Timedependent Models

of the

Polar Cap and Low Latitude Ionosphere

PhD Thesis

by

Martin Fankhauser

Institute of Communication Networks and Satellite Communications
Graz University of Technology

Head of the Institute: Univ.-Prof. Dipl.-Ing. Dr.techn. Otto Koudelka

Supervisor: Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Martin Friedrich

Graz, August 2010

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

---------- ----------

Abstract

The goal of the thesis was to develop two empirical ionosphere models based on neural networks. One for low latitude based on data from Arecibo (Puerto Rico, 18.34°N 66.75°W, 150 - 700 km) and another model for high latitude based on data from Svalbard (Spitzbergen, 78.09°N 16.02°E, 100 - 700 km, including rocket data from Heiss Island). Artificial neural networks are able to "learn" relations between input parameters, generalize, and ultimately to predict outputs for new input data.
In this case inputs for diurnal time, seasonal time, solar activity, magnetic index and altitude are used as input parameters to predict the dependent output, the electron density. To achieve wide validity it is essential to have a large amount of training data which covers most of the ranges of the input parameters. The models are single station model and strictly speaking only valid for one latitude (18.2°, 78.09° respectively). As an additional topic of research the night decay of the electron density was analysed.

Zusammenfassung

Ziel der vorliegenden Arbeit war es zwei empirische Ionosphärenmodelle zu entwickeln, welche sich auf künstliche neuronale Netze stützen. Ein Modell beschreibt das Verhalten der Ionosphäre in niedrigen Breiten von ca. 150 km bis ca. 700 km (Daten aus Arecibo, Puerto Rico, 18.34°N 66.75°W).
Das zweite Modell beschreibt das Verhalten der Ionosphäre in hohen Breiten (Daten aus Svalbard, Spitsbergen, 78.09°N 16.02°E + Raketen Heiss). Künstliche neuronale Netze sind in der Lage Zusammenhänge von Eingangsparametern zu "lernen" und zu generalisieren. Somit ist es möglich Voraussagen für neue Eingangsparameterkonstellationen abzugeben.
Im konkreten Fall werden Jahreszeit, Tageszeit, Sonnenaktivität, Magnetfeldstörung und geometrische Höhe als Eingabe Parameter verwendet um die davon abhängige Elektronendichte zu beschreiben. Um einen möglichst großen Gültigkeitsbereich gewährleisten zu können, ist es nötig eine große Menge an Trainingsdaten zu benützen.
Da nur Daten von jeweils einem Beobachtungspunkt verwendet wurden, sind diese Modelle strenggenommen nur für den jeweiligen Breitengrad (18.2°, 78.09°) gültig. Als zusätzlicher Forschungsgegenstand wurde die Abnahme der Elektronendichte in der Nacht untersucht.

Special thanks to my supervisor Prof. Martin Friedrich, for supporting me patiently throughout the preparation of this thesis. I want to thank my mother for supporting me all my life and specially when I was studying.

Contents

1 Introduction

Radio wave propagation in the upper atmosphere is strongly influenced by the ionosphere. The ionosphere is the ionized part of the atmosphere and it is of interest to know its behavior to calculate radio wave propagation.
In this thesis the development of two new ionospheric models (single station) is presented. Two widely different stations are selected, Arecibo (Puerto Rico, 18.34°N 66.75°W) and Svalbard (Spitsbergen, 78.09°N 16.02°E) not least because edited raw data for these stations are available.

The attempt is to develop two models based on artificial neural networks which can predict electron densities from 100 km to 700 km. A short introduction of the physical basics and existing models is given in this chapter.

In the Ch. 2 "Neural Networks" a short explaination of artificial neural networks is given to understand the basic idea of this modelling approach. To gain an overview about structure of datasets in Ch. 4 "Data sources" many statistics and distributions are provided. Later in Ch. 3 "Other Models" the most important facts about the neutral atmosphere model MSIS and IRI model are summarized.
Ch. 5 "Input Parameters and Preprocessing" includes all steps of Data Preprocessing and calculation of physically appropriate Input Parameters. As a main Ch. 6 "Model Development" provides all important steps of optimisation of the model design regarding combinations of input parameters, architecture, learning parameters and time dependent inputs.

Finally the Results (Arecibo, Ch. 7 and Svalbard, Ch. 8) of the Models are dicussed, and limitations and comparisons to IRI are shown.

1.1 Atmosphere

At first an overview over different parts of the atmosphere is given. Traditionally meteorology divides the atmosphere into following layers:

- Troposphere (below approx. 12 km altitude)

- Stratosphere (12 to 50 km altitude)
- Mesosphere (50 to 80 km altitude)
- Thermosphere (80 to 500 km altitude)
- Exosphere (above 500 km altitude)

A planetary atmosphere can be described by four properties: pressure, density, temperature and composition. The different layers are traditionally named by the variation of temperature with height. Different regions are called spheres and the upper limit layers are called pauses. Up to 12 km the atmosphere's density is high enough to make weather possible (*troposphere*), in it the temperature decreases with about 10 K km^{-1}. Above it is the *stratosphere*, where the temperature increases with height. At about 50 km there is a maximum of temperature because of the absorption of solar ultra-violet radiation (ozone layer); this region is called the *stratopause*. In the *mesosphere* the temperature decreases again to 180 K (-90° C) or even lower at a height of about 80 to 90 km. This coldest region of the atmosphere is called *mesopause*. In the *thermosphere* solar ultra-violet radiation increases temperature. The temperature in the thermosphere is above 1000 K which makes thermosphere the hottest part of all these different regions which can also be seen in Fig.1.1.

In the field of wave propagation the atmosphere is divided into different ionized layers as can be seen in (Fig.1.1).

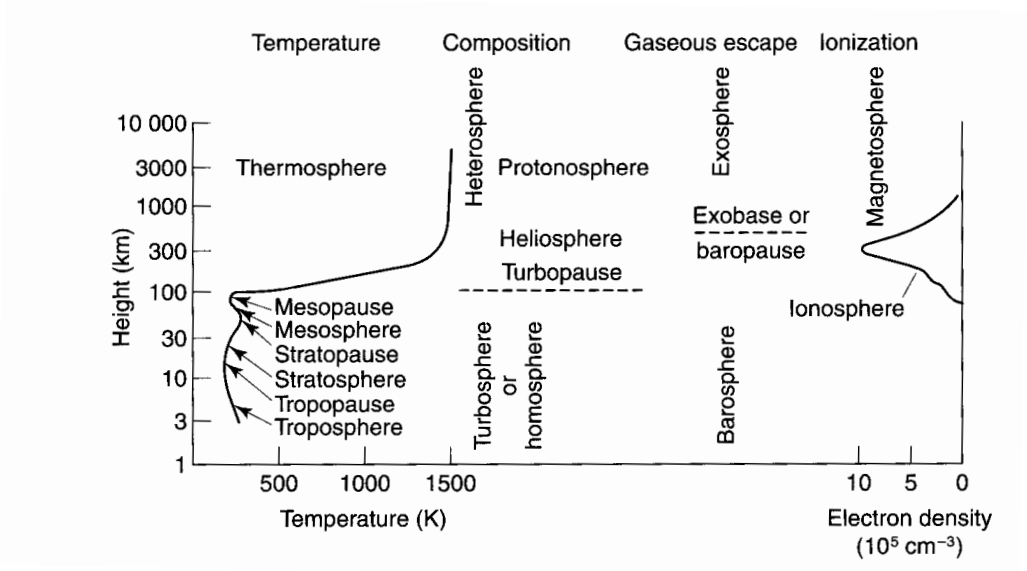


Figure 1.1 Nomenclature of the upper atmosphere based on temperature, composition, mixing and ionisation [Hargreaves, (1992)].

1.2 Ionosphere

The ionized part of the atmosphere is called ionosphere. The name comes from the free electrons and positive ions in this layer. Electrons and ions are generally in equal numbers (quasi-neutrality). The majority of the particles are uncharged, but the charged ones have a crucial influence on the electrical properties in this layer. These charged particles make radio communication over large distances possible by reflecting radio waves.

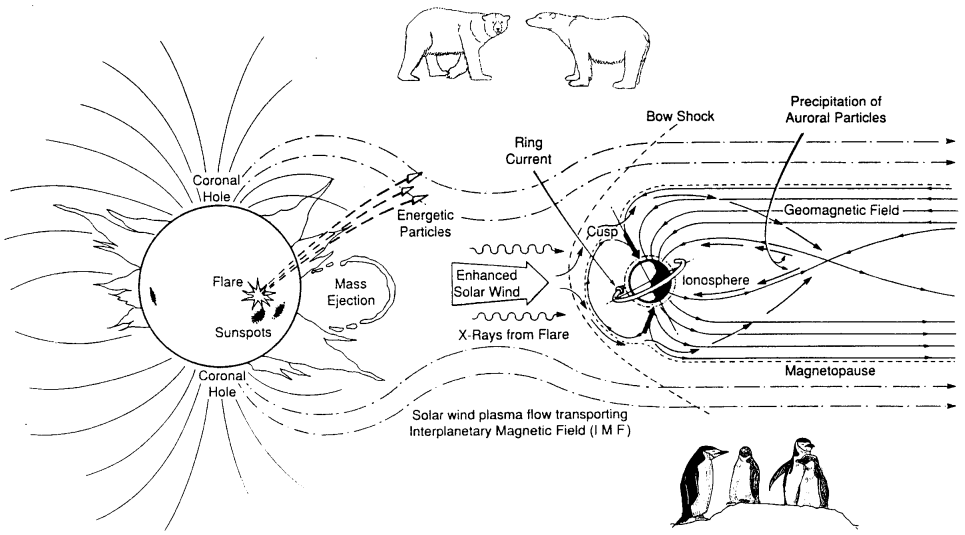


Figure 1.2 The earth in the solar wind [Hargreaves, (1992)].

With the development of radio communication the ionosphere became more and more interesting and the history of ionospheric research began[Hargreaves and Hunsucker, (2003)]. Marconi's transatlantic communication experiment (from Cornwall to Newfoundland) showed that there must be a reflective ionized layer because of the earth's curvature there is no line of sight (Fig 1.3).

The ionosphere acts as a mirror situated between 100 to 400 km above the earth's surface as can be seen in Fig. 1.3. The spatial distribution of charged particles determines the behavior regarding radio waves. This spatial distribution varies due to several influencing parameters. These parameters are primarily:

- day time (zenith angle, Chapman production, ...)
- season (daily integrated insolation, day number, ...)
- Earth's magnetic field (ap, kp, dst, ...)
- solar activity (Rz, $F107$, ...)

- latitude

In this thesis from these determining parameters a model is developed. In Chapter Development 6 all design steps and decisions are described. Examples of electron distributions (electron profiles) are given in Figs. 1.6.

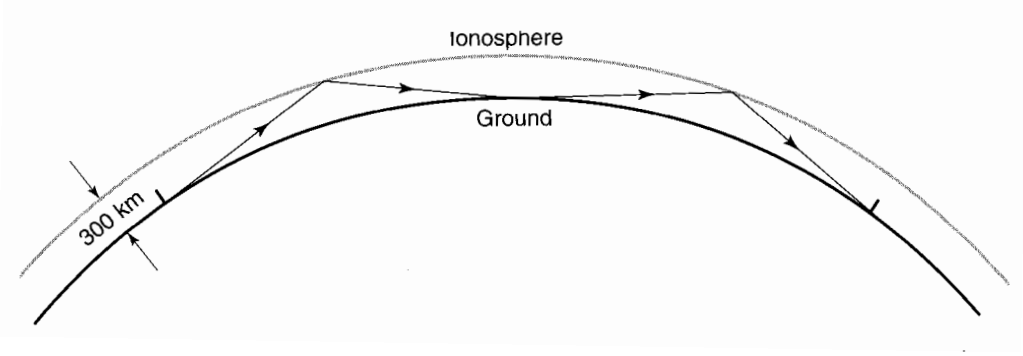


Figure 1.3 Long Range Radio Communication by reflections between ionosphere and ground [Hargreaves and Hunsucker, (2003)].

Radio Communication (terrestrial)

When using high frequencies, the ionosphere reflects the transmitted radio ray. The ray returns to Earth's surface and may again be reflected to the ionosphere.

A radio wave in the ionosphere forces the electrons to oscillate with the same frequency as the incident radio wave. Some of the radio wave energy is lost due to collisions with neutral particle by the oscillation. Total reflection can occur when the plasma frequency of the ionosphere equals the radio frequency.

The limiting frequency at or below a radio wave is reflected at vertical incidence by an ionospheric layer is called critical frequency. If the transmitted frequency is higher than the plasma frequency of the ionosphere, the electrons cannot respond fast enough, and they are not able to reradiate the signal. This critical frequency is calculated in an approximate way as shown below.

The Appleton-Hartree formula as defined by Appleton:

$$n^2 = 1 - \frac{X}{1 - iZ - \frac{\frac{1}{2}Y^2 \sin^2\theta}{1-X-iZ} \pm \frac{1}{1-X-iZ}\left(\frac{1}{4}Y^4 \sin^4\theta + Y^2\cos^2\theta\,(1-X-iZ)^2\right)^{1/2}}$$

Where n is the complex and dual refractive index and v is the electron collision frequency.

$$X = \frac{\omega_0^2}{\omega^2}$$

$$Y = \frac{\omega_H}{\omega}$$

$$Z = \frac{v}{\omega}$$

This simplified relation is derived from the Appleton-Hartree formula [Lassen, (1926)] with the simplification that used frequencies are much higher than the collision frequency v (Z=0) and much higher than the gyro frequency ω_H (Y=0). The imaginary part and the dual property disappear.

$$f_{critical} = 9.10^{-3}\sqrt{N}$$

Where N_e = electron density per cm^{-3} and $f_{critical}$ is in MHz. The Maximum Usable Frequency (MUF) is defined as the upper frequency limit that can be used for transmission between two points at a specified time.

$$f_{muf} = \frac{f_{critical}}{sin(I)}$$

where I = the angle of the wave relative to the horizon.

These relations show that it is necessary to know the electron density N_e at a specific altitude dependent on solar activity, magnetic field, season and daytime. If it is possible to predict electron densities, it is in conseqence possible to calculate wave propagation.

Satellite communication

If we sum up the electron densities along a line we obtain the so-called total electron content (TEC). A plasma reduces the speed of a radio signal traveling through it. The greater its density and/or the lower the radio frequency, the more it is delayed. This signal delay introduces an error in measurements made by positioning systems like GLONASS or Global Positioning Systems (GPS). Further details can be found in [Jakowski et al. (2002)].

