Estimation of tropospheric profiles using COSMIC GPS radio occultation data with neural networks

Fabrizio Pelliccia¹, Stefania Bonafoni¹, Patrizia Basili¹, Roberta Anniballe¹, Nazzareno Pierdicca² and Piero Ciotti³

¹Dept. of Electronic and Information Engineering, Univ. of Perugia, via Duranti 93 06125, Perugia, Italy
E-mail: stefania.bonafoni@diei.unipg.it
²Dept. of Electrical Engineering, Univ. “Sapienza” of Rome, via Eudossiana 18 00184, Rome, Italy
³Dept. of Electrical and Information Engineering, Univ. of L’Aquila, 67040 Poggio di Roio, L’Aquila, Italy

Abstract
In this work we have proposed a method based on neural networks to directly retrieve dry and wet refractivity and dry pressure profiles in troposphere (which in turn can be used to obtain temperature and humidity profiles) by using COSMIC GPS radio occultation (RO) data. To overcome the constraint of an independent knowledge of one atmospheric parameter at each GPS occultation, we trained different neural networks with refractivity profiles as input, while the targets were the dry and wet refractivity profiles and the dry pressure profiles obtained from the ECMWF data. Finally, selecting two cases, we have compared the estimated profiles, for each neural network approach, with the corresponding radiosounding profiles.

Keywords: Radio occultation, GPS-LEO, atmospheric profiling, neural network.

Stima di profili troposferici da Radio Occultamento GPS COSMIC tramite reti neurali

Riassunto
In questo lavoro abbiamo proposto un metodo basato su reti neurali per determinare profili di rifrattività, temperatura, pressione e umidità a partire da dati di radio occultazione (RO) COSMIC GPS. Abbiamo allenato diverse reti neurali, ponendo in ingresso profili di rifrattività ottenuti dai parametri geometrici di RO e in uscita profili di rifrattività in condizioni di aria secca e umida e profili di pressione parziale di aria secca, estrapolati da dati ECMWF, con lo scopo di aggirare il vincolo della necessità di conoscere un parametro atmosferico da fonti indipendenti, ad ogni RO. Infine, selezionando due casi, abbiamo confrontato i profili stimati, per ogni tipo di rete neurale, con quelli forniti da corrispondenti radiosondaggi.

Parole chiave: Radio occultazione, GPS-LEO, profili atmosferici, reti neurali.

Introduction
Radio occultation (RO) with Global Positioning System (GPS) is a sounding technique for the atmospheric profiling useful for numerical weather prediction and climatological
The system employs GPS receivers placed on Low-Earth Orbit (LEO) satellites to sound the Earth’s neutral atmosphere and ionosphere. The aim is the evaluation of the additional delay affecting a radio signal when passing through the atmosphere due to the refractivity index variations [Gorbunov and Sokolovskiy, 1993; Rius et al., 1998]. GPS occultations works under all-weather conditions due to the insensitivity of the GPS signal wavelength to scattering by clouds and precipitation, with a horizontal resolution of about 200 km in the direction along the occulted link and a resolution of 1 km or better in the cross-link and vertical directions [Kursinski et al., 1997].

The GPS-RO technique is exploited to obtain profiles of refractivity, temperature, pressure and humidity in the atmosphere at global scale, and several investigations have demonstrated that the retrieval accuracies are comparable to traditional atmospheric remote sensing techniques [Hardy et al., 1994; Kursinski et al., 1995]. Even though the atmospheric refractivity profiling by radio occultation is a well-defined problem, care must be taken to analyze factors affecting the occulted signal (multipath, satellite motion etc.) and to compute temperature and particularly humidity profiles from refractivity [Kursinski et al., 1995]. The accuracy of atmospheric profile estimation is affected by the presence of water vapour in the atmosphere, that complicates the interpretation of the refractivity [O’Sullivan et al., 2000]. Refractivity profiles can be converted in a straightforward way into pressure and temperature profiles in dry regions, but when the water vapour is not negligible, the retrieval of atmospheric profiles from GPS refractivity measurements is possible only given an independent knowledge of temperature profile derived from independent observations (i.e. radiosoundings or data from atmospheric numerical modeling).

