A neural network based ionospheric model for Arecibo

M. Friedrich[a], M. Fankhauser[a], E. Oyeyemi[b,c], and L.A. McKinnell[b,c]

[a]*Graz University of Technology Inffeldgasse 12, A–8010 Graz, Austria, e-mail: martin.friedrich@tugraz.at*
[b]*Hermanus Magnetic Observatory, Hermanus, South Africa*
[c]*Rhodes University, Grahamstown, South Africa*

ABSTRACT

The Arecibo Observatory (18°N, 66°W) has the world's largest single dish antenna (300 m dia.). Beyond radio astronomy it can also operate as an incoherent scatter radar and in that mode its figure-of-merit makes it also one of the most powerful world-wide. For the present purpose the electron density data available on the web, from the beginning with the first erratic measurements in 1966 up to 2004 inclusive, were downloaded. The measurements range from about 100 km to beyond 700 km and are essentially evenly distributed, *i.e*. not dedicated to measure specific geophysical events. From manually edited / inspected data a neural network (NN) was established with season, hour of the day, solar activity and *Kp* as the input parameters. The performance of this model is checked against a - likewise NN based - global model of *foF2*, a measure of the maximum electron density of the ionosphere. Considering the diverse data sources and assumptions of the two models it can be concluded that they agree remarkably well.

1. INTRODUCTION

Traditionally the ionosphere was investigated from the ground by ionosondes which - at least in principle - yield electron densities up to the height of the maximum density (hmF_2). From the ground electron density profiles beyond the *F*-region peak can only be obtained by much more complex installations such as incoherent scatter radars. The largest and one of the most powerful is the Arecibo Observatory in Puerto Rico. Various approaches to process data in order to establish empirical models are conceivable, such as simply binning the data or to assume physics-based analytical function whose coefficients are obtained from the data. We here use the neural network (NN) approach to establish an empirical, single-station model for Arecibo. A Neural Network (NN) is essentially a weighted, multi-dimensional interpolation procedure. The advantage of the method is that one does not need to "guess" a function for an expected dependence, but rather let the procedure find the relation between the input parameters "suspected" to be relevant for the desired output. The penalty of the method is that extrapolated results, *i.e.* outside the so-called input space, are usually utterly unrealistic.

For communication purposes the peak electron density is the most important parameter since its value determines up to what frequency the ionosphere can be used as a reflector. Traditionally this parameter is obtained by vertical sounding with an ionosonde and recording the highest frequency that the ionosphere reflects at vertical incidence (foF_2). The most commonly used models to predict foF_2 are from URSI and CCIR; both of these analytical models are based on ionosonde data of but limited time coverage. The global foF_2 model by Oyeyemi *et al*. (2005) was already found to better predict measured (ionosonde) foF_2 than both of the above models; here it is tested against the corresponding results obtained from a single station model in its early stage of development.

2. NEURAL NETWORK MODELS

Separate empirical models, using the NN technique, were developed entirely independently, one for the whole ionosphere above Arecibo and the other for foF_2 from a large number of globally distributed stations. The technique of NNs has been employed by several groups for modelling ionospheric parameters (Wintoft and Cander, 1999; McKinnell and Poole, 2000; 2001; Tulunay *et al*., 2000; 2006). In this paper, we do not attempt to explain NNs, other than to describe the method as a computer code that is trained to learn the relationship between a given set of input parameters and a corresponding output parameter. The reader is referred to Haykin (1994) for more information on the NN procedure. The purpose of this paper is to compare two NN-based ionospheric models: one a single station model for the electron densities in the altitude range from 100 km to 700 km, and the other a global model for the peak electron density.

