Therefore, these pixels have been excluded in the validation of the versions shown in this paper. The Global Land Cover supplied by the LANDSAF/MSG group of the Valencia University, contains 25 classes of land cover (among them “bare areas” that includes desert areas), has been used to differentiate desert and non desert pixels.

Current version V1.3 is then compared to the latest version V2.0 for the validation behaviour. The 2D-histograms for LPW_TPW parameter are shown in Figure 3 for the two databases: sea pixels (top) and non-desert land pixels (bottom). Figures on the left show the 2D-histograms for the current version (V1.3) and on the right, the 2D-histograms with the local bias correction (V2.0) are shown. The 2D-histogram performances and the statistical parameters (Figure 4, 5 and Table 1, 2) significantly improve when the local bias correction is applied.

The statistical parameters between each LPW parameter and its equivalent from the ECMWF analysis profiles were calculated and their spatial distributions were represented (see examples for LPW_TPW parameter in Figure 4 and Figure 5 below) over sea and non-desert land pixels for both previous and latest version. Significant improvements are observed from one version to the other: note the general bias reduction as well as the correlation increase and rms error decrease over MSG-N area. Best performances are appreciated over the Baltic area but there is still a bad behaviour over the North-Atlantic region, probably as a consequence of lacking clear-air pixels. Similar improvements are observed over land, where an excellent behaviour is remarkable over North Europe in spite of the cloudy pixel predominance. The rmse are comparable to those over the Peninsula, as it was required in the Product Assessment Review Workshop in 2005. However, a slight tendency to negative biases is observed, as well as a poor behaviour over some pixels which need a further check.
Figure 3: LPW_TPW 2D-histogram comparison of previous version V1.3 (left) and latest version V2.0 (right) over sea-pixels (top) and over non-desert land pixels (bottom)

Figure 4: Comparison of LPW_TPW statistical parameter distribution for previous version V1.3 (left) and latest version V2.0 (right) over sea-pixels (top: bias, middle: correlation, bottom: rmse)
The correlation coefficients and the rms error are included in Table 1 (sea pixels) and in Table 2 (land non-desert pixels) for the four LPW parameters. The correlation coefficient increases and the rms error decreases when the local bias correction is applied in both database (sea and non-desert land) for the four LPW parameters, being more significant the improvements obtained over land pixels.

<table>
<thead>
<tr>
<th>LPW parameters</th>
<th>LPW_BL</th>
<th>LPW_ML</th>
<th>LPW_HL</th>
<th>LPW_TPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm version</td>
<td>V1.3</td>
<td>V2.0</td>
<td>V1.3</td>
<td>V2.0</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.7295</td>
<td>0.7881</td>
<td>0.7448</td>
<td>0.8207</td>
</tr>
<tr>
<td>RMS Error</td>
<td>2.8308</td>
<td>2.5636</td>
<td>2.9198</td>
<td>2.3141</td>
</tr>
</tbody>
</table>

Table 1: Statistical parameters of LPW parameters over sea pixels: comparison of previous version V1.3 and latest version V2.0 (which includes local bias correction)

<table>
<thead>
<tr>
<th>LPW parameters</th>
<th>LPW_BL</th>
<th>LPW_ML</th>
<th>LPW_HL</th>
<th>LPW_TPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm version</td>
<td>V1.3</td>
<td>V2.0</td>
<td>V1.3</td>
<td>V2.0</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.7458</td>
<td>0.7924</td>
<td>0.6823</td>
<td>0.7870</td>
</tr>
<tr>
<td>RMS Error</td>
<td>3.0373</td>
<td>2.5610</td>
<td>3.5941</td>
<td>2.7253</td>
</tr>
</tbody>
</table>

Table 2: Statistical parameters of LPW parameters in non-desert land pixels: comparison of previous version V1.3 and latest version V2.0 (which includes local bias correction)

**LPW COMPARISON WITH RADIOSONDE**

Radiosonde profiles from 102 RAOB stations over MSG-N region at 00 and 12 UTC for the period of 1st July 2005-31st December 2006 are used (provided by the University of Wyoming). LPW has been reprocessed using version 1.3 and 2.0 for that period (CMA SAFNWC v1.3 has been used to depict clear and cloudy pixels). The stations over desert areas have been removed through the desert mask.
The NWCSAF LPW parameters obtained have been compared with the LPW equivalent parameters calculated from the radiosonde data. The statistical parameters obtained with radiosonde over non desert pixels are similar for every LPW parameter. In Figure 6 it is represented the 2-D scatterplots and statistical parameters for LPW_ML and LPW_TPW parameters, showing the improvements from one version to the other (see the statistical parameter values on the tables beside the figures).

![Figure 6: 2-D histogram of LPW parameters versus their equivalent parameters from radiosonde: ML (top) and TPW (bottom) for previous (left) and latest version (right). Tables with statistical parameter results are shown on the right.](image)

During this work it was noticed a significant difference between the station height values (real topography, SEVIRI height and model topography). This could explain the poor behaviour of some stations. Hence, further analysis is required in order to check which stations are reliable.