In this paper, we have proposed a new retrieval algorithm based on multilayer perceptron neural networks to derive profile of atmospheric parameters from RO refractivity overcoming the requirement for external information availability at each GPS occultation. In particular, we have implemented and compared different neural network trainings, employing three neural networks at each approach. The inputs are refractivity profiles computed from the occultation parameters observed by the COSMIC (Constellation Observing System for Meteorology Ionosphere and Climate) satellites and provided by the COSMIC Data Analysis and Archive Center (CDAAC) of Boulder (Colorado) [Anthes et al., 2000]. The targets employed in the training are the dry and wet refractivity profiles, together with the dry pressure ones, obtained from the space-time co-located European Centre for Medium-Range Weather Forecast (ECMWF) analysis data.

The neural network training and the following independent test were performed over the entire land area between Tropics, split into a desert and vegetation zone, by using the available data set of 445 refractivity profiles on summer 2006, from July 17 to August 18. The choice of splitting the entire available data set into desert and vegetation data sets leads to design neural networks more efficient in the ability to retrieve the atmospheric profiles of the two zones, since trained with homogeneous atmospheric conditions.

To evaluate the performances of the different approaches, we have computed errors affecting the estimated profiles with respect to ECMWF analysis, assumed as the truth. Such a choice of ECMWF data as reference in the comparisons was also adopted by other authors [Kursinski and Hajj, 2001; Kuo et al., 2004; Beyerle et al., 2006], considering that these data provide global coverage and high spatial resolution reconstruction of the atmosphere. We have also compared the estimated profiles with the ones provided by radiosoundings,
but only few cases matched the temporal and spatial correspondence between RO events and radiosondes. 

The proposed technique shows the possibility to estimate tropospheric profiles included the wet ones only from RO refractivity, after the settlement of the training phase of neural networks, and hence the possibility to increase the atmospheric observations, thanks to a wide spatial coverage of RO soundings. For this purpose, the employment of neural networks proved useful and hence different training approaches were tested and evaluated.

**Atmospheric refractivity profiles from RO**

GPS-RO observations are performed in a limb-scanning mode, where in the geometrical optics approximation a ray passing through the atmosphere is refracted due to the vertical refractive profile variations. The overall effect of the atmosphere can be characterized by a total bending angle profile \( \alpha \), an asymptotic impact parameter \( a \) and a tangent radius \( r_p \), as shown in Figure 1, and their variations depend primarily on the vertical profile of refractive index [Kursinski et al., 1997].

\[
\begin{align*}
\text{Figure 1 – Instantaneous GPS-LEO occultation parameters.}
\end{align*}
\]

With the assumption of local spherical symmetry, the refraction index profile \( n \) can be retrieved from profile of \( \alpha \) as a function of \( a \) during an occultation by using an Abel transformation as in [Fjeldbo et al., 1971]:

\[
n(r_p) = \exp\left\{ \frac{1}{\pi} \int_{a_p}^{\infty} \frac{a(a)}{\sqrt{a^2 - a^2_p}} \, da \right\} \quad [1]
\]

where \( a_p = n(r_p) \) \( r_p \) is the impact parameter for the ray whose tangent radius is \( r_p \). The refractivity profile is then \( N=(n-1) \times 10^6 \).

First, we have collected 445 FORMOSAT-3/COSMIC radio occultation events provided by CDAAC [COSMIC website], covering the inter-tropical land area during the period from July 17 to August 18, 2006. The FORMOSAT-3/COSMIC is a joint Taiwan-U.S. mission of six micro-satellites, launched on April 2006, with onboard receivers registering phase and amplitude of radio waves at the two GPS carrier frequencies (1575.42 and 1227.6 MHz).
Then we have computed the refractivity profile as in [1] using the impact parameters and the bending angles profiles provided by CDAAC, vertically spaced from 4 m to 50 m in the low and high atmosphere, respectively.