2.1. ARECIBO MODEL

In the *D*-region the lifetime of free electrons is of the order of seconds and below, say, 100 km it can therefore safely be assumed that electron densities can adequately be described by the geophysical parameters prevailing at the time one seeks to describe. At greater heights, such as the *F*-region, this assumption will generally *not* hold because of the much longer lifetime of free electrons and ions. In this preliminary modelling effort we nonetheless assume steady-state, fully being aware that this is largely not an acceptable simplification, but will serve as complex test of the quality of the data to be used later in a more sophisticated version of the model. All of the about 100,000 profiles were interactively inspected and - if necessary - edited. About 1.6 Mio. data points survived the manual editing processes and are used in the model. The seasonal and diurnal distribution of the data is fairly even (not more than a factor of two between the best/poorest hour of the day, or the best /poorest covered month of the year). We seek to describe the whole ionosphere – as covered by the Arecibo incoherent scatter radar – by the following six geophysical parameters:

- day number (season)
- hour (in UT)
- Chapman function (solar zenith angle)
- altitude
- geomagnetic index *Kp*
- solar activity ($F_{10.7}$).

In order to bypass the necessity of establishing both cyclic components in season (day of the year) and hour number (time of the day), the same data set is added before and after the period of interest (day or year, respectively). In the final result these “artificially” added pieces are neglected. With this procedure the results at the beginning and the end of the year (or the day) wrap around the ends with values differing only a few percent. The present network architecture consists of two hidden layers with 15 nodes each. The data were grouped in different sets for training, training validation and test. The training procedure is continued until the sum of all rms error factors (of each data point *vs.* its corresponding model value) after consecutive iterations no longer decreases (Fankhauser, 2005); continued iterations would make the NN memorise the training data and thus lose generalisation capability. Figure 1 shows the rms error factor as a function of altitude of the model to be described in the following.

2.2. GLOBAL foF_2 MODEL

The model of foF_2, described in detail by Oyeyemi *et al*. (2005), presents the first attempt at predicting the peak electron density on a global basis using NNs and a much longer time series than previous models. The 59 stations used in the NN training cover latitudes from 77.9°S to 74.7°N and the data are from the years 1964 to 1986. In contrast to the Arecibo model sketched above, this model does consider long-term effects of geophysical parameters. Notably the solar activity (sunspot number) and the geomagnetic index *Ap* are used as running means of the time preceding the prediction, namely two months for the sunspot number and two days for *Ap*. Other input parameters are geographic latitude, geomagnetic latitude *and* longitude, day number, and universal time, the latter two in both cyclic components. Three hidden layers with a total of 90 nodes were found to yield the best results with the aim of minimising the root mean square (rms) error *difference* between the measured and predicted foF_2; this differs from the desired optimum of the Arecibo model where the *ratio* between measured and predicted electron density is sought (actually, mathematically the difference in $\log N_e$).

It should also be noted here that no Arecibo data were included in the 59 stations used to train the global foF_2 model. Therefore this is a true test of how the global model performs within a region not covered by the training set in comparison with a single station model designed for that particular area. Also, after training, a NN should provide the best average prediction for a given set of inputs; *i.e.* a NN should provide a generalised solution.

3. COMPARISON TO MEASUREMENTS

The averaged rms error factor (data point/corresponding model value) of all data points is not a very revealing quantity, but also its variation with altitude (Fig. 1) is not as elucidating as a direct comparison between a series of measurements and the corresponding prediction. In Figure 2 we show contour plots of the measured electron densities of a whole UT day (February 4th, 1989) together with what the single station Arecibo model predicts with the geophysical parameters prevailing during that day. The bottom panel depicts the prediction error expressed as the ratio between measured data and corresponding prediction (on a logarithmic scale, hence "0" is a perfect prediction). In this example the error is at most a factor of 3. In the third panel - with all due caution - one can see