1.2.1 Chemical Composition

The upper atmosphere mainly consists of oxygen and nitrogen in molecular or atomic form, or helium and hydrogen at greater heights. Up to 100 km the major species are in almost the same proportion as at ground level.

With increasing altitude molecular oxygen is dissociated to atomic oxygen by ultraviolet radiation betweeen 102.7 and 175.9 nm. The atomic and molecular forms reach equal concentration at about 125 km, and above the atomic form dominates. In Fig. 1.4 densities of all minor and major species of molecular and atomic forms are plotted.

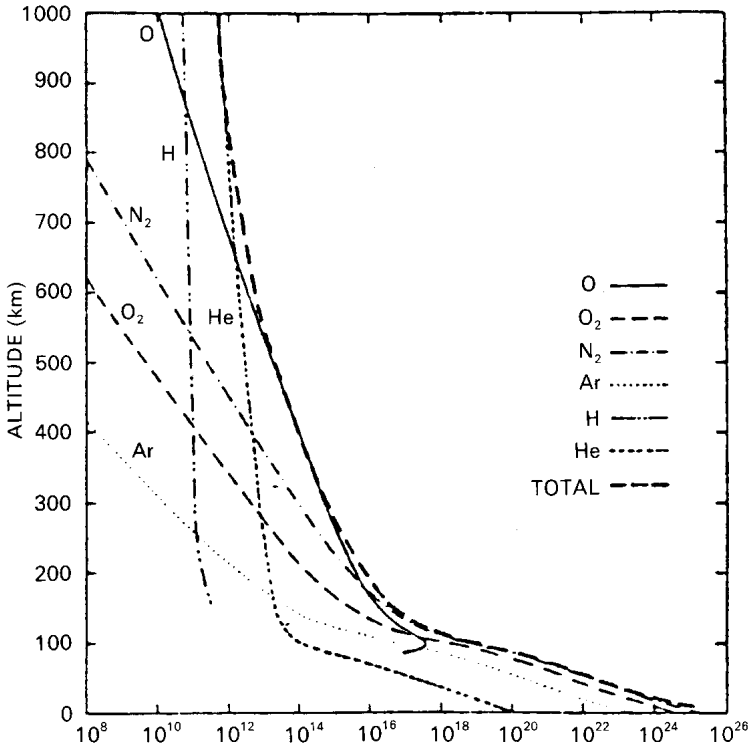


Figure 1.4 Number density (m^{-3}), Atmospheric composition to 1000 km for a typical temperature profile (US Standard Atmosphere, 1976).

1.2.2 Variation of Electron Density

Due to the insolation the ionosphere is changing electron density with day and season. At nighttime electrons and ions recombine and the electron density decreases. If the solar activity is high, more high energy radiation is emitted. This larger radiation causes a larger electron production which can be seen in Fig. 1.5.

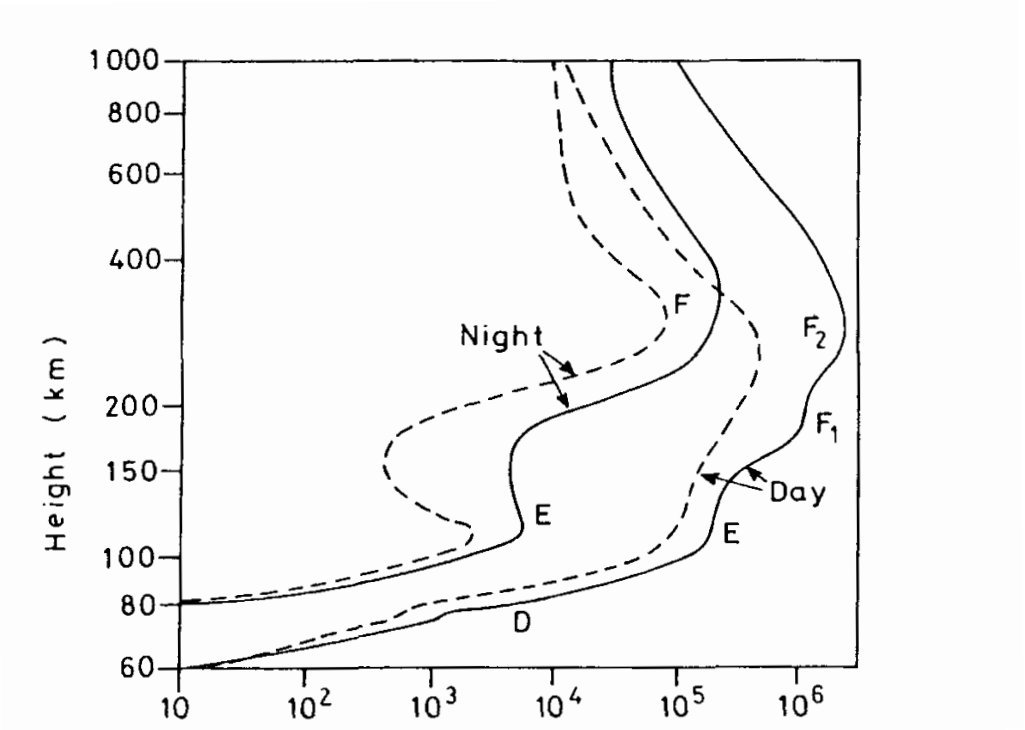


Figure 1.5 Typical vertical profiles of electron density in the mid-latitude ionosphere: - sunspot maximum, – sunspot minimum (After W. Swider, Wallchart Aerospace Environment, US Air Force Geophysics Laboratory).

To give an idea how electron density varies in space and time dimension a plot of a whole day is shown in (Fig. 1.6).

Sunset is around 2100 and sunrise is about 1030 (UTC). When the sun disappears electron densities decrease as can be seen in Fig. 1.6.

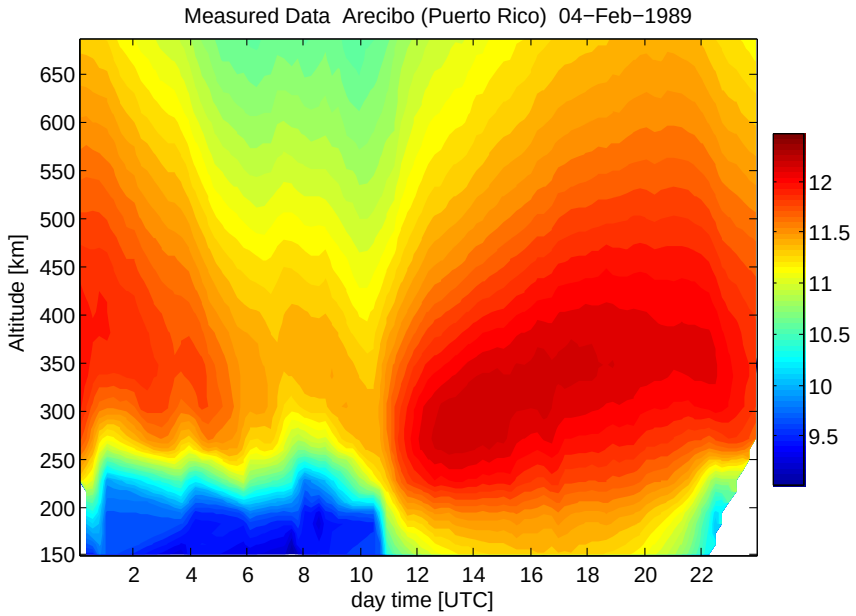


Figure 1.6 Electron density plot over a whole day. Local midnight is at 0430, noon at 1630, densities in $log(m^{-3})$.

1.2.3 Perturbations caused by the Sun

Strong solar flares can occur when the Sun is in the active part of its cycle. These flares cause so-called Sudden Ionospheric Disturbances, SID's. Hard X-Rays (Sudden Ionospheric Disturbances, SID) hit the sunlit side of Earth. High frequency (3 - 30 MHz) radio blackouts caused by released electrons due to penetration of X-Rays to D-region. Very Low Frequency (3 - 30 kHz) signals are reflected during this time by the D rather than by the E layer. Signal loss through the D layer is reduced. When the X-ray emissions end, signal strengths return to normal due to rapid recombination of electrons in D layer.

Polar Cap Absorption (PCA). 15 minutes to 2 hours later high-energy protons reach the earth. The protons spiral around and down the magnetic field lines of the Earth and penetrate into the atmosphere near the magnetic pole. The ionisation in D and E layer is increased. PCA's typically last from one hour to several days, with an average of around

24 to 36 hours [USGS, (2005)].

2 Neural networks

This chapter explains the fundamental idea of modelling physical relations with artificial neural networks. Artificial neural networks are mathematical structures based on the functionality of the human brain.

2.1 Theory of neural networks

Technically speaking artificial neural networks are parallel computing units which consist of connected simple processors called neurons. Each simple processor only knows the signals provided by its inputs. If many of these simple processors work together in a coordinated network they can solve complex problems like pattern recognition or empirical modelling.

The theoretical principles were developed in the 1920's but not until the late 1980's artificial neural networks gained relevance in practice. Such networks can be implemented as hardware units or as software simulations. Due to very ready availibility of high performance CPU's the variant of software simulations is used more often as here in this thesis.

The main difference between neural networks running on simulators and conventional software is that here the solutions of problems are "learned" from real data and not programmed like in hard-coded software. If neural networks are used their solution is better or it is not possible to solve the problem with hard coded software.

There are many applications where artificial neural networks are used such as recognition of credit card fraud, stock predictions, credit evaluation, machine maintainance monitoring, optical pattern recognition (to collect highway fees) and many more.

There are many different architectures and training algorithms but we will restrict further explainations to Feed Forward Networks used for empirical modelling.

2.2 Architecture

A simple processing unit called neuron has several inputs and one output. There are two rules, one how to combine the inputs and one how to calculate the output value out of this combination. A simple network is displayed in Fig. 2.1. One blue square represents one neuron. All outputs of the neurons of one layer are linked to every input of the next layer. Every neuron in the first layer (input layer, networks input) has just one input. There a set of input data (a vector) is fed to the network. The output of the one neuron in the last layer is at the same time output of the neural network.
A full data vector consists of input and output data, in our case the input data contains information about time of the day, season, solar activity, geomagnetic index and altitude. The output data of the neural network is just the electron density N_e.

In a training cycle data (n inputs and 1 output) is presented to the neural network datapoint by datapoint. The difference between the output of the neural network and the measured electron density is squared and summed over all datapoints of the training dataset and is called summed squared error (SSE).

The connections between the neurons of the networks do not transmit a signal one to one, they act as valves and are called weights. During the training process these weights are optimized according to a special training algorithm called Backpropagation to achieve a minimum summed squared error.

The main problem is to find a network architecture that is suitable for the particular task. We want a model that can predict electron densities. Therefore it is important that the network is detailed enough to "learn" all relevant physical relations. On the other hand we want to achieve a good generalisation, meaning that the predictions should be based on physical relations and not too much on singular occurences is training data.

To achieve a good generalisation the network has to be simple enough not to be able to "learn" singular occurences. Further it is necessary to monitor the training process to avoid overfitting. Therefore two pattern sets are used, one for training and one for validation. In the training process the weights are adjusted to minimize the SSE. In the validation process the SSE for the validation patternset is determined. This is done after every training cycle. At the end the network with the lowest SSE_{val} is used for the model. Further information is given in Chapter Development 6.1.

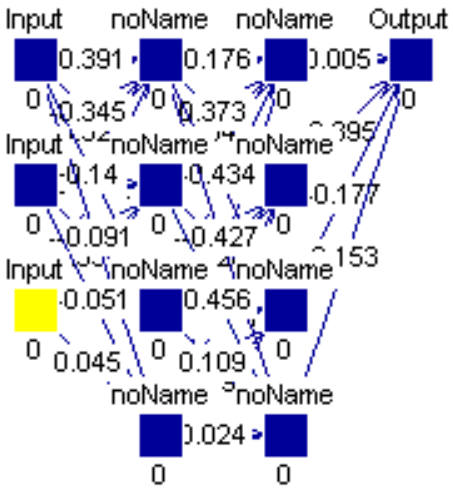


Figure 2.1 Architecture of a simple neural network with 3 input neurons and 2 hidden layers with 4 neurons each (Screenshot from SNNS Creator GUI).

How a training process can look is shown in Fig. 2.2. The upper trace shows the SSE for the training pattern set, the lower curve the SSE for the validation pattern set. Each trace consists of 5 lines because a model consists of 5 neural networks, each with different random starting weights to avoid influences caused by starting weights. As model output a mean value of all five neural network outputs is calculated.

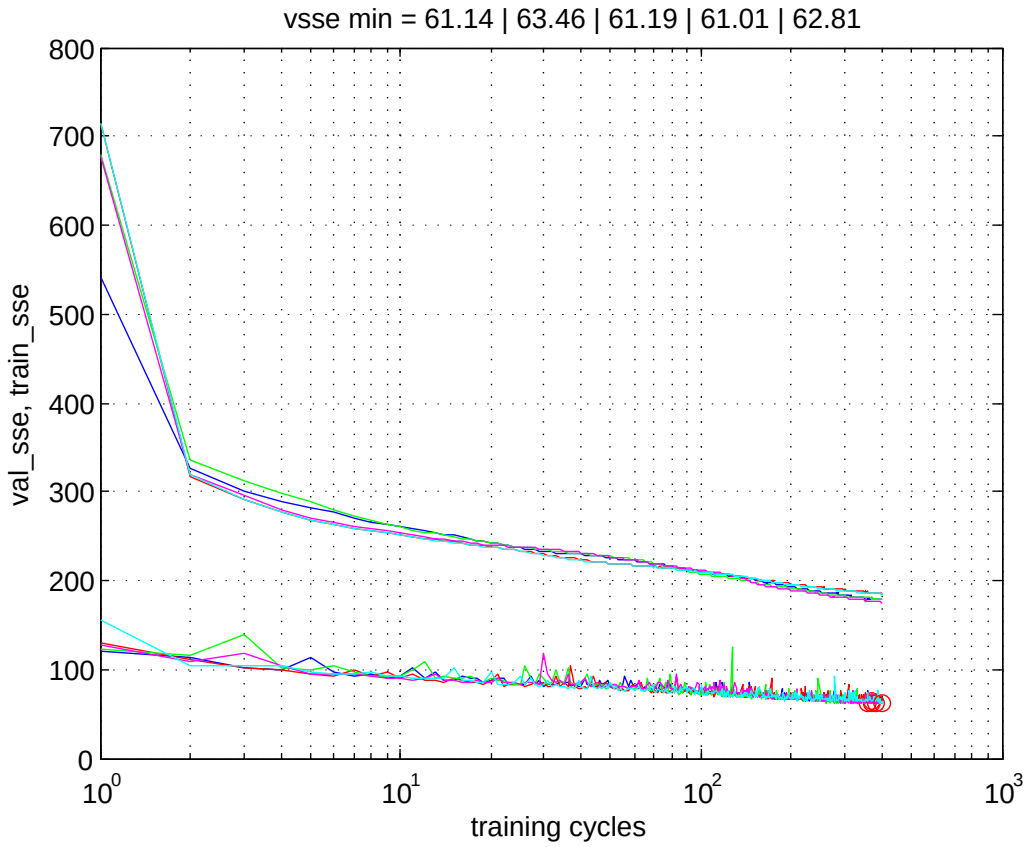


Figure 2.2 Characteristics of SSE training (upper trace) and validation (lower trace), 5 lines each for 5 neural networks, 5 networks with lowest SSE_{val} are saved (red circles) and, model output is the mean of the 5 network outputs.