Finally, we have produced a complete database matching GPS occultations with the corresponding ECMWF observations co-located in time and space, with a maximum time difference of 1 hour and a geographical coordinate distance less than 0.5°. We have used ECMWF analysis data with full vertical resolution of 91 levels, horizontal resolution of 0.5° and time resolution of 6 hours.

In order to train and test the neural networks we have created two databases (desert and vegetation) of refractivity profiles sampled at the same altitude intervals for each RO observation, with a vertical resolution of 200 m. As a result, the desert database contains 98 profiles with 56 fixed altitude levels, representing the atmosphere from 1.1 to 12.1 km, i.e. the troposphere, and the vegetation database contains 347 profiles with 54 fixed altitude levels, representing the atmosphere from 1.4 to 12 km.

We have chosen RO with the minimum altitude of 1.1 km and 1.4 km over desert and vegetation area respectively as a trade-off between the requirement of an adequate number of observations to train the networks and the need of tropospheric profile estimation as much as possible close to the surface, considering that several RO observations do not reach low atmospheric levels.

**Atmospheric profiles from refractivity**

The atmospheric refractivity at microwave wavelength is given by [Smith and Weintraub, 1953]:

$$N = 77.6 \frac{P_d}{T} + 72 \frac{P_w}{T} + 3.75 \cdot 10^3 \frac{P_w}{T^2}$$ \[2\]

where $P_d$ is the pressure of dry air in mb, $P_w$ the partial pressure of water vapour in mb, $T$ is the atmospheric temperature in Kelvin. To solve for $T$, $P_d$ and $P_w$ given $N$, we use the additional constraints of ideal gas and hydrostatic equilibrium laws, respectively, as:

$$\rho = \frac{P}{T} \frac{M_d}{R_0} + \frac{P_w}{T} \frac{(M_w - M_d)}{R_0}$$ \[3\]

$$dP(z) = -g\rho(z)dz$$ \[4\]

where $\rho(z)$ is the air density in kg m$^{-3}$, $P = P_d + P_w$, $M_d$ and $M_w$ are respectively the mean molecular mass of dry air and water vapour, $R_0$ is the universal gas constant, $g$ the gravitation acceleration. Given $N$, we have a system of three equations and four unknowns ($T$, $P_d$, $P_w$ and $\rho$), and therefore it is necessary to have an independent knowledge of one of the four parameters to solve the atmospheric profiling problem [Kursinski and Hajj, 2001; Vespe et al., 2002]. For instance, some authors exploit the temperature profile derived from independent observations or weather analysis [Kursinski et al., 1997]. In this study, a method based on neural networks is proposed to retrieve atmospheric profiles overcoming the need of knowing the temperature profile.
at each GPS occultation.

**Target profiles from ECMWF data**

The targets for the neural network training and the references for the following test are the dry and wet refractivity profiles and the dry pressure profiles. These profiles were obtained by processing analysis data provided by the ECMWF, such as the logarithm of pressure, the specific humidity and the temperature, belonging to the “ECMWF 91 model levels” data set, representing the atmosphere from 75 km to the ground with a vertical resolution spanning from 5 km to 25 m in the high and low atmosphere, respectively and with a spatial grid of 0.5° [ECMWF website].

From ECMWF analysis data we have computed the water vapour partial pressure profile \( P_w \) by using the relationship between \( P_w \) and mixing ratio:

\[
P_w = \frac{wP}{R_{dry} + w} \quad [5]
\]

where \( w \) is the mixing ratio in kg kg\(^{-1}\) derived from the specific humidity, \( R_{dry} \) and \( R_{vap} \) are the gas constant for dry air and water vapour respectively and \( P \) is the total pressure in mb. Hence, the dry pressure profile has been obtained by subtracting the partial pressure of water vapour from the total pressure profile. Finally we have computed the refractivity by using [2], interpolating at the same vertical levels of the refractivity profiles obtained from Abel transformation.