downward propagating enhancements. For the purpose of amplifying minor variations in the *D*- and *E*-regions a similar exercise has been made earlier by Friedrich *et al*. (2006) with data from EISCAT-mainland. Figure 3 depicts the variation of the NmF_2, the value of the peak electron density of the same day as in the previous figure. The nocturnal electron density decay until sunrise is better predicted by the local model, whereas - perhaps a little surprisingly - the global model, interrogated for Arecibo, better reproduces the rapid increase at sunrise. Figure 4 shows a comparison of the NmF_2 prediction between the two models. We chose the year 1993 because it is best covered by Arecibo measurements; hence the local model is assumed to yield the most realistic results for the geophysical conditions of this time period. Depicted are the midnight and noon NmF_2 according to the two models; however we do not show a comparison with the measured data due to the only intermittent operation of the incoherent scatter radar. The first impression is that the local single station model shows more structure than the values obtained from the global foF_2 model. This is not unexpected since the global model is partly based on smoothed input parameters, whereas the local model uses the *Kp* and $F_{10.7}$ at exactly 04 and 16 UT (*i.e*. local midnight and noon, respectively). In Figure 5 we show the predicted diurnal variation of electron densities at various altitudes for a spring day and median solar and magnetic activity. The variations look reasonable: there is a constant decay at night and "humps" at twilight in the *D*- and lower *E*-region, a feature also predicted for sunrise by theoretical models (*cf*. Ogawa and Shimazaki, 1975). This effect is caused by detachment of electrons from negative ions by sunlight in the visible range. The prediction for 80 km at night is obviously erroneous since it shows values *larger* than at 120 or 160 km. But also the night values for 120 km must be too large since both the International Reference Ionosphere (IRI) and the dedicated *D*- and *E*-region model FIRI (Friedrich and Torkar, 2001) predict densities of the order of only 10^9 m^{-3}. This feature must be attributed to the effective threshold of the Arecibo radar of about $3 \cdot 10^9$ m^{-3}, notably after the manual editing. Finally, in Figure 6 we show the solar activity dependence at spring equinox and median *Kp*, both for day and night. Perhaps the most interesting feature is that the solar activity dependence of the *F*-region peak is larger at night than during the day. This rather puzzling result is supported by a qualitatively very similar characteristic predicted by the global foF_2 model.

4. CONCLUSIONS

The results of two entirely separate ionospheric models, built from different data but using the same modelling technique, are compared, one is a single-station, full ionospheric model, and the other describes the global behaviour of the peak electron density. The agreement of the comparable parameter (NmF_2) is encouraging and differences are explicable by the different inputs. The comparison with IRI, or rather with the foF_2 models within IRI, has already been made against the dedicated global foF_2 model (Oyeyemi *et al*., 2005) and generally showed significantly better agreement with the large set of ionosonde-based data; a similar comparison with the Arecibo model does not deem to be necessary at this stage.

REFERENCES

Fankhauser, M., A Steady-State Ionospheric Model Based on Arecibo Incoherent Scatter Radar Using Neural Networks, MSc thesis, Graz University of Technology, 2005.

Friedrich, M., Torkar, K.M., FIRI: A Semiempirical Model of the Lower Ionosphere, *J. geophys. Res.* **106 (A10)**, pp. 21409-21418, 2001.

Friedrich, M., Egger, G., McKinnell, L.A., Belova, E., Perturbations in EISCAT Electron Density Data Visualised by Normalisation, *Adv. Space Res*., in press, 2006.

Haykin, S., Neural Networks: A Comprehensive Foundation, Macmillan, Indianapolis, 1994.

McKinnell, L.A., Poole, A.W.V, The Development of a Neural Network Based Short Term foF2 Forecast Program, *Phys. Chem. Earth.* **25**, pp. 287-290, 2000.

McKinnell, L.A., Poole, A.W.V, Ionospheric Variability and Electron Density Profile Studies with Neural Networks, *Adv. Space Res.* **27**, pp. 83-90, 2001.

Ogawa, T., Shimazaki, T., Diurnal Variation of Odd Nitrogen and Ionic Densities in the Mesosphere and Lower Thermosphere: Simultaneous Solution of Photochemical-Diffusive Equations, *J. geophys. Res*. **80**, pp. 3945-3960, 1975.

Oyeyemi, E.O., Poole, A.W.V., McKinnell, L.A., On the Global Model for foF_2 Using Neural Networks, *Radio Sci*. **40**, doi:10.1029/2004RD003223, 2005.