**LPW_TPW COMPARISON WITH INTEGRATED WATER VAPOUR FROM GPS**

The Global positioning system (GPS) ground-based receivers can work as meteorological sensors. GPS estimations of total precipitable water vapour are available with a high temporal resolution (few minutes), and they are not adversely affected by the presence of clouds. Therefore, European network of Global Positioning System (GPS) receivers are now routinely used to provide near-real-time estimations of precipitable water vapour. It can be used as another independent water vapour measurement source.

NWCSAF LPW_TPW parameter has also been compared with the total column Integrated Water Vapour derived from GPS (IWV_GPS) for two centres (GFZ and METO, 178 GPS sites) in the frame of TOUGH Project. This comparison was done for a 1-year period (1st May 2005- 30th April 2006). The statistical results for one year of IWV GPS from GFZ and METO processing centres are shown in Table 3.

The purpose of this checking is to know the accuracy of LPW_TPW and to identify the potential causes of the discrepancies in order to improve the algorithm. The better shapes of the V2.0 scatterplots for both processing centres (Figure 7 and Figure 8) show the improvements achieved with
the new correction. The LPW_TPW was systematically wetter than the GPS IWV (especially over United Kingdom) on the previous version but this is corrected with the new version (slightly negative biases are found). Best agreement is found between LPW_TPW and IWV from GFZ centre. It is likely that the main discrepancies reflect the impact of the lack of clear-air pixels and bad cloud decontaminations on the LPW algorithm for METO data. Anyhow, the station height dependence for the rmse and bias is not clear.

Figure 7: 2-D histogram comparison between LPW_TPW from SAFNWC and IWV from GPS centre GFZ (Germany) for both previous (V1.3) and latest version (V2.0)

Figure 8: 2-D histogram comparison between LPW_TPW from SAFNWC and IWV from GPS centre METO(UK) for both previous (V1.3) and latest version (V2.0)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>LPW_TPW vs GFZ_</th>
<th>LPW_TPW vs METO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm version</td>
<td>V1.3</td>
<td>V2.0</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.8360</td>
<td>0.8517</td>
</tr>
<tr>
<td>RMSE</td>
<td>4.3384</td>
<td>3.9237</td>
</tr>
<tr>
<td>Bias</td>
<td>1.3665</td>
<td>-0.7529</td>
</tr>
<tr>
<td></td>
<td>2.8542</td>
<td>-1.1120</td>
</tr>
</tbody>
</table>

Table 3: Statistical parameters at each GPS processing centre (METO and GFZ) for both versions.

It can also be noted that the correlation between SAFNWC LPW_TPW and IWV GPS of GFZ (0.8517) is greater than the correlations obtained in the validations previously shown (ECMWF and radiosonde) for version V2.0. In a similar way, the rms error (3.9237) for GFZ is below the values obtained in the other validations (ECMWF and radiosonde). However, the fact of considering different validating periods hinders any conclusion.

CONCLUSIONS

A global radiance bias correction (bias between observed and simulated radiances) were included in previous versions. Validation of those versions revealed a spatial variability on the biases. This paper
shows the improvements of the Layer Precipitable Parameter (LPW) when a local bias correction is implemented in its algorithm (latest version 2.0 which will be included in the 2008 delivery).

A validation of the LPW parameters was performed by means of the equivalent parameters obtained from ECMWF analysis profiles, radiosondes and GPS (1-year period databases at least). The inter-comparison was performed with both previous version V1.3 and latest version V2.0, in order to better appreciate the improvements observed for every LPW parameter.

The statistical parameters obtained with radiosonde and ECMWF on non desert pixels are similar for LPW_BL, LPW_ML, LPW_HL and LPW_TPW. The statistical parameters also show that the parameter accuracy increases when the local bias correction is applied: a general increase of correlation is observed as well as a reduction of bias and rms error over both sea and non-desert land pixels, especially over North-Europe. On the contrary, desert areas and their emissivity problems associated need to be carefully analysed in future works. This may be done with the help of other data sources as IASI, atlas emissivities, etc.

The inter-comparison with GPS Integrated Water Vapour supports the previous results, showing an improvement for the latest version V2.0 in both processing centres (GFZ and METO). At every validation a slight tendency to negative biases is highlighted, which means an overcorrection of radiances at some points. In addition to this, there are some regions were the performance is not satisfactory enough (low correlation, high rmse). Hence, a further study of these zones is required for future works in order to improve the LPW parameter accuracy.

According to the decisions of the last NWCSAF Steering Group Meeting, the PGE07 version 2.0 coded in the 2008 delivery will only have three outputs (LPW_BL, LPW_ML, LPW_HL), the PGE06 being replaced by the current LPW_TPW. The latter will run in the 2008 delivery as PGE06 version 3.0.

ACKNOWLEDGEMENTS

This study was carried out in the SAFNWC validation activities framework. The SAFNWC is an agreement of co-operation between EUMETSAT and INM (the SAFNWC consortium is constituted by Météo-France, SMHI, ZAMG and INM).

We would like to thank the University of Wyoming (Department of Atmospheric Science) which kindly provided us the radiosonde data.

Thanks also to the INM Numerical Prediction Department which supplied us the GPS data in the frame of the TOUGH Project. TOUGH is a shared-cost project (contract EVG1-CT-2002-00080) co-funded by the Research DG of the European Commission within the RDT activities of the Environment and Sustainable Development subprogram (5th Framework Programme).

REFERENCES