**Neural network approaches**

As previously described, to solve the atmospheric profiling problem from RO overcoming the need of external information, we have considered different neural network approaches for both desert and vegetation databases. For each database and each approach, we have trained three neural networks where predictors are the total refractivity profiles \( N \) computed from RO data using [1], and the targets are the dry pressure profiles \( P_d \) computed from ECMWF data and the dry \( N_d \) and wet \( N_w \) refractivity profiles, respectively defined by:

\[
N_d = 77.6 \frac{P_d}{T} \quad [6]
\]

and

\[
N_w = 72 \frac{P_w}{T} + 3.75 \cdot 10^5 \frac{P_w}{T^2} \quad [7]
\]

Over desert area between Tropics, the neural network training and test were performed by using 88 profiles for the training and the remaining 10, randomly selected, for the independent test, that represent 90% and 10% of the entire desert data set. Over vegetation area between Tropics, we have proceeded in the same way by choosing 312 profiles for the training and the remaining 35 for the independent test.
The different approaches for the neural networks learning are described in the following sub-sections.

**Early stopping technique**

At first, we have applied the early stopping technique [Demuth et al., 2008] to define the optimal number of training epochs dividing the available dataset in two disjoint subsets: the training set and the validation set. The first one is used for the learning itself, the second one to choose the number of training epochs. Learning ends when the error on the validation set begins to rise even if the error on the training set could be further reduced, improving the ability of generalization of the network. Since overtraining could occur even on the validation set, a further test subset should be used to assess the capacity of generalization of the network. Then we have divided the training desert data set and the training vegetation data set in three subsets respectively: the training subset used for the learning itself, the validation subset and the test subset, by assigning them randomly the 70% (62 and 218 samples for desert and vegetation lands respectively), the 15% (13 and 47 samples) and the 15% (13 and 47 samples) of the whole data sets, respectively.

Instead of the standard back-propagation, we have used the Levenberg-Marquardt optimization that is often the fastest back-propagation algorithm for training moderate-sized feed-forward neural networks, in agreement with the early stopping technique [Marquardt, 1963; Hagan and Menhaj, 1994].

**Early stopping technique with Principal Component Analysis**

In the second approach, we have applied the same early stopping technique as described above, with the same division of the available data set, but we have processed the input and target matrices by the Principal Component Analysis (PCA). The PCA decomposes the 56-level and 54-level profiles, for desert and vegetation area respectively, on a basis of empirical orthogonal functions called principal components [Smith and Woolf, 1976]. The PCA permits a reduction of the number of descriptive profile levels by exploiting the correlation among values at different altitudes, ensuring a faster processing and a reduction of computer memory requirements in comparison with the original data (full profiles). We have chosen to employ a number of principal components representing the 99.9% of the total variance of the original data [Demuth et al., 2008], leading to the use of only 15 and 14 principal components for the total refractivity instead of the original 56 and 54 levels, over desert and vegetation lands respectively. Concerning the neural network targets, the number of components for dry refractivity, wet refractivity and dry pressure profiles are 11, 14 and 8 for the desert area and 11, 11 and 7 for the vegetation area.

**Cross validation technique**

Finally, we have applied the cross validation technique, together with the early stopping, useful in the case the available data set contains few profiles for training and testing neural networks. Cross validation technique consists in dividing the whole considered data set in K subsets: the training is performed with the profiles belonging to K-1 subsets and the validation with the profiles of the remaining subset. This process recurs K times by changing the validation set every time. The percentage of profiles used to train the neural networks is 1-1/K and, in this way, all the profiles are used in the learning and validating phase in turn.
The only drawback of this technique is the need to repeat the learning K times, increasing the computational costs. For this reason, only profiles reduced with the PCA technique were employed, so cutting down the overall processing time. Therefore, we have trained the neural networks with the cross validation technique applying both K=4 and K=8, that is with a validation subset corresponding to 25% and 12.5% of training data set. Then we have divided the training desert data set (88 samples) and the training vegetation data set (312 samples) in the training subset and in the validation subset by assigning them, for K=4, the 75% (66 and 234 samples for desert and vegetation lands respectively) and the 25% (22 and 78 samples) of the whole data sets. Similarly, for K=8, the 87.5% (77 and 273 samples for desert and vegetation lands respectively) and the 12.5% (11 and 39 samples) of data sets.