Tulunay, E., Özkaptan, C., Tulunay, Y., Temporal and Spatial Forecasting of the foF2 Values up to Twenty Four Hours in Advance, *Phys. Chem. Earth.* **25**, pp. 281-285, 2000.

Tulunay, E., Senalp, E.T., Radicella, S.M., Tulunay, Y., Forecasting Total Electron Content Maps by Neural Network Technique, *Radio Sci.* **41 (4)**, RS4016, 2006.

Wintoft, P., Cander, Lj., Short-Term Prediction of foF2 Using Time Delay Neural Networks, *Phys. Chem. Earth.* **24**, pp. 343-347, 1999.

Figure Caption

Figure 1. RMS error factor of the single-station model for Arecibo as a function of altitude.

Figure 2. Lines of constant electron densities on February 4th, 1989. *Top*: Arecibo incoherent scatter radar measurements, *centre*: corresponding prediction of the local model, *bottom*: ratio between data and corresponding model values. Note the downward propagating enhancement structures apparent in the bottom panel.

Figure 3. Maximum electron density during February 4th, 1989, above Arecibo together with the corresponding predictions of the two models.

Figure 4. Predicted peak electron densities according to the two models for noon and midnight during 1993. *Top*: Arecibo model, *bottom*: foF_2 model.

Figure 5. Diurnal variation of electron densities at various altitudes in spring at medium solar and magnetic activity. At the lowest depicted altitude (80 km) there is an obvious problem caused by the radar's threshold (night values at 80 km being larger than at 120 and 160 km).

Figure 6. Solar activity dependence at spring equinox and medium *Kp*. *Left*: midnight, *right*: noon. Note that at night the variation of the *F*-region peak is larger than during the day.