For detailed information about training parameters and architecture used for the ionospheric models see Figs. 6.1.1 and 6.1.2.

3 Used Models

In this chapter other models used in this thesis are shortly described.

For comparison purposes the IRI model is used too. With the IRI model it is possible to get electron densities for every data point of the Arecibo data. More information about IRI can be found in Ch. 3.1.

The original data from Arecibo and Svalbard consist of time [UTC], altitude [km] and electron density. To have the air pressure as well, the MSIS model is used to calculate pressure from altitude, time and geographic coordinates. More information about MSIS can be found in Ch. 3.2.

3.1 IRI

The IRI[1] model is the result of a cooperation of URSI[2] and COSPAR[3] (Working Group 4). It is the most extensive attempt to develop an empirical ionosphere model. IRI is defined for non-auroral latitudes, at day time up from 65 km and in night time up from 80 km.

In the IRI model the following input parameters are used: latitude, longitude, zenith angle, solar activity, season time and geometric height. Beyond electron density the model yields ion composition (O^+, H^+, He^+, NO^+, $O_2{}^+$) and ion-/electron temperature. In this thesis only electron density is used for comparison purposes.

The data on which the IRI model is based consists of thousands of electron profiles (not manually cleaned data). In F-layer data come mainly from satellite measurements. Data from D-layer originate mainly from rockets. In E- and F1-layer data stem from incoherent scatter radars and ionosondes.

[1] International Reference Ionosphere
[2] Union Radio Scientifique Internationale
[3] Committee on Space Research

The IRI model attampts to approximate the electron density profiles by an analytic function, which leads to dependent height layers that causes poor results in lower altitudes. An example of data produced by the IRI model (IRI 2007) can be seen in Fig. 3.1.

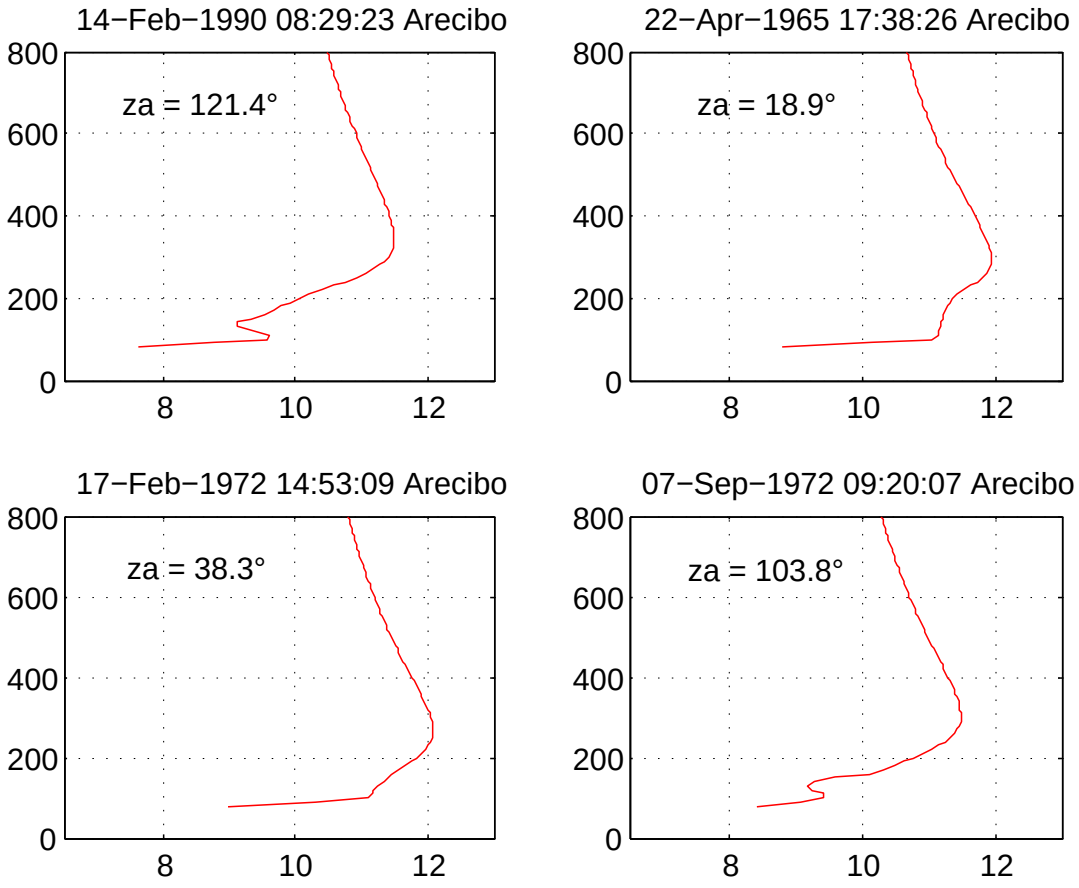


Figure 3.1 IRI output for the coordinates of Arecibo, 4 electron density profiles, valley at night time. Local midnight at 0430, noon at 1630.

The plots in Fig. 3.1 show electron maxima at 10^{12} per m^3 at an altitude of around 270 km. There is also a valley at night time at 170 km. The IRI model in this thesis is only used to test whether the neural networks are able to approximate clean data with no distortions but with similar structure as the measured real life data. The IRI model is also used to evaluate the developed models based on neural network regarding their availability of predicting electron densities. Further information about IRI 2007 can be found at [Bilitza, (2009)].

3.2 MSIS

To test the relevance of altitude and air pressure for the ionosphere, a model which provides air pressure is needed. MSIS is actually several models which describe the neutral atmosphere in terms of temperature, densities and composition (N_2, O, O_2, He, Ar and H) from the ground to the thermosphere. The following inputs are needed:

- geographic coordinates
- day of the year
- altitude
- solar activity
- geomagnetic index kp

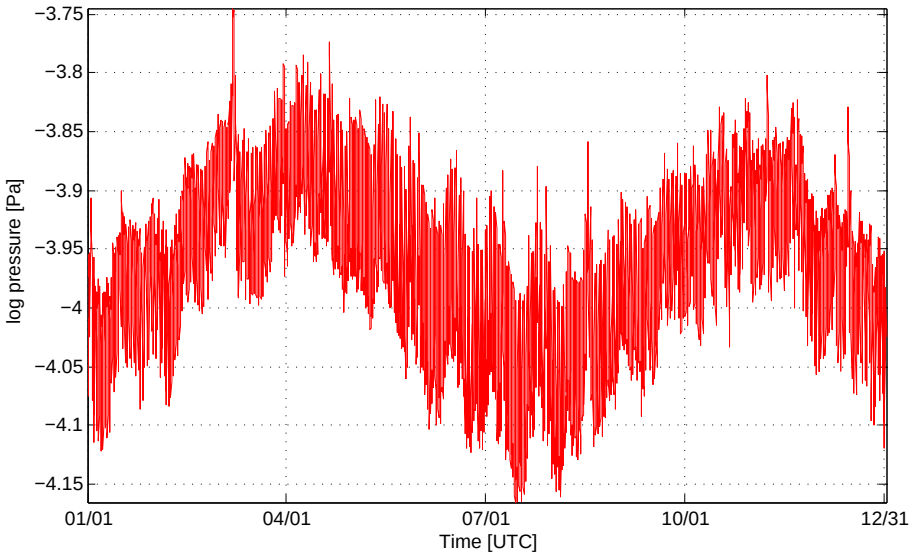


Figure 3.2 Variation of the pressure during the year 1970 at 200 km over Arecibo according the the empirical model MSIS. Locations inside the tropic of the Cancer experience two insolation maxima (c.f. Fig. 5.6).

The data on which the models are based includes measurements from the ground, from rockets and from satellites. Measurement instruments were incoherent scatter radars, mass spectrometers, solar ultraviolet occultation, pressure gauges, falling spheres and grenade detonations. The MSIS models were originally developed by A.E. Hedin at NASA Goddard Space flight Center. For detailed information of several MSIS models see [NASA Goddard].

For this thesis the version NRLMSISE-00 Model 2001 was used. To see an example of MSIS density outputs pressure is plotted for one year (Arecibo, altitude 200 km) in Fig. 3.2.

Due to large variations between day and night, a second (zoomed) plot is displayed in Fig. 3.3.

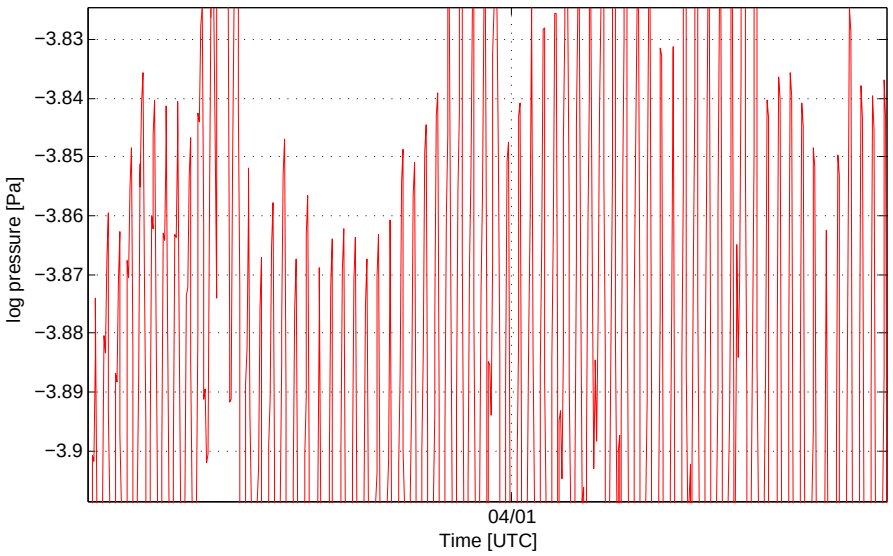


Figure 3.3 MSIS pressure changes in the year 1970, zoomed range to display diurnal changes (Arecibo, 200 km).

4 Datasources

Any empirical model hinges on a good and extensive data base. The approach to develop a model by using artificial neural networks requires good data coverage for all the model space. Because due to their optimization algorithms based on least square methods, the network weights are adjusted to fit as good as possible to the training data. Further the knowledge of data statistics is important to give limitations and validity ranges for the models, because artificial neural networks yield unreasonable results for extrapolation. In this chapter the measured data is analysed and presented to give an overview how well all ranges of the input parameters are covered. Note that always the logarithmic electron density is used.

4.1 Arecibo Electron Density Profiles

The first amount of data comes from Arecibo (about 18° N) and was measured from 1966/7/4 to 2002/4/16 and consists of 1,585,824 datapoints in 98,554 electron density profiles (after editing). The raw dataset was manually edited by Prof. Friedrich to eliminate obviously unrealistic datapoints.

4.1.1 Arecibo Observatory

The electron density data was measured by the NAIC[1] in Arecibo. Arecibo is a town in Puerto Rico, the observatory is in the interior at the (18° 20' 36" N, 66° 45' 11" W).

The idea to build the observatory came from professor William E. Gordon, who was interested in the study of the ionosphere. Gordon's research during the fifties led him to the idea of radar back scatter measurements to gain more information about the ionosphere.

[1] National Astronomy and Ionosphere Center

Figure 4.1 The Arecibo observatory (http://www.naic.edu/)

Begiinning in summer 1960 a huge reflector (about 300 m diameter) was built into a natural dell. Three years later the obsevatory started its measurements under the direction of Gordon.

The surface of the reflector consists of 40,000 perforated aluminium panels, each measuring about 3 by 6 feet (about 1 by 2 meters). To bear the weight of the panels steel cables are strung across the underlying Karst sinkhole. It is a spherical (not parabolic!) reflector.

Suspended 450 feet (about 150 meters) above the reflector is the 900 ton platform. Fixed on this platform the bow-shaped azimuth arm containing the feed is mounted.

The dish in Arecibo is the world's largest, but for incoherent scatter radar purpose only part of it can be illuminated by the feed. Depending on the feed position another part of the dish acts as reflector. Thus the Arecibo incoherent scatter radar does not have a significant lower treshold than for example the EISCAT[2] installation in Tromsø,

[2]European Incoherent Scatter

or ALTAIR[3] in the Kwajalein Atoll. The Arecibo system can operate at frequencies ranging from 50 MHz (6 m wavelength) to 10 GHz (3 cm wavelength). In 1974 a new high precision surface for the reflector was installed together with a high frequency planetary radar transmitter. Further information can be found at [NAIC 2010]. The 1962 built transmitter has a fixed center frequency of 430 MHz and a maximum total peak pulse output power of 2.5 MW.

4.1.2 Incoherent Scatter Radar ISR

High energy ultraviolet radiation from the Sun separates electrons from some of the atoms and molecules in the atmosphere, and these free electrons can scatter radio waves.