**Choice of neural network architectures**

For each technique, we have considered feed-forward neural networks having, besides the input layer, a number (1 to 3) of hidden layers with tan-sigmoid transfer functions and an output layer with linear transfer functions [Demuth e al. 2008]. To select the most suitable architecture, we have used a growing technique adding 1 neuron in the first hidden layer at each training session, until a maximum of 20 neurons for each hidden layer. We have considered a maximum of 3 hidden layers, choosing among the possible combinations the architecture with the lower Root Mean Square (RMS) error computed comparing the network outputs of the test session with the corresponding ECMWF profiles, where the test session employs the profiles not used in the training phase (10 and 35 observations from desert and vegetation databases, respectively).

The best neural network topologies in terms of performances for the dry refractivity, wet refractivity and dry pressure profiles retrieval are reported in Table 1 and in Table 2 for desert and vegetation area, respectively.

**Results**

Performance and generalization capability of each neural network approach meet the eye in Figure 2 and Figure 3, where the RMS error profiles for \( N (N=N_d+N_w) \) employing the independent test set of 10 and 35 occultations for desert and vegetation zones are shown superimposed to the corresponding ECMWF standard deviation profiles. The ECMWF standard deviation can be assumed as an index of the climatological variability of a given atmospheric parameter, than a good accuracy of a retrieval algorithm is obtained when its RMS error is clearly below it.

As underlined in the previous sections, choosing to train three networks for each approach enables us to retrieve atmospheric profiles without the constraint of temperature profile availability at each GPS occultation, only using [2]. In particular, with the availability of \( N_d, N_w \) and \( P_d \), first we can solve for temperature \( T \) in a straightforward way from the dry refractivity defined in [6], and then for partial pressure of water vapour \( P_w \) from the wet refractivity defined in [7].

In Figure 4 and Figure 5 the RMS error profiles for \( T \) and in Figure 6 and Figure 7 the RMS error profiles for \( P_w \) are shown superimposed to the corresponding ECMWF standard deviation profiles, for desert and vegetation zone respectively.
Table 1 – Best neural network topologies over desert zone obtained for each neural network approach: input, HL1, HL2, HL3 and output columns report the number of neurons for the input, hidden layer 1, hidden layer 2, hidden layer 3 and output, respectively. Each approach employs 3 neural networks, named N Dry (for dry refractivity estimation), N Wet (for wet refractivity estimation) and P Dry (for dry pressure estimation).

<table>
<thead>
<tr>
<th>Desert zone</th>
<th>EARLY STOPPING - PCA</th>
<th>EARLY STOPPING – Full Profile</th>
<th>CROSS VALIDATION (25%) - PCA</th>
<th>CROSS VALIDATION (12.5%) - PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Input</td>
<td>HL1</td>
<td>HL2</td>
<td>HL3</td>
</tr>
<tr>
<td>N Dry</td>
<td>15</td>
<td>11</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>N Wet</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>P Dry</td>
<td>15</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

By comparing the proposed approaches, it’s noticeable that each one produces good results, especially the cross validation with the PCA approach, particularly suited for this event characterized by the shortage of profiles belonging to the data sets, most of all the desert ones. However, considering these approaches from the standpoint of the computational cost in terms of time and memory requirements, the cross validation with the PCA is less expensive than the early stopping with the full profiles.