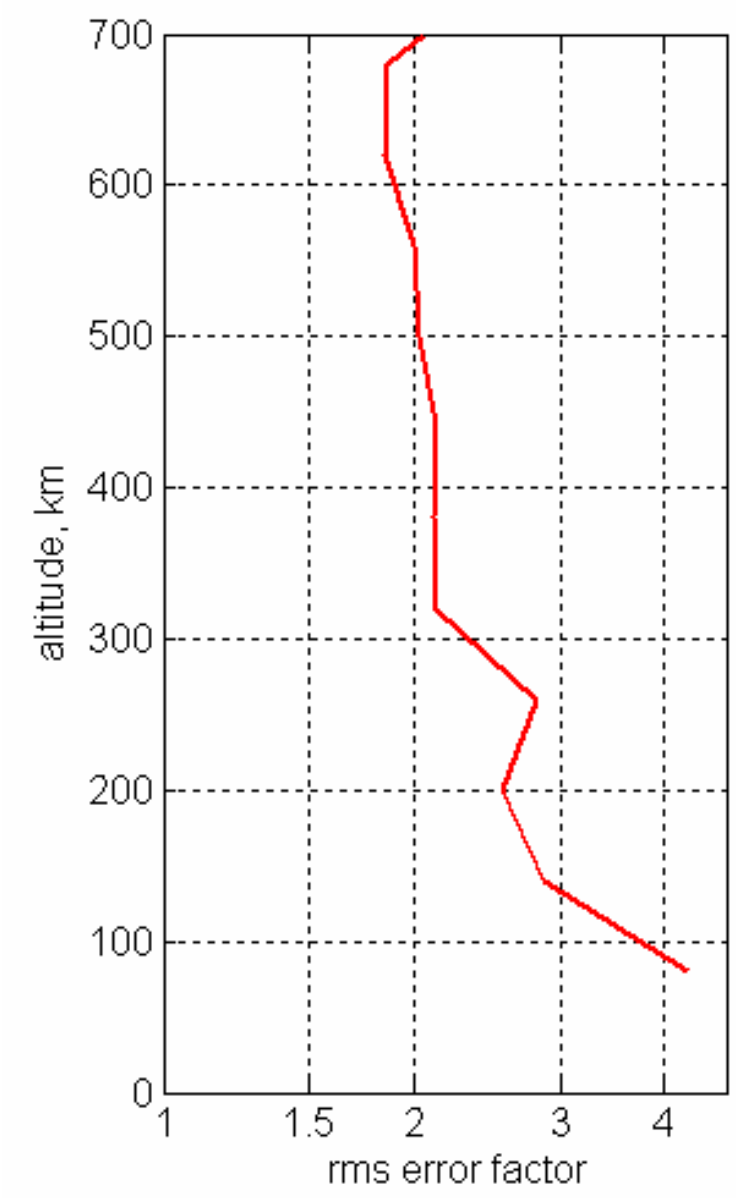

Figure 1. RMS error factor (data/prediction) of the single-station model for Arecibo as a function of altitude.