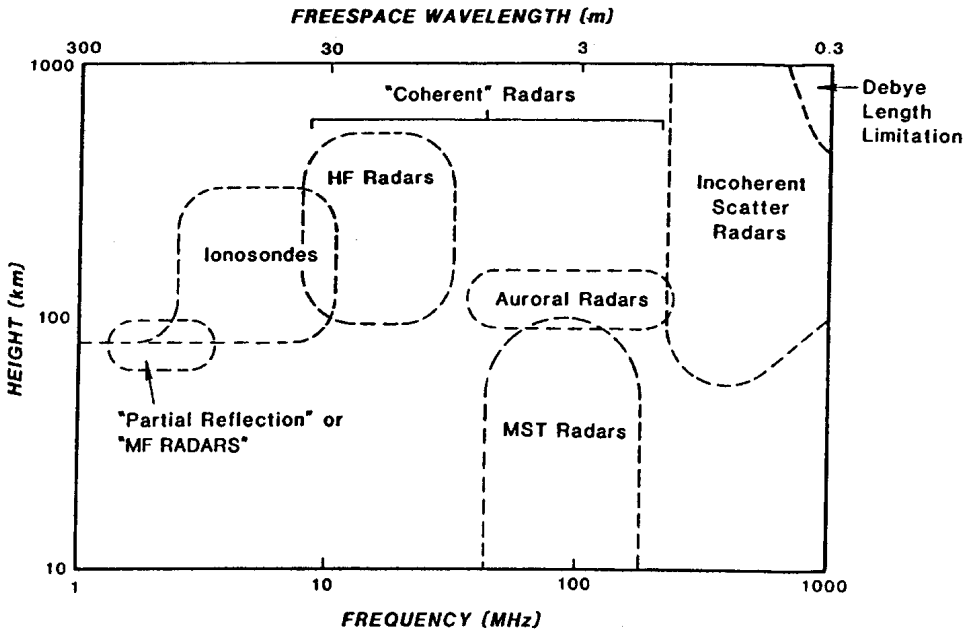


Figure 4.2 Height-frequency diagram to compare various ionospheric radar probes [Schlegel, (1984)]

Characteristics of the ionosphere are measured by the Arecibo 430 MHz radar using incoherent scatter. The transmitted waves reach the ionosphere and each of the millions of electrons in the ionosphere illuminated by the beam act as a target. The radar echo of all the little targets is reflected by the dish to the receiver. The received power correlates with the amount of free electrons.

[3]ARPA Long-Range Tracking and Instrumentation Radar

The ions and electrons in the ionosphere are split mainly by ultraviolet radiation from the sun. The light electrons respond to the radar waves and scatter it. The heavy ions are too slow, but they hold electrons nearby, that is why overall motions of electrons also show motion of ions.

Through signal processing with Fourier transformations and autocorrelations it is possible to calculate from the radar echo the properties of the ionospheric plasma [NAIC 2010].

A diagram of working point and application range of incoherent scatter radars compared to other radar probes is shown in Fig.4.2.

4.1.3 Arecibo Statistics and Distributions

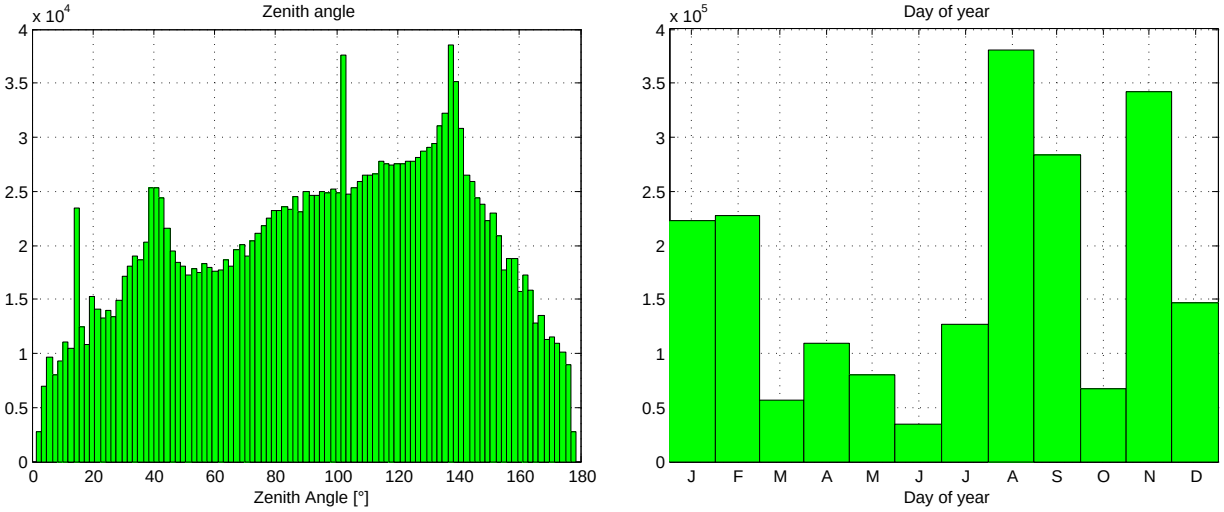
Table 4.1 shows the mean value, the 1^{st} sextile, the 1^{st} quartile, the 3^{rd} quartile and the 5^{th} sextile for all relevant parameters of the Arecibo data.

Variable	**Unit**	**1.Sextile**	**1.Quartile**	**Mean**	**3.Quartile**	**5.Sextile**
zenith angle	°	42.6	58.9	95.5	133.0	142.2
10.7 cm flux		74.9	80.1	129.4	176.2	193.2
dst		-40.0	-33.0	-21.1	-4.0	0.0
kp		1.3	1.7	2.7	3.7	4.0
ap		5.0	6.0	17.2	22.0	27
alt	km	138.0	140.0	324.1	447.1	515.8
$logN_e$	m^{-3}	9.89	10.56	10.95	11.63	11.82

Table 4.1 Most important input parameters are analysed for range coverage (Arecibo).

To develop a reliable model it is important to achieve a good coverage of input parameters ranges. On the following pages distributions for all relevant parameters are shown.

Distribution of zenith angle and day number.

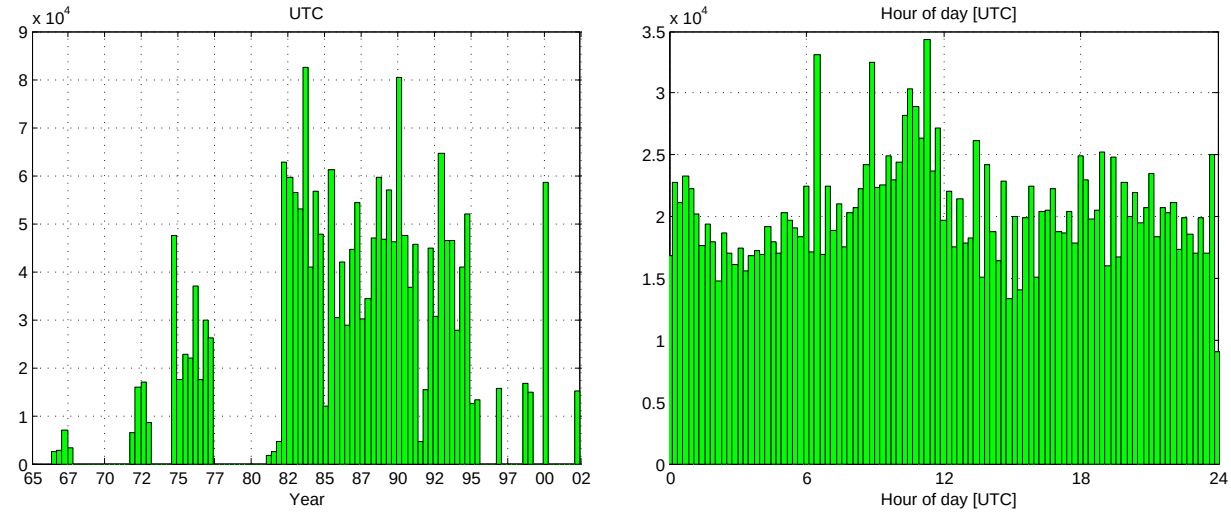


(a) Zenith angle, mean = 95.5 1. Q = 58.9 3. Q = 133.0 1. S = 42.6 5. S = 142.2

(b) Day of year

Figure 4.3 Arecibo, distribution of zenith angle and day number

Distribution of UT and hour number.

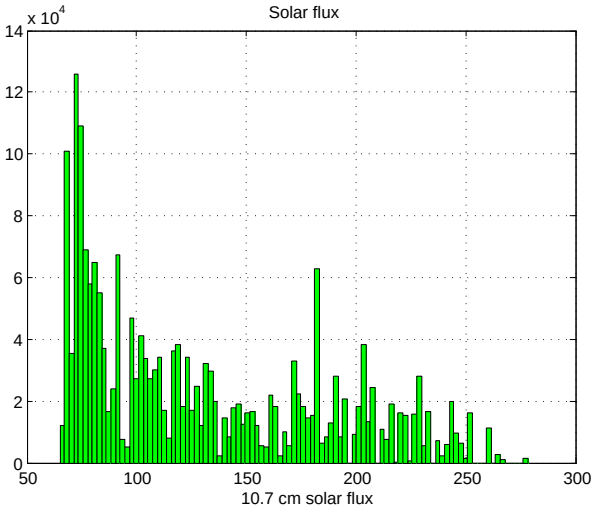


(a) UTC

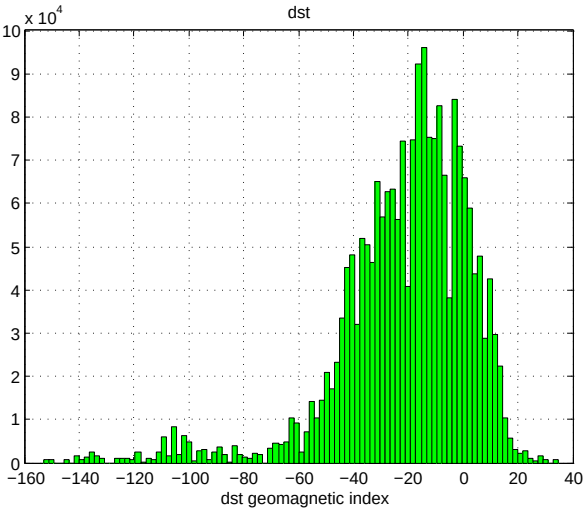
(b) Hour of day UTC

Figure 4.4 Arecibo, distribution of UT and hour number

Distribution of F107 flux and dst.



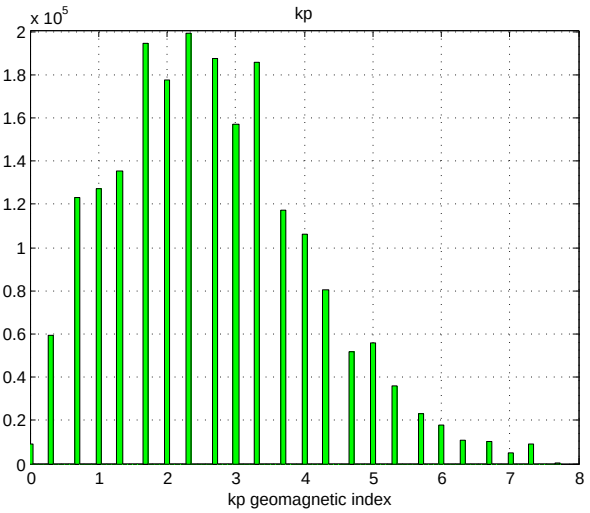
(a) Solar flux, mean = 129.4 1. Q = 80.1 3. Q = 176.2 1. S = 74.9 5. S = 193.2



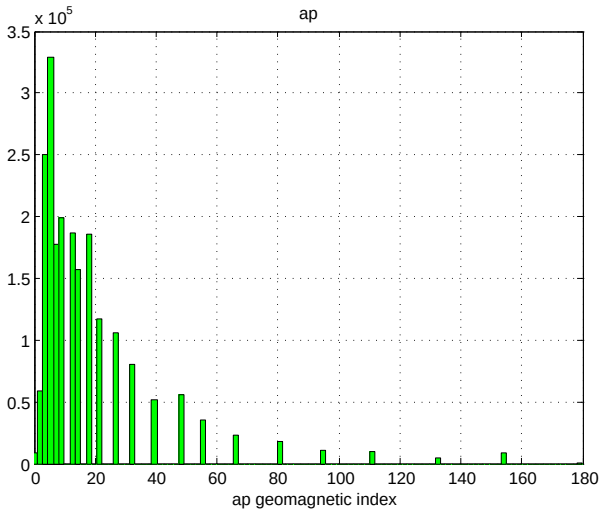
(b) dst, mean = -21.1 1. Q = -33.0 3. Q = -4.0 1. S = -40.0 5. S = 0.0

Figure 4.5 Arecibo, distribution of F107 flux and dst

Distribution of kp and ap.



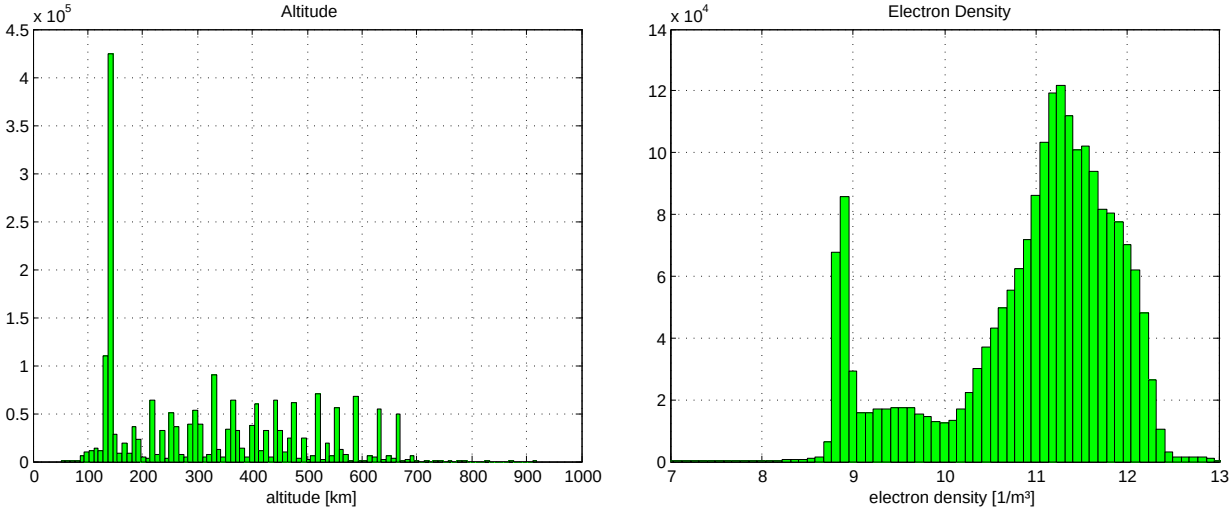
(a) kp, mean = 2.7 1. Q = 1.7 3. Q = 3.7 1. S = 1.3 5. S = 4.0



(b) ap, mean = 17.2 1. Q = 6.0 3. Q = 22.0 1. S = 5.0 5. S = 27.0

Figure 4.6 Arecibo, distribution of kp and ap

Distribution of altitude and electron density.



(a) Altitude, mean = 324.1 1. Q = 140.0 3. Q = 447.1 1. S = 138.0 5. S = 515.8

(b) Log electron density, mean = 10.95 1. Q = 10.56 3. Q = 11.63 1. S = 9.89 5. S = 11.82

Figure 4.7 Arecibo, distribution of altitude and electron density

4.1.4 Minimum Altitude

For the interpretation of the model results in Chapter 7 it is important to know the distribution of the minimum altitudes of Arecibo electron density profiles.
In Figure 4.8 all lowest datapoints of all profiles are displayed in a scatter plot. Because of the much lower electron densities at night the electron density profiles after sundown begin at significant higher altitudes.