Finally we have compared pressure, temperature and water vapor partial pressure profiles obtained from the independent test with the ones provided by radiosoundings (RAOBs). The RAOB data are provided by National Center for Atmospheric Research (NCAR) mass store and collected by CDAAC. Considering both desert and vegetation zone radio occultation events of the independent test phase (45 events), we have found 15 matches between RO soundings and RAOBs: finally we have investigated only two cases showing a distance between the “occultation point” and the radiosonde launch site lower than 50 km. The “occultation point” is defined as the point on the Earth’s surface to which the retrieved refractivity profile is assigned, located under the perigee point of the bended ray where the excess path exceed 500 m. (Kuo et al. 2004). The two analyzed cases are relative to events occurred over Hanoi (Vietnam) on August 1 2006, and over Rio de Janeiro (Brazil) on August 15 2006. The distance between the occultation point and the RAOB is of 18.29 km for Rio de Janeiro and of 34.49 km for Hanoi. In most examined cases these distances are included in a range from 100 km to 400 km and for this reason these events have not been considered.
Table 2 – Best neural network topologies over vegetation zone obtained for each neural network approach: input, HL1, HL2, HL3 and output columns report the number of neurons for the input, hidden layer 1, hidden layer 2, hidden layer 3 and output, respectively. Each approach employs 3 neural networks, named N Dry (for dry refractivity estimation), N Wet (for wet refractivity estimation) and P Dry (for dry pressure estimation).

<table>
<thead>
<tr>
<th>Vegetation zone</th>
<th>Input</th>
<th>HL1</th>
<th>HL2</th>
<th>HL3</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EARLY STOPPING - PCA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Dry</td>
<td>14</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>N Wet</td>
<td>14</td>
<td>12</td>
<td>8</td>
<td>-</td>
<td>11</td>
</tr>
<tr>
<td>P Dry</td>
<td>14</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td><strong>EARLY STOPPING – Full Profile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Dry</td>
<td>54</td>
<td>15</td>
<td>8</td>
<td>-</td>
<td>54</td>
</tr>
<tr>
<td>N Wet</td>
<td>54</td>
<td>12</td>
<td>5</td>
<td>-</td>
<td>54</td>
</tr>
<tr>
<td>P Dry</td>
<td>54</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>54</td>
</tr>
<tr>
<td><strong>CROSS VALIDATION (25%) - PCA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Dry</td>
<td>14</td>
<td>10</td>
<td>10</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>N Wet</td>
<td>14</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>11</td>
</tr>
<tr>
<td>P Dry</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td><strong>CROSS VALIDATION (12.5%) - PCA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Dry</td>
<td>14</td>
<td>17</td>
<td>9</td>
<td>-</td>
<td>11</td>
</tr>
<tr>
<td>N Wet</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>P Dry</td>
<td>14</td>
<td>13</td>
<td>11</td>
<td>-</td>
<td>7</td>
</tr>
</tbody>
</table>

![Figure 2](image)

**Figure 2** – Desert zone, neural network independent test (10 occultations): RMS error profiles for $N$ using the different neural network approaches.

Figure 8, Figure 9 and Figure 10 show the difference of pressure, temperature and water vapor pressure profiles respectively between estimations and radiosounding over Hanoi. Figure 11, Figure 12 and Figure 13 represent the same differences over Rio de Janeiro.
The comparisons show a large error in the lower part of the profiles (around 2 km from surface), especially for the water vapor pressure, while the comparison between the absolute mean difference values of the analyzed profiles, reported in Table 3, highlights the different behaviors of the training approaches. Also, in both cases, the early stopping technique with full profiles provides the best performances in terms of temperature, but the neural network performances are penalized by the discrepancy near the surface.

**Conclusions**

In this work, we have proposed a method to estimate profiles of refractivity, temperature, pressure of dry air and of water vapor in the troposphere from FORMOSAT-3/COSMIC.
GPS radio occultations, in particular we have considered measurements taken over the entire land area between Tropics during summer time.
To overcome the necessity to know the true temperature profile at each occultation, we have trained three neural networks, by using different approaches, with targets that permit to solve the atmospheric profiling problem reflecting the physical constraints learned from the ECMWF analysis data.