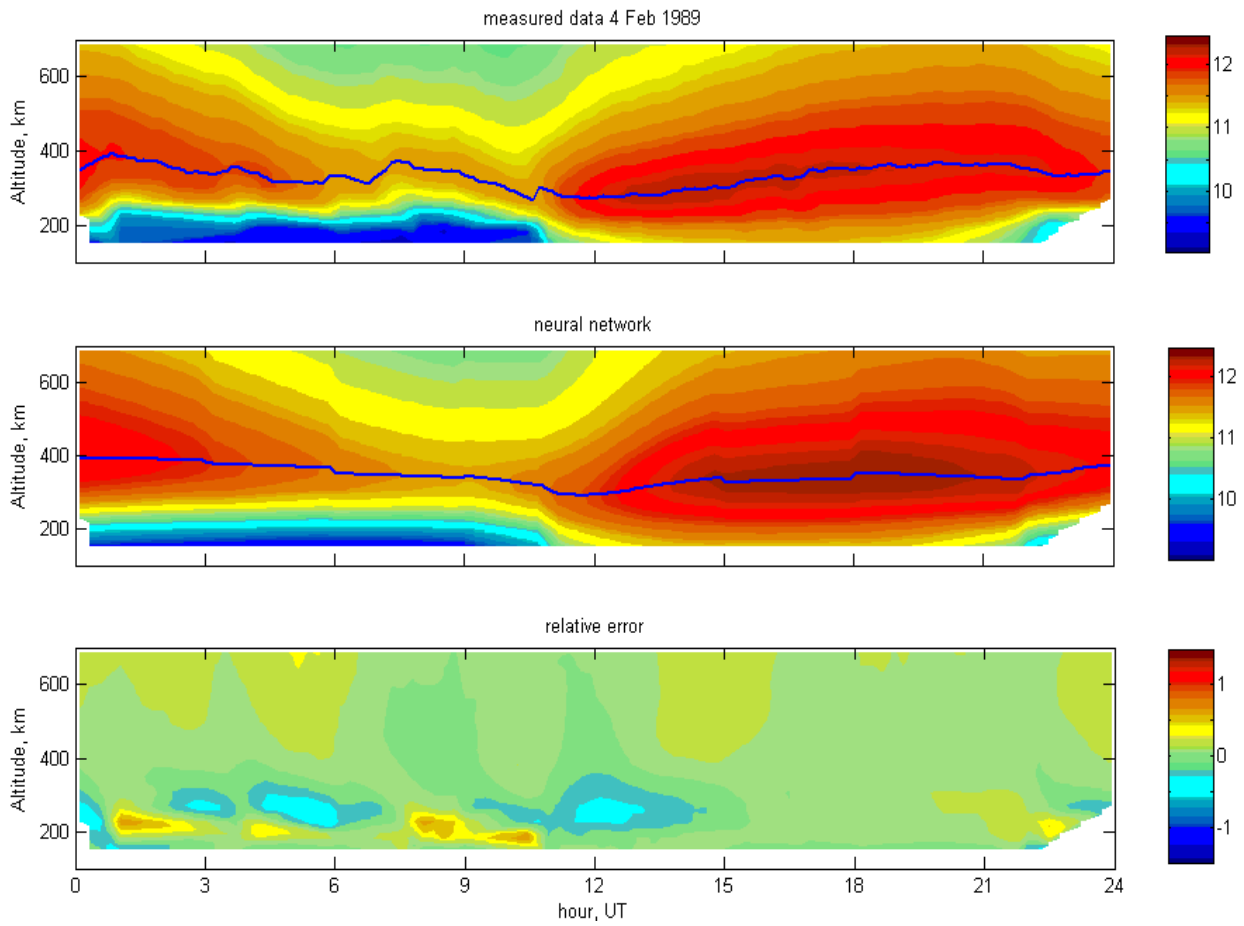

Figure 2. Lines of constant electron densities on February 4th, 1989. *Top panel:* Arecibo incoherent scatter radar measurements, *centre panel:* corresponding prediction of the local model, *bottom*: ratio between data and corresponding model values. Note the downward propagating enhancement structures apparent in the bottom panel.

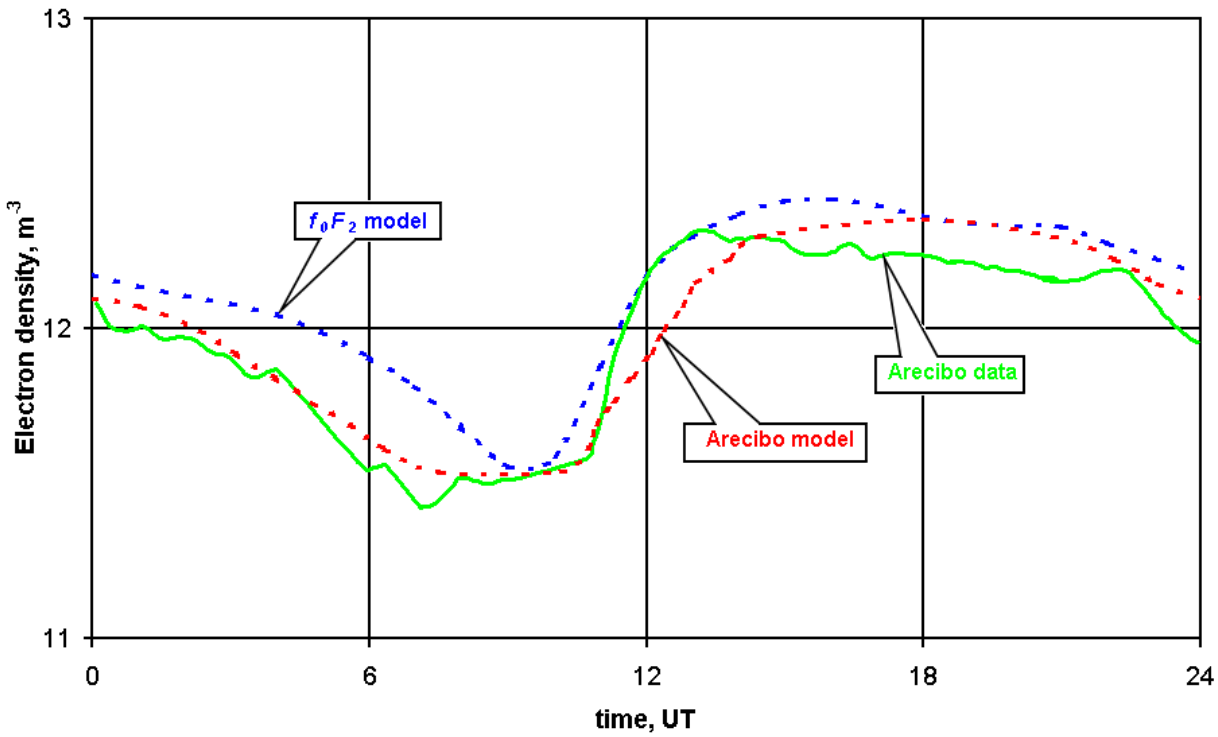

Figure 3. Maximum electron density during February 4th, 1989, above Arecibo together with the corresponding predictions of the two models.