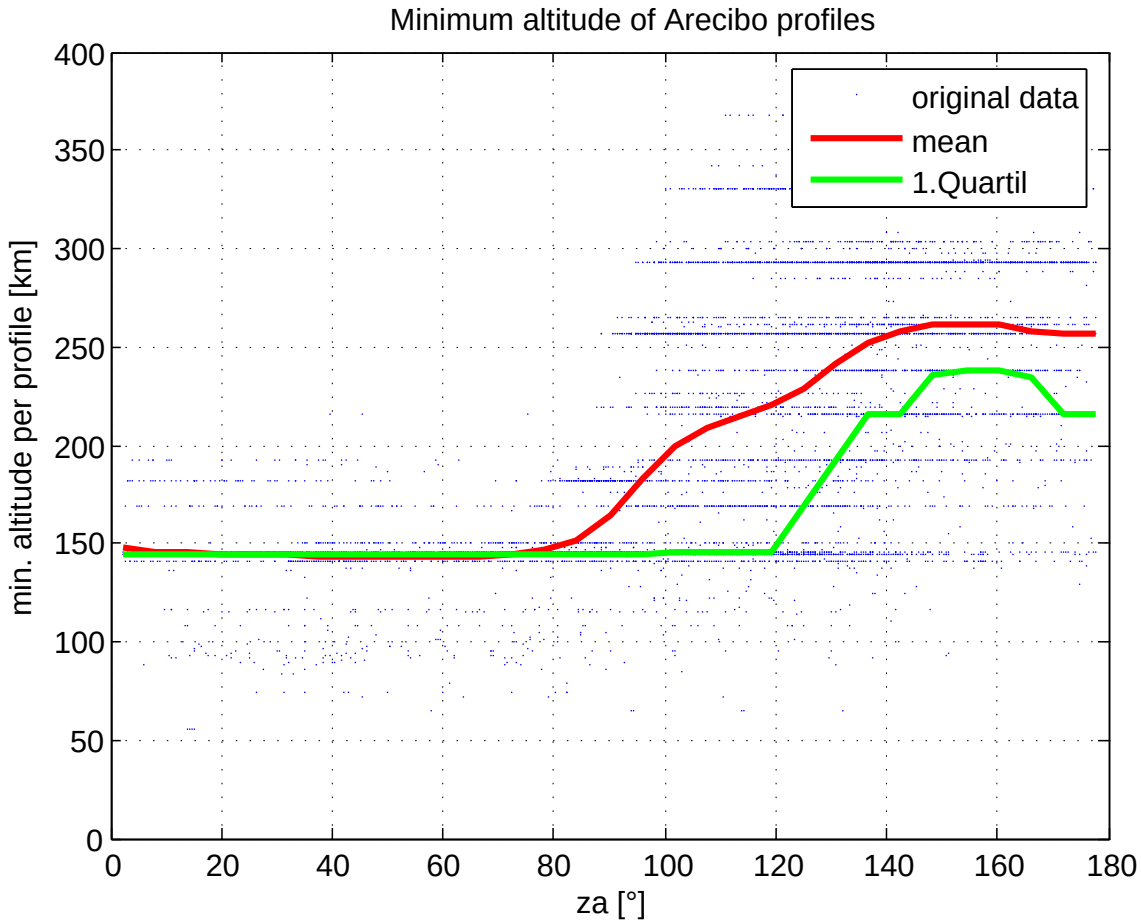


Figure 4.8 Distribution of minimum altitudes of the electron density measurements as a function of solar zenith angle. Note the considerable higher starting heights at the measurements compared to the Svalbard data. Due to low electron densities measurements are started at higher altitudes at night.

4.2 Svalbard Electron Density Profiles

4.2.1 Svalbard observatory

The second source of data was measured at the Kjell Henriksen Observatory (KHO) in Svalbard (Longyearbyen) (Fig. 4.9) (78° 09' 11" N, 16° 01' 44" E, 78.15305° N, 16.02889° E) at 445 m altitude [EISCAT]. The data are measured with an incoherent scatter radar (ISR) as is described in Ch. 4.1.2. The feed in this installation is a Cassegrain type. The transmitter has a fixed center frequency at 500 MHz and a peak power of 1 MW.

The ISR data used here start in 1996 and end in 2004. The data from the incoherent scatter radar is supplemented by data (7049 datapoints) from 371 rocket flights starting on 28. 2. 1979 and ending at 2. 3. 1994. These rockets were launched at Heiss Island (80.62° N, 58.05° E), slightly further north than EISCAT Svalbard, but at almost the same geomagnetic latitude. Further details can be found in [Danilov et al. (2003)].

Figure 4.9 The archipelago of Svalbard in the north of Norway.

One profile from a rocket flight on 6. 7. 2003 launched near Svalbard (coordinates of measurement: 78.91° N 11.93° E, conducted by Prof. Friedrich) was added (39 datapoints) as well(4.11).

In the total number of data points used for the Polar Cap Model consists of 1,459,262 datapoints in 162,507 electron density profiles after editing. The raw dataset was manually edited by Prof. Friedrich to exclude obviously unrealistic datapoints.

The UT timestamps of the rocket flights are corrected to coordinates of Svalbard EISCAT radar, to treat them like data from radar measurement. Further details to EISCAT can be found in [Schlegel, (1995)].

To have an idea of values of the rocket measurements all electron density profiles are plotted in Fig. 4.10. Data is available down to a minimum altitude of 50 km and an electron density of approximately 6.8 10^6 m^{-3}.

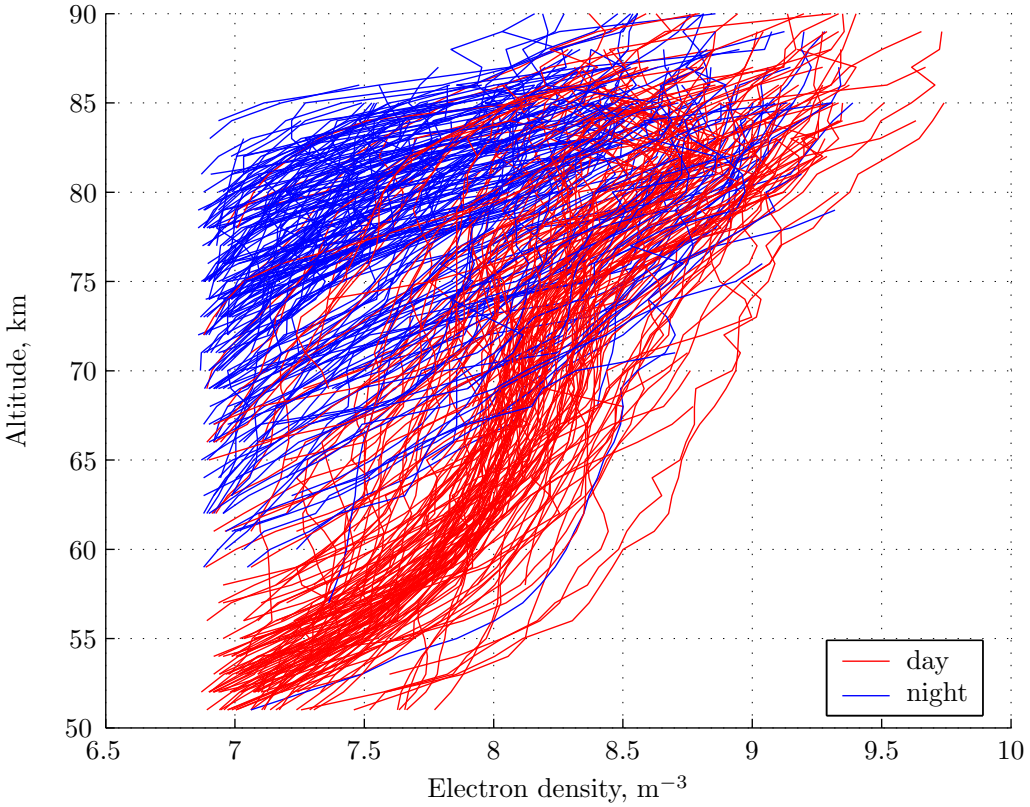


Figure 4.10 All data from rocket flights at Heiss Island (Egger, 2004). The threshold electron density is slightly below 10^7, the lower altitude limit is determined by the probe deployment and the upper limit coincides with the apogee of the rocket (M100-B).

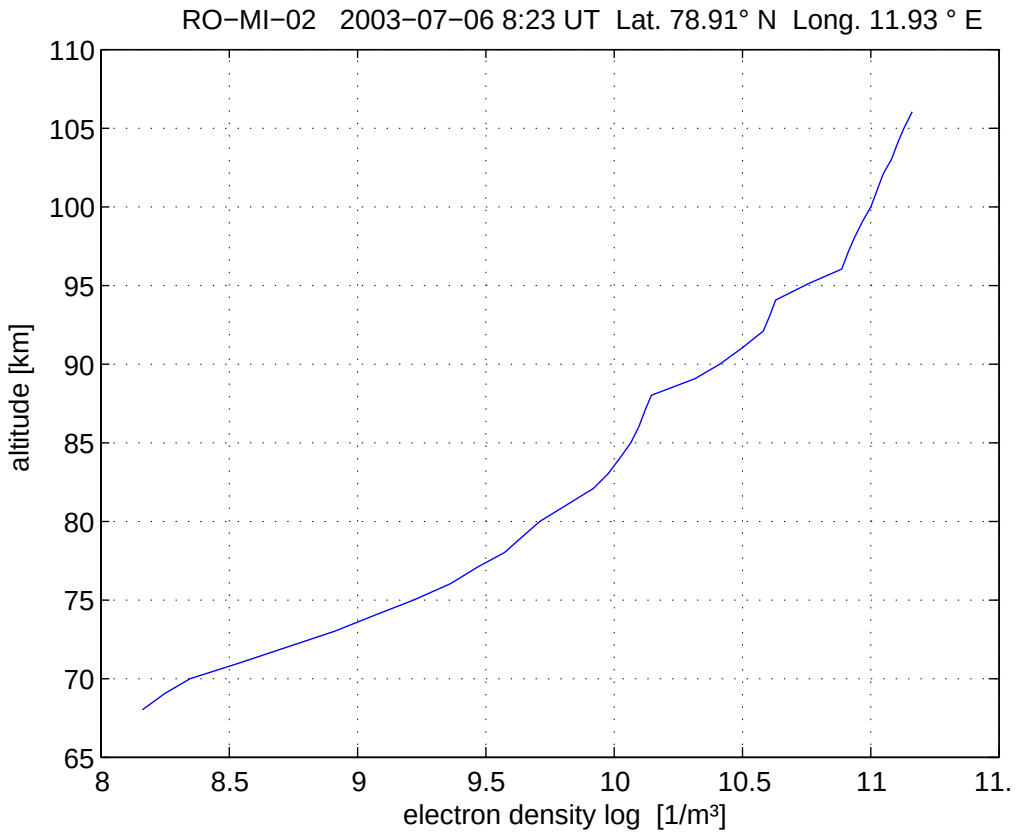


Figure 4.11 Rocket flight 6. 7. 2003, Svalbard. This is the only profile in the dataset based on a wave propagation experiment and hence inherently unbiased.

4.2.2 Svalbard Statistics and Distributions

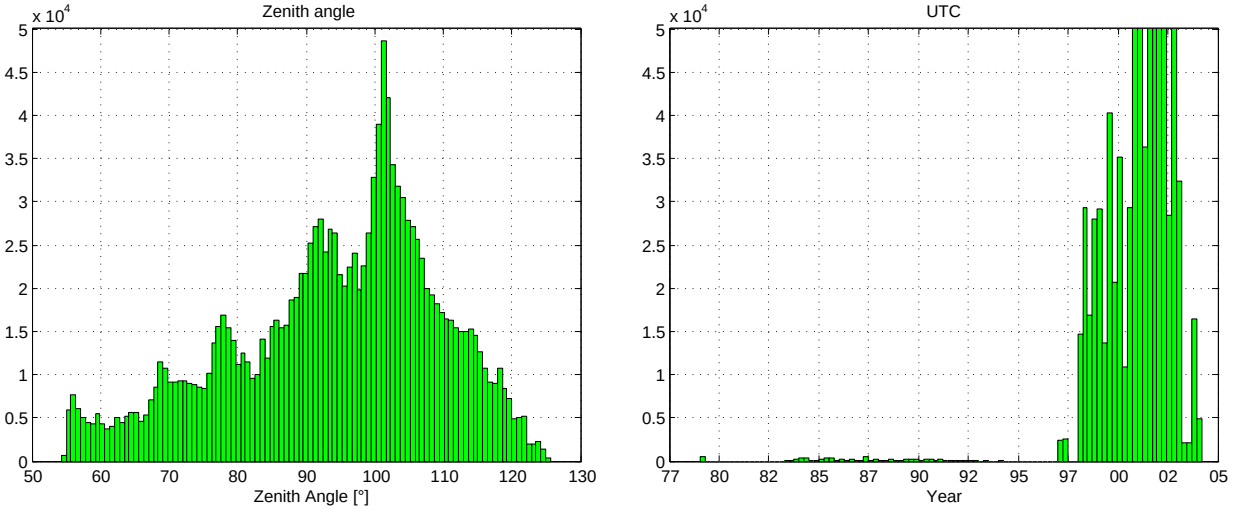
Table 4.2 shows the mean value, the 1^{st} sextile, the 1^{st} quartile, the 3^{rd} quartile and the 5^{th} sextile for all relevant parameters.

Variable	Unit	1.Sextile	1.Quartile	Mean	3.Quartile	5.Sextile
zenith angle	°	77.9	84.8	94.1	104.8	108.4
10.7 cm flux		137.9	147.4	174.5	197.5	210.6
dst		-39.0	-30.0	-19.9	-2.0	2.0
kp		0.7	1.0	2.2	3.0	3.3
ap		3.0	4.0	13.2	15.0	18.0
alt	km	126.0	141.0	273.6	348.0	451.0
$logN_e$	m^{-3}	10.47	10.64	10.94	11.33	11.46

Table 4.2 Most important input parameters are analysed for range coverage (Svalbard).

To develop a reliable model it is important to achieve a good coverage of input parameters ranges. On the following pages distributions for all relevant parameters are shown.

Distribution of zenith angle and year.

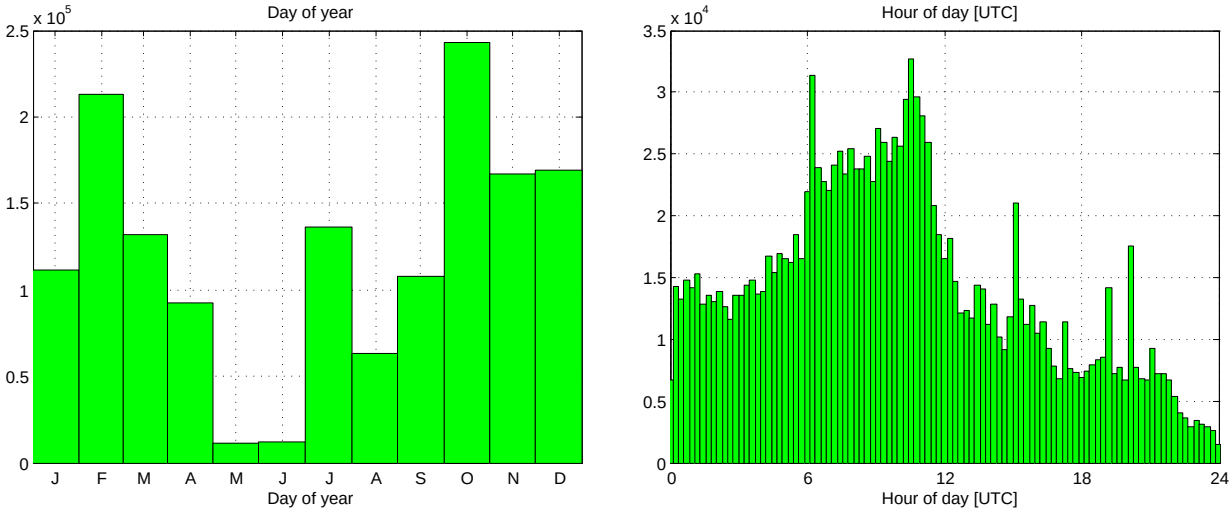


(a) Zenith angle, mean = 94.1 1. Q = 84.8 3. Q = 104.8 1. S = 77.9 5. S = 108.4

(b) Year

Figure 4.12 Svalbard, distribution of zenith angle and year. The few data before 1997 are the rocket experiments from Heiss Island.

Distribution of UT and hour number.

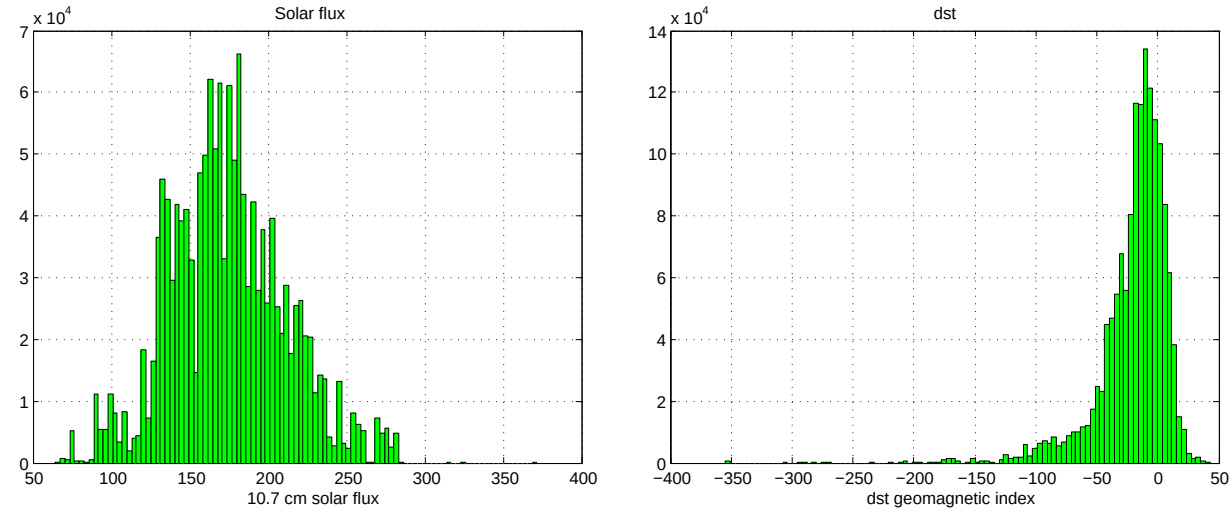


(a) Day of year

(b) Hour of day

Figure 4.13 Svalbard, distribution of UT and hour number

Distribution of F107 flux and dst. The data used here are somewhat biased towards high solar activity prevalent in the first years of assertion.

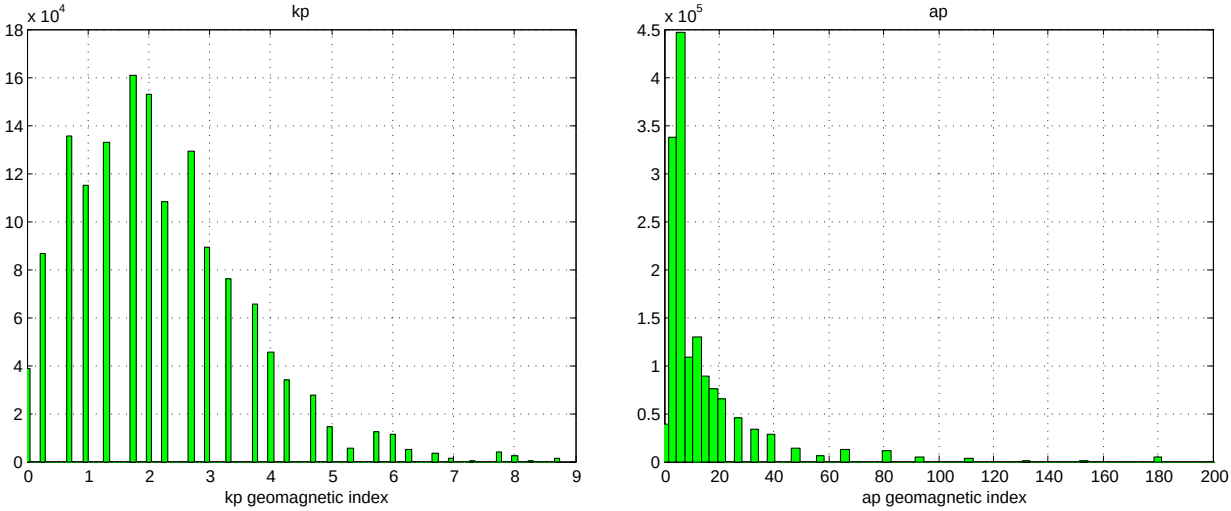


(a) Sun flux, mean = 174.5 1. Q = 147.4 3. Q = 197.5 1. S = 137.9 5. S = 210.6

(b) dst, mean = -19.9 1. Q = -30.0 3. Q = -2.0 1. S = -39.0 5. S = 2.0

Figure 4.14 Svalbard, distribution of F107 flux and dst

Distribution of kp and ap.

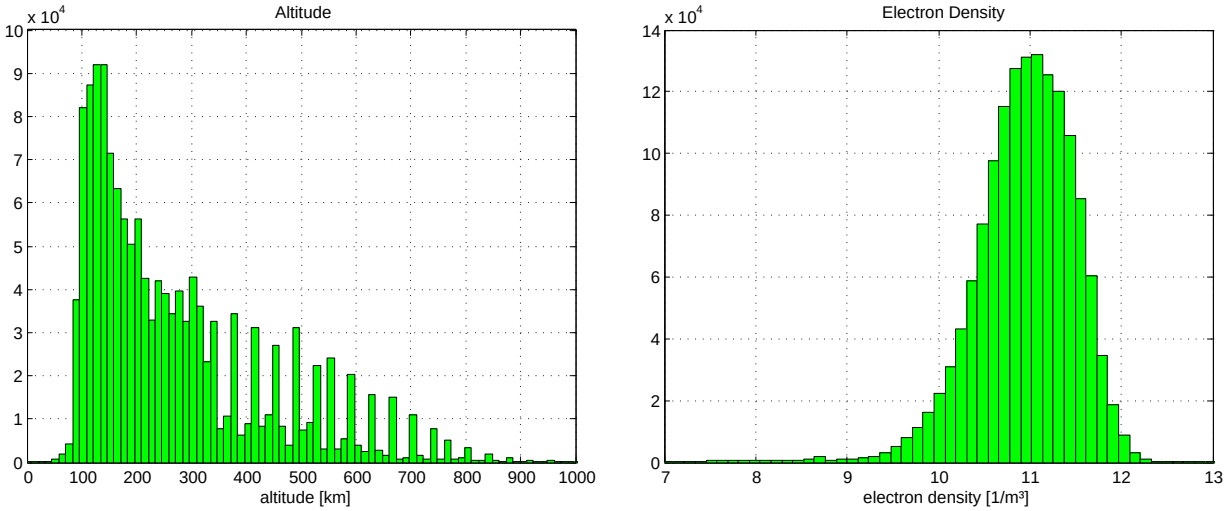


(a) kp, mean = 2.2 1. Q = 1.0 3. Q = 3.0 1. S = 0.7 5. S = 3.3

(b) ap, mean = 13.2 1. Q = 4.0 3. Q = 15.0 1. S = 3.0 5. S = 18.0

Figure 4.15 Svalbard, distribution of kp and ap

Distribution of altitude and electron density.



(a) Altitude, mean = 273.6 1. Q = 141.0 3. Q = 348.0 1. S = 126.0 5. S = 451.0

(b) Log electron density, mean = 10.94 1. Q = 10.64 3. Q = 11.33 1. S = 10.47 5. S = 11.46

Figure 4.16 Svalbard, distribution of altitude and electron density

4.2.3 Minimum Altitude

For interpretation of model results in Chapter 8 it is important to know the distribution of the minimum altitudes of Svalbard electron density profiles.
In Figure 4.17 all lowest datapoints of all profiles are displayed in a scatter plot. In Svalbard a large number of profiles reaches to at least 90 km due to measurements from rocket flights.

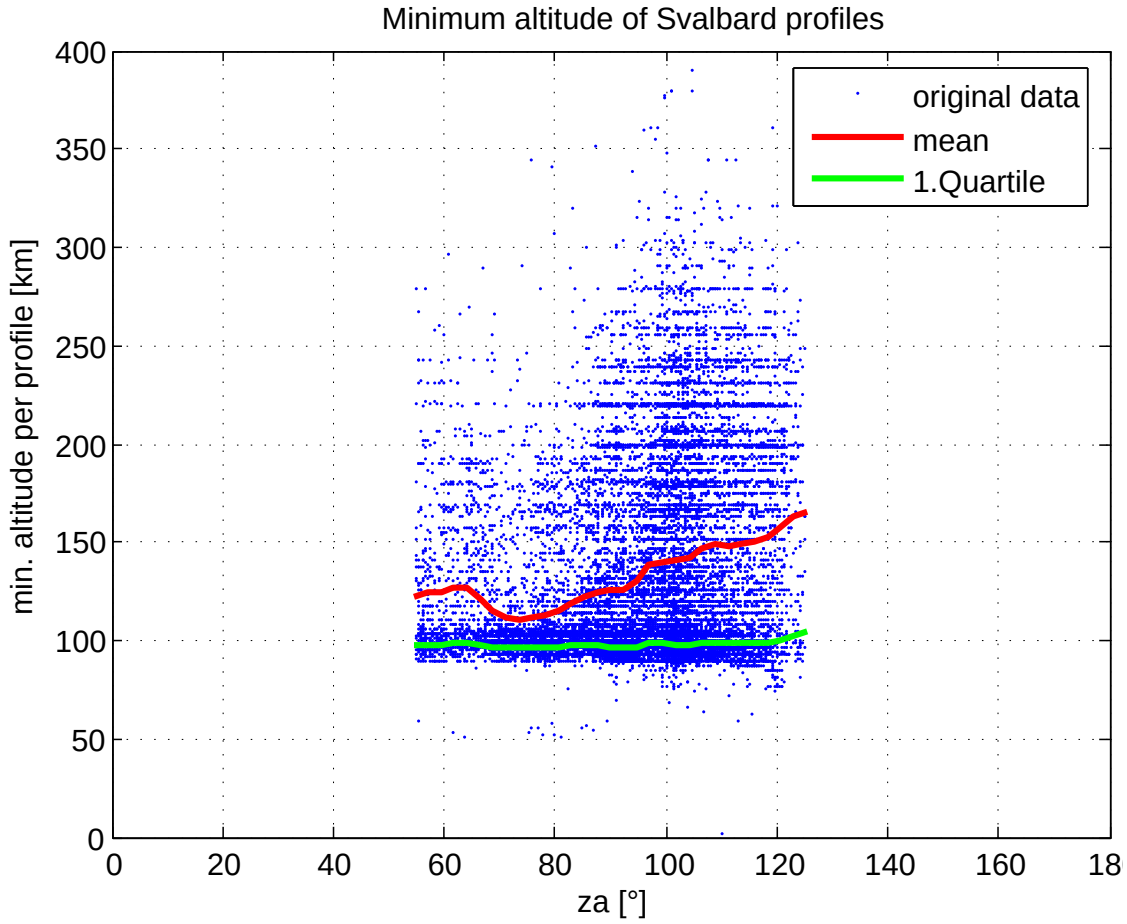


Figure 4.17 Variation of minimum heights including rocket data from Heiss Island. Note the considerable lower starting heights at the measurements compared to the Arecibo data. Due to low electron densities measurements are started at higher altitudes at night.

4.3 Conclusion

In this chapter the structure of the two datasets is shown. The Arecibo dataset consists of 2,077,902 datapoints in 98,554 electron density profiles (Ch. 4.1.3) and the Svalbard dataset consists of 1,459,223 datapoints in 162,506 electron density profiles (Ch. 4.2.2). These two datasets are used to train the artificial neural network and can consequently be called substructures of the two ionospheric models.

5 Input Parameters and Preprocessing

The better the problem is described through preprocessing, the more successfully the artificial neural network will solve the problem. For example, to just provide the number of hour (0 to 23) is not very effective, because the values do not show that it is a circular effect.

5.1 Zenith angle

The zenith angle is defined as angle between sun and zenith which means that the zenith angle is zero for highest solar position. At 90° the Sun's position is at the horizon.

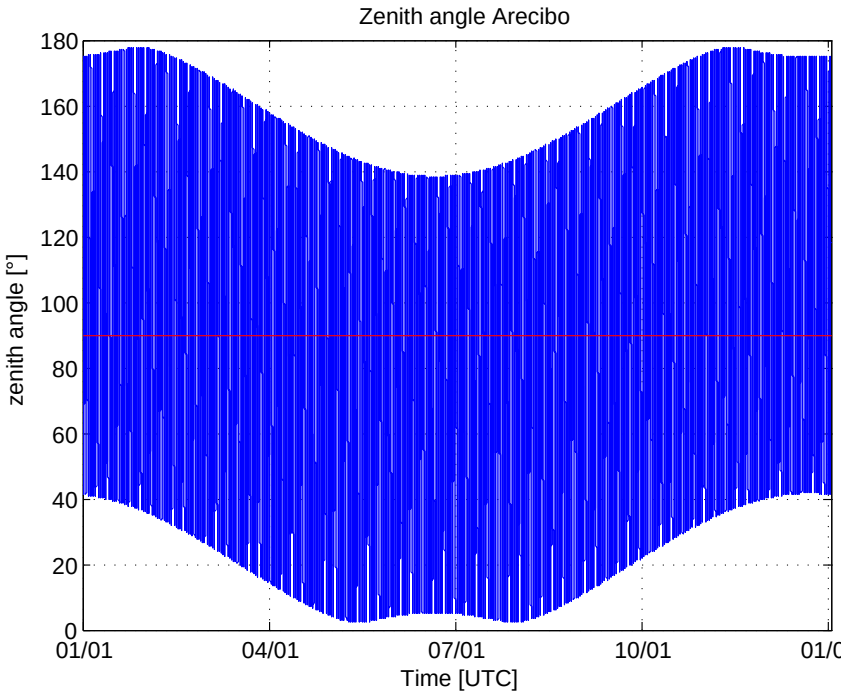


Figure 5.1 Diurnal variation of the solar zenith angle during a full year at Arecibo. Note the large diurnal variation and that the extrema occur *twice* a year.

In Figure 5.1 the zenith angle for one year is plotted (Arecibo). Due to low latitude the sun reaches relatively small (and very large) zenith angles throughout the year.

In Figure 5.2 the zenith angle for one year is plotted (Svalbard). Due to a high latitude (beyond polar circle) the sun reaches middle zenith angles of 55° at day time and 80° at night time.

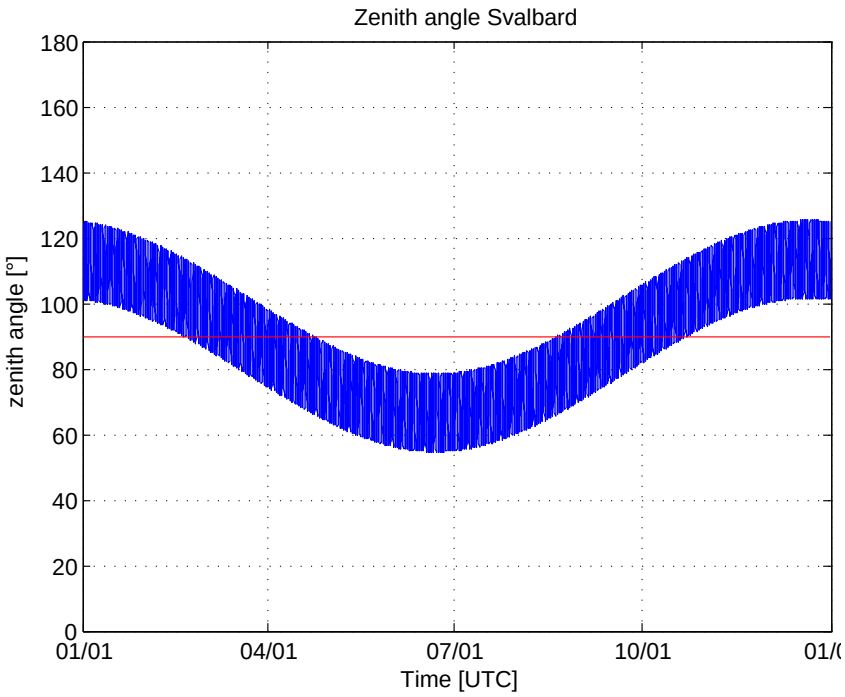


Figure 5.2 Solar zenith angle of a full year at Svalbard. The seasonal variation is much larger than the diurnal.

This 24 hour day leads to a very high insolation in midsummer at polar cap. This effect can be seen in Figures 5.6 and 5.7. On the other hand the zenith angle never reaches 90° around midwinter (polar night).

5.2 Chapman production

As input parameter for the diurnal variations of electron density the Chapman production was choosen because of its direct correlation.

The Chapman grazing incidence function as defined 1931 by Chapman:

$$Ch(x,\chi) = x\ sin\chi \int_0^x e^{x-x\ sin\chi/sin\alpha} cosec^2\alpha d\alpha \tag{5.1}$$

where $x = (R+h)/H$ and χ is the solar zenith angle, R is the Radius of the Earth, h is the altitude and H is the scale height of the absorbing gas. Unfortunately Equation 5.1 is not analytically solvable, thus a numerical approach was taken. The inverse Chapman grazing incidence function is shown in Fig. 5.3.

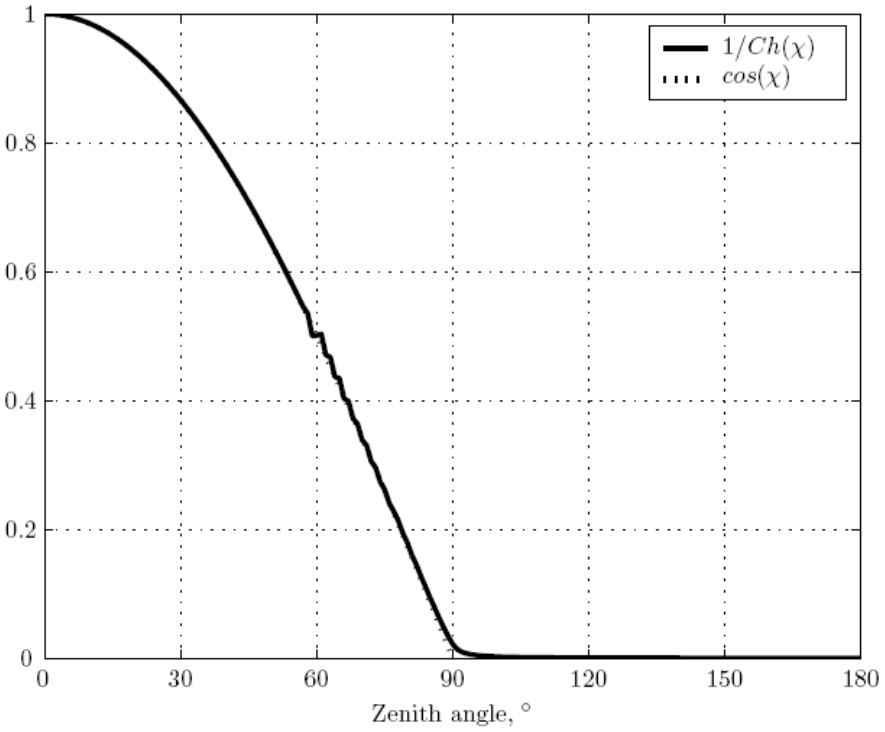


Figure 5.3 Inverse Chapman grazing incidence function vs. solar zenith angle with comparison to cosine (H = 5 km, Altitude = 100 km) [Egger, (2004)].

The Chapman production means electron-ion pair production by ionization. As a simplification an approximation function was used here. It consists mainly of the cosine of day time zenith angles.
The characteristics of the Chapman production at spring equinox is plotted in Figures 5.4 (Arecibo) and 5.5 (Svalbard).

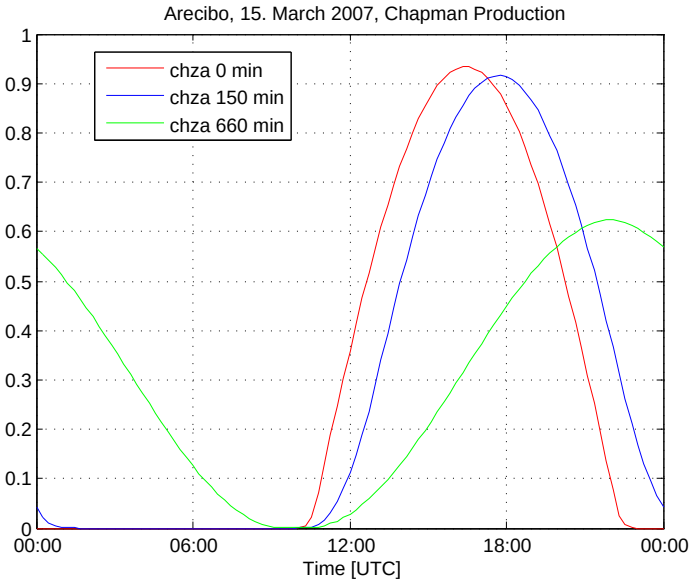


Figure 5.4 Chapman production, Averages: 0, 2.5 h, 11 h, Arecibo midnight: 04:30 noon: 16:30

To allow a diurnal asymmetry for the model a second diurnal input parameter is needed. Therefore an 11 hour average of Chapman production is calculated. With this pair of input parameters the neural network can "learn" a difference between a morning datapoint and an evening datapoint for the same zenith angle. If we would use the zenith angle as an input parameter the neural network would not have this information. Detailed information to Chapman production can be found in [Titheridge, (1988)].

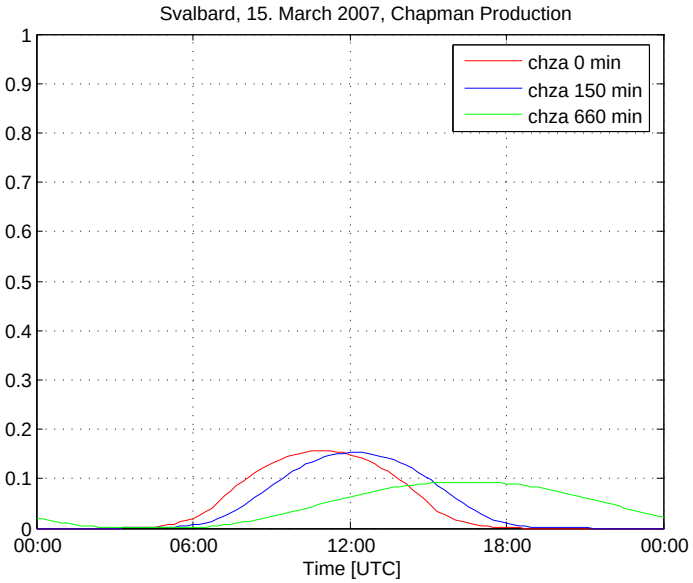


Figure 5.5 Chapman production, Averages: 0, 2.5 h, 11 h, Svalbard

5.3 Daily integrated insolation (Dii)

The daily integrated insolation is used to provide the neural network information about the seasonal location of the datapoint. The cosine of the zenith angle is integrated over one day from sunrise to sunset.

A daily integrated insolation of value "1" means, that the Sun would shine 24 hours with an zenith angle of 0° (i.e. 1.18 10^8 Ws/m^2).

In Figure 5.6 values of daily integrated insolation are coded in colour for every day of the year and all latitudes. In this contour plot the variation in Sun - Earth distance (of about ±5%) is also considered which makes the plot is asymmetric.

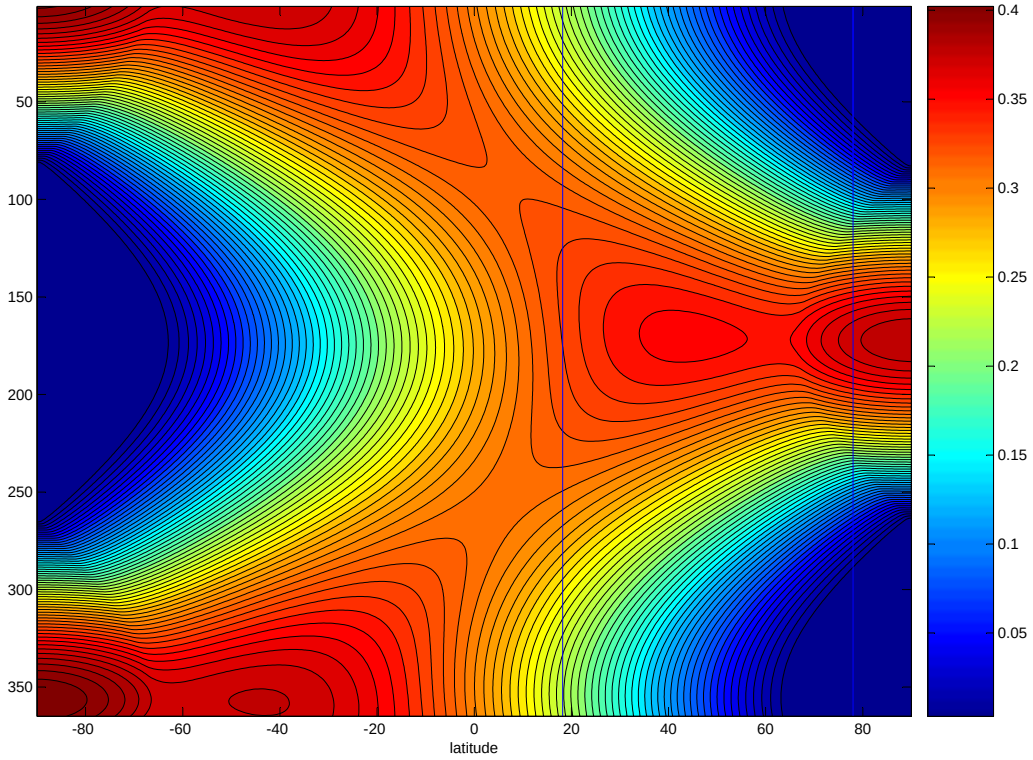


Figure 5.6 Daily integrated insolations for any latitude and day of year. Two lines mark the latitudes of Arecibo and Svalbard, respectively. Note that in the equatorial region the daily integrated insolation has two maxima in a year (more pronounced in the southern hemisphere).

In Figure 5.7 the seasonal characteristics are plotted for the latitudes of Arecibo and Svalbard.

To achieve a seasonal asymmetry we use a pair of seasonal input parameters. At first the daily integrated insolation and as second input a 30 day average value of the daily integrated insolation is used. With this pair of input parameters the neural network can distinguish between spring and autumn values.

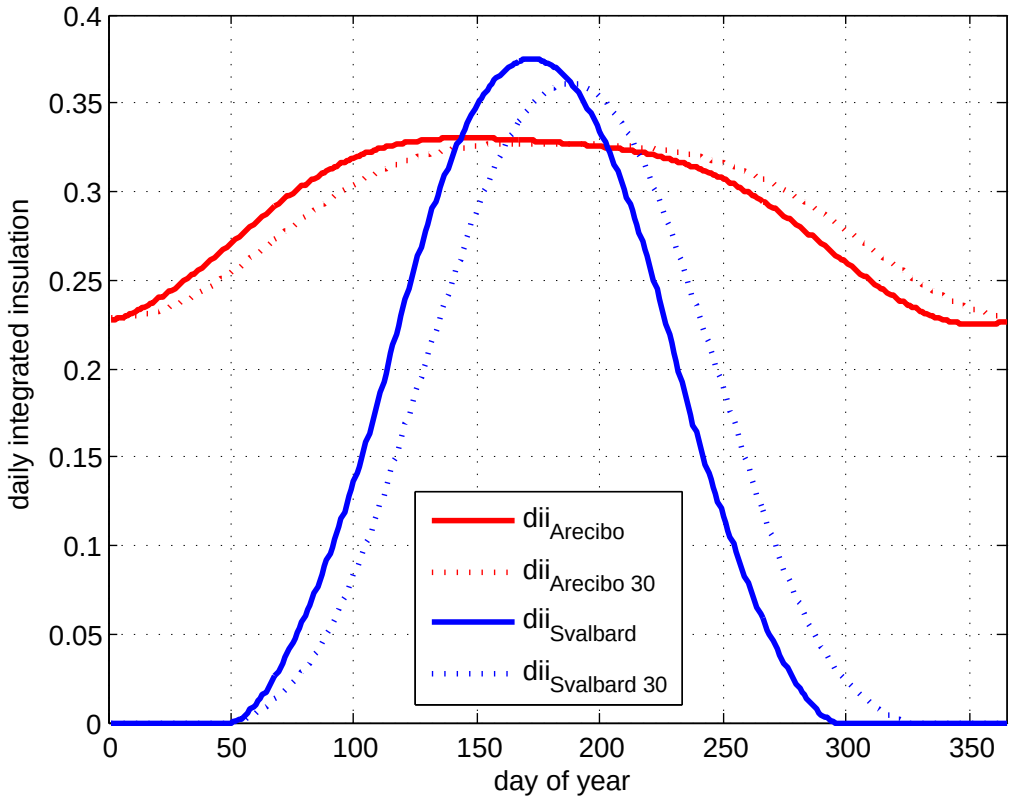


Figure 5.7 Daily integrated insolation for Arecibo and Svalbard (average of the 30 preceeding days is plotted).

5.4 Solar activity (10.7 cm flux)

As a measure for solar activity the value for the solar flux at 2800 MHz is used. The Sun changes its magnetic field every 11 years. At this time solar activity increases and more radiation and particles are emitted.

The solar flux is measured since 1947 by the NRC[1]. One precise measurement is taken at local noon (20:00 UT in Penticton since 1991). This flux is called "observed flux". Out of this raw value an so-called "adjusted flux" which considers the variation of Sun - Earth distance is calculated (corresponding to 1 AU); AU = astronomical unit = mean Sun-Earth distance (149.6 10^6 km).

[1] National Research Council of Canada

As input parameter for the neural network the *observed* flux is used, because for the ionosphere the intensity of the solar flux hitting the Earth is relevant.

The solar flux from 1960 to 2000 is plotted in Figure 5.8.

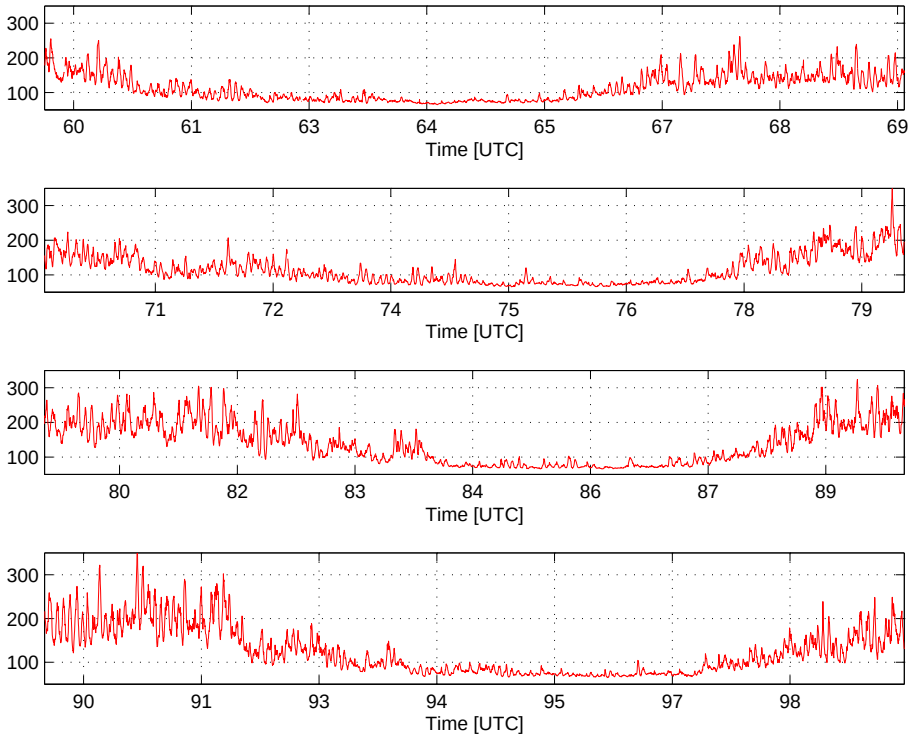


Figure 5.8 10.7 cm solar flux (1960 - 2000).

In Figure 5.9 the solar flux and several running means are plotted for some months in 1991. Further information how to choose average time can be found in Ch. 6.2.5.

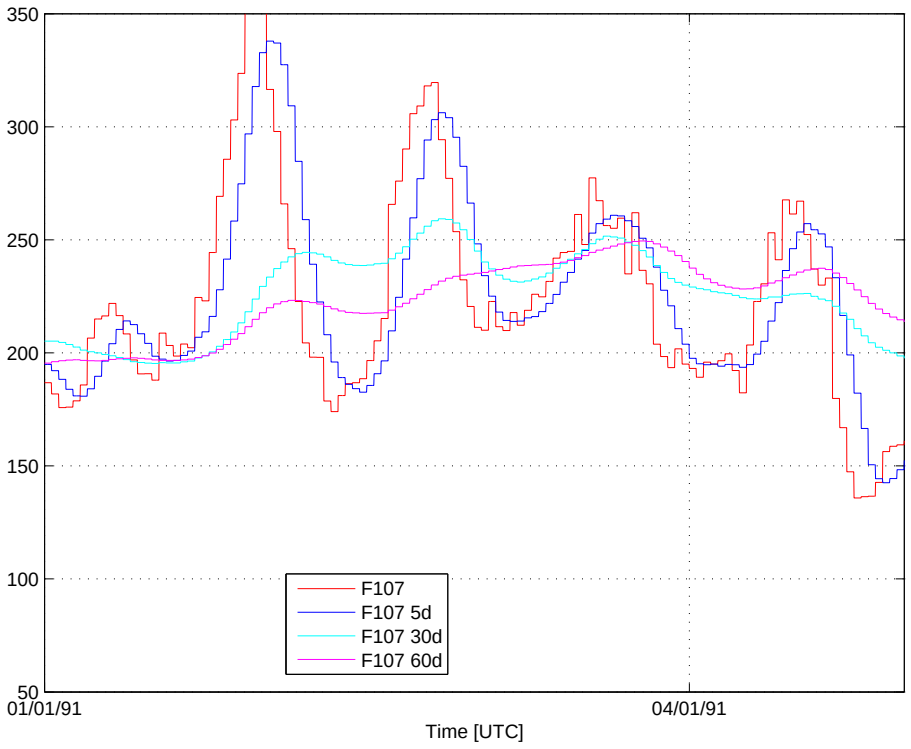


Figure 5.9 10.7 cm solar flux, average calculation range [days], (year 1991 JAN-APR).

5.5 Magnetic indices ap, kp, Dst

Disturbances reflected in deviations of the Earth's magnetic field occur due to currents in the E-region, in turn due to enhanced electron densities. To provide information of magnetic conditions geomagnetic indices are used.

The *Dst* index of magnetic storms gives information about the reduction of the magnetic H component at the equator due to the ring current. *Dst* is measured in units of nanotesla and is mainly interesting for singular magnetic storms. Further information is given in [Hargreaves and Hunsucker, (2003)].

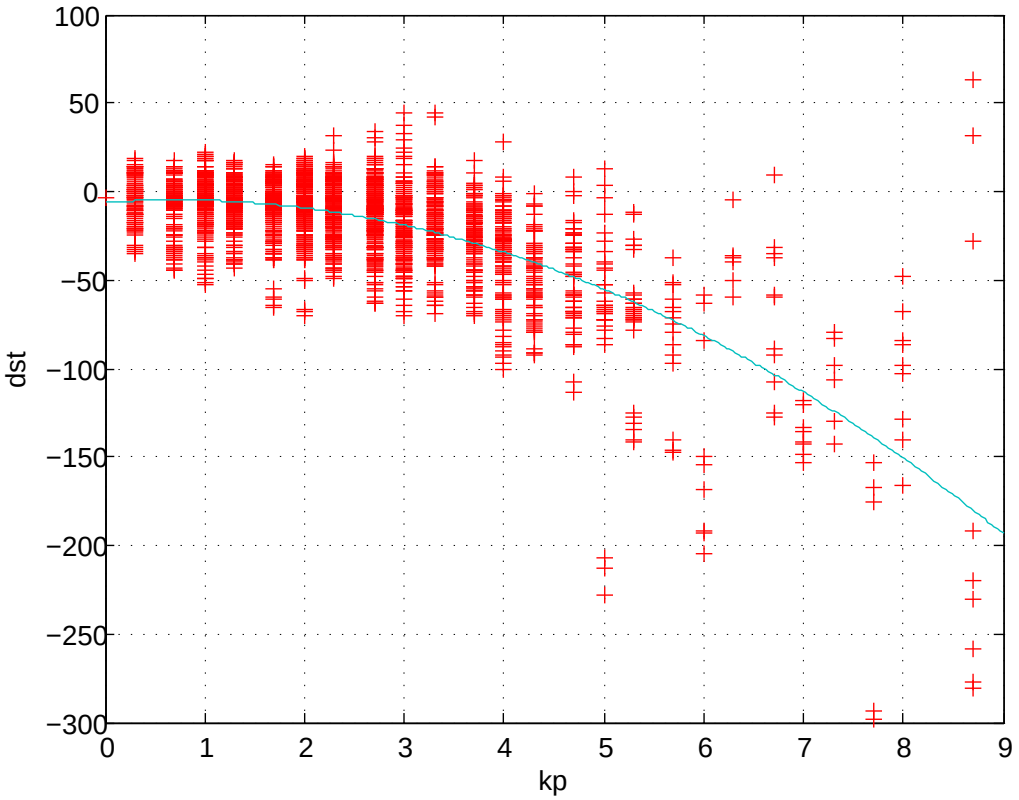


Figure 5.10 Relation between the geomagnetic indices Dst and kp from 1979 to 2003.

The geomagnetic indices *ap* and *kp* give information on disturbances in the geomagnetic field. The index kp is measured by 13 non-auroral observatories all over the world. It is a mean value over 3 hours and was introduced by J. Bartels in 1949 and published by the IAGA[2] (formerly IATME[3]) in 1951 [Egger, (2004)].

In Figure 5.10 the relation between *kp* and *Dst* is shown.

Figure 5.11 shows characteristics of the geomagnetic indices *ap*, *kp* and *Dst* for some months in 1991. When a magnetic disturbance occurs, all indices display significant excursions.

[2]International Association for Geomagnetism and Aeronomy
[3]International Association for Terrestrial Magnetism and Electricity

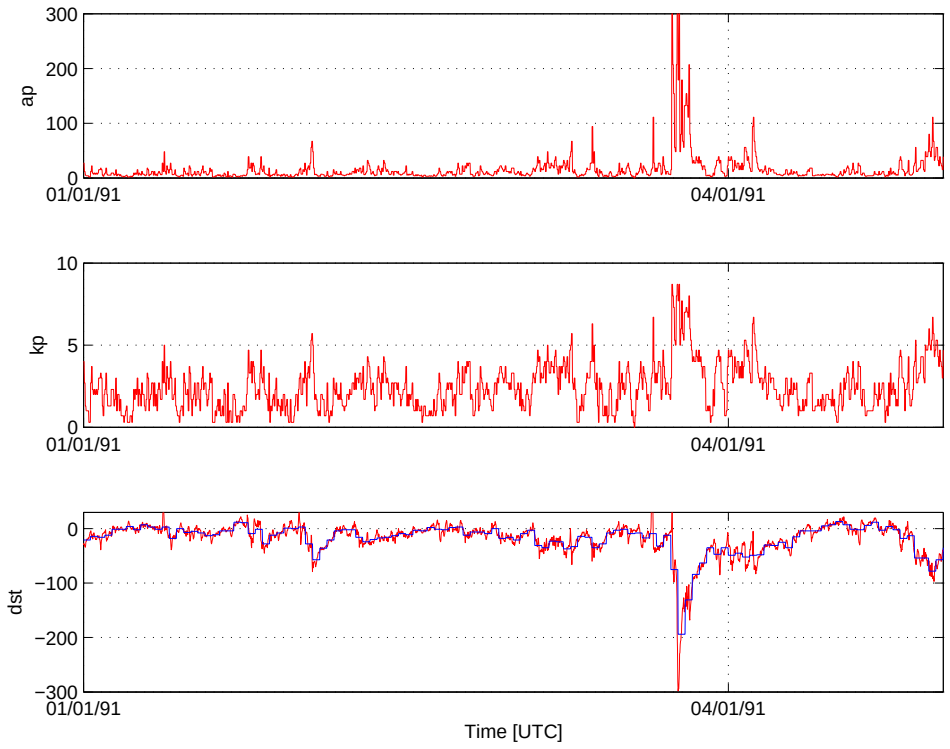


Figure 5.11 ap, kp and dst, geomagnetic indices (year 1991 JAN-APR)

The geomagnetic index *ap* is derived from *kp* via a fixed relation. The relation and the data can be found at NOAA[4].

Due to a better ability to describe normal situations, ap was choosen as input parameter to provide information for geomagnetic conditions. To achieve time dependence an average value of ap is used as well to model more inert characteristics of electron density.

[4]National Oceanic and Atmospheric Administration, http://www.ngdc.noaa.gov/stp/GEOMAG/kp_ap.html

In Figure 5.12 ap and several running mean values are plotted.