The results have shown good performances of the neural networks training them both with the principal component analysis for a fast and less expensive approach, and without them.
preserving all the atmospheric profile informative content, exhibiting a fairly good accuracy for temperature and partial pressure of water vapor profiles.

![Image](image_url)

**Figure 7 – Vegetation zone, neural network independent test** (35 occultations): RMS error profiles for $P_w$ using the different neural network approaches.

**Table 3 – Absolute mean differences of pressure ($P$), temperature ($T$) and wet pressure ($P_{Wet}$) between radiosounding profiles and estimates from RO (independent test) using the different approaches. E.S.: Early Stopping; C.V.: Cross Validation.**

<table>
<thead>
<tr>
<th>Cases</th>
<th>Parameters</th>
<th>Absolute mean differences between RAOB profiles and estimates from RO</th>
<th>Absolute mean differences between RAOB profiles and estimates from RO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>E.S. PCA</td>
<td>E.S. Full Profile</td>
</tr>
<tr>
<td>Hanoi</td>
<td>$P$</td>
<td>1.29 mb</td>
<td>4.55 mb</td>
</tr>
<tr>
<td></td>
<td>$T$</td>
<td>3.41 K</td>
<td>0.75 K</td>
</tr>
<tr>
<td></td>
<td>$P_{Wet}$</td>
<td>1.02 mb</td>
<td>1.28 mb</td>
</tr>
<tr>
<td>Rio de Janeiro</td>
<td>$P$</td>
<td>1.57 mb</td>
<td>0.89 mb</td>
</tr>
<tr>
<td></td>
<td>$T$</td>
<td>3.20 K</td>
<td>1.83 K</td>
</tr>
<tr>
<td></td>
<td>$P_{Wet}$</td>
<td>1.16 mb</td>
<td>1.28 mb</td>
</tr>
</tbody>
</table>

Furthermore, such results can be improved with the future availability of a larger data set for training the neural networks, covering a wider variety of atmospheric conditions spanning a larger seasonal and geographical extension.

The comparisons with radiosondes have shown a large error in the lower part of the profile, especially for the water vapor pressure, even if a greater number of events to compare
would be necessary for a better understanding.

In conclusion, our analysis aims to show the possibility to retrieve each atmospheric parameter included the wet ones only from RO refractivity, after the settlement of the training phase of neural networks, and then the ability to increase the atmospheric observations thanks to a wide spatial coverage of RO soundings on the Earth (only COSMIC mission produces 1500-2500 profiles per day). The limit of this approach is that the exploitation of the informative contribution brought by RO soundings is in some way connected to the necessary employment of the ECMWF atmospheric model profiles as targets for the neural network training.

![Figure 8](image)

Figure 8 – Difference profiles for $P$ between RAOB and estimation from RO using the different neural network approaches. Hanoi (Vietnam), August 1 2006.

![Figure 9](image)

Figure 9 – Difference profiles for $T$ between RAOB and estimation from RO using the different neural network approaches. Hanoi (Vietnam), August 1 2006.

**Acknowledgements**

The work has been sponsored by the ASI, Italian Space Agency, through the Thales Alenia Space. We wish to thank COSMIC Data Analysis and Archive Center (CDAAC) of Boulder (Colorado) for the availability of the occultation data.
Figure 10 – Difference profiles for $P_w$ between RAOB and estimation from RO using the different neural network approaches. Hanoi (Vietnam), August 1 2006.

Figure 11 – Difference profiles for $P$ between RAOB and estimation from RO using the different neural network approaches. Rio de Janeiro (Brazil), August 15 2006.

Figure 12 – Difference profiles for $T$ between RAOB and estimation from RO using the different neural network approaches. Rio de Janeiro (Brazil), August 15 2006.
Figure 13 – Difference profiles for $P_w$ between RAOB and estimation from RO using the different neural network approaches. Rio de Janeiro (Brazil), August 15 2006.

References
ECMWF website. Available via www.ecmwf.int.

Received 07/04/2009, accepted 11/09/2009.