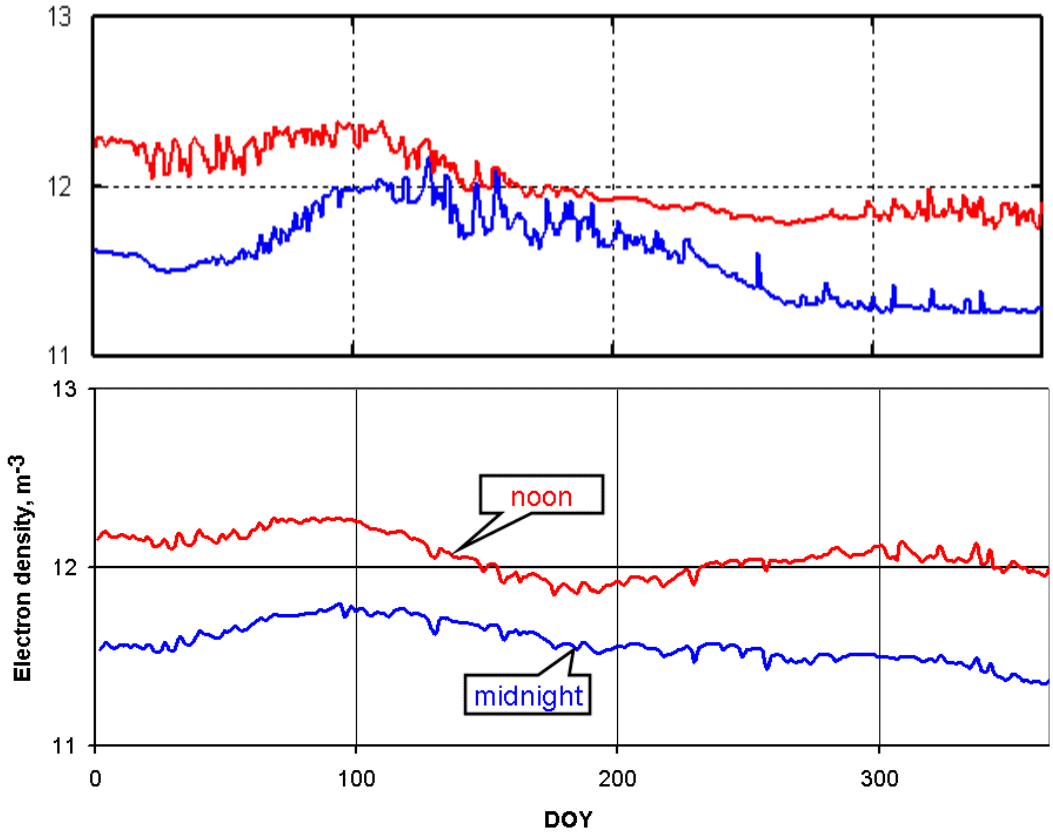

Figure 4. Predicted peak electron densities according to the two models for noon and midnight during 1993. *Top panel*: Arecibo model, *bottom panel: foF*$_2$ model.

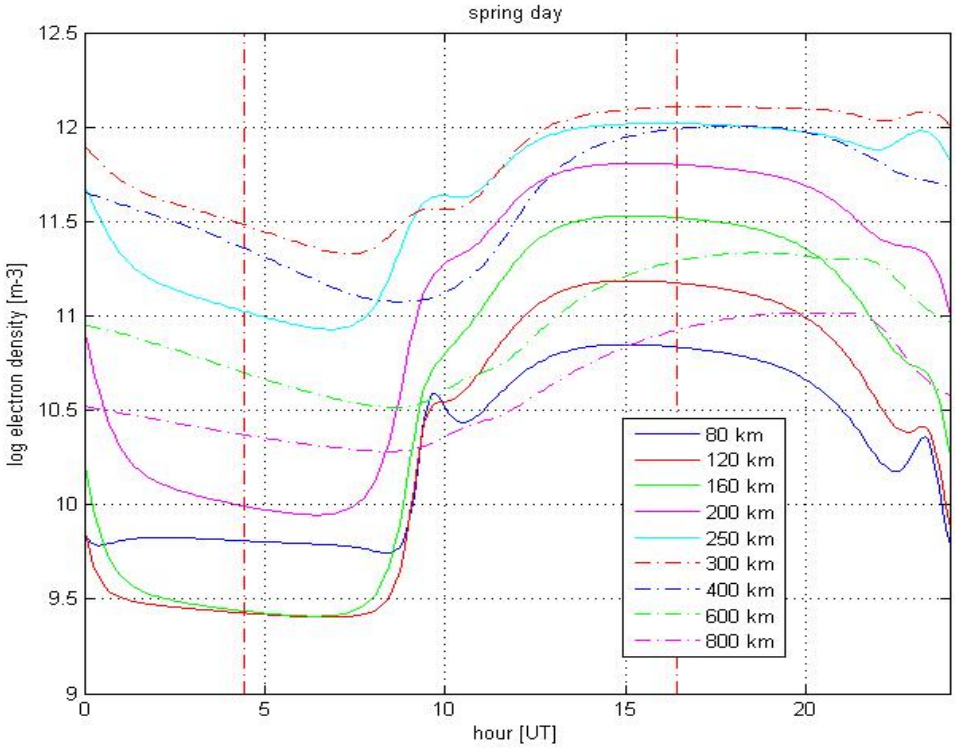

Figure 5. Diurnal variation of electron densities at various altitudes in spring at medium solar and magnetic activity. At the lowest depicted altitude (80 km) there is an obvious problem caused by the instrument's threshold (night values at 80 km being larger than at 120 and 160 km).

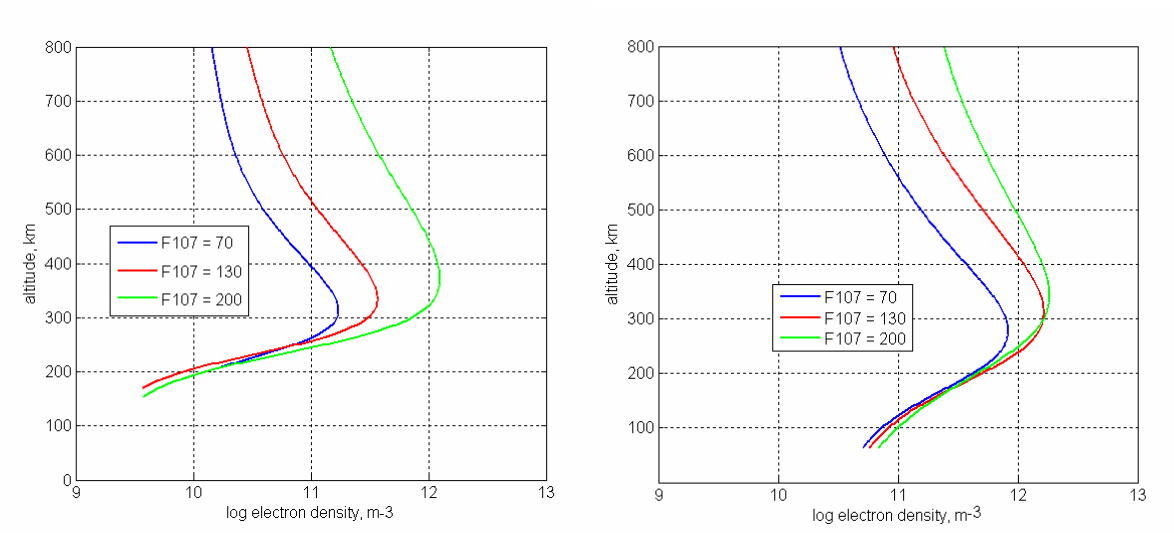

Figure 6. Solar activity dependence at spring equinox and medium Kp. *Left*: midnight, *right*: noon. Note that at night the variation of the *F*-region peak is larger than during the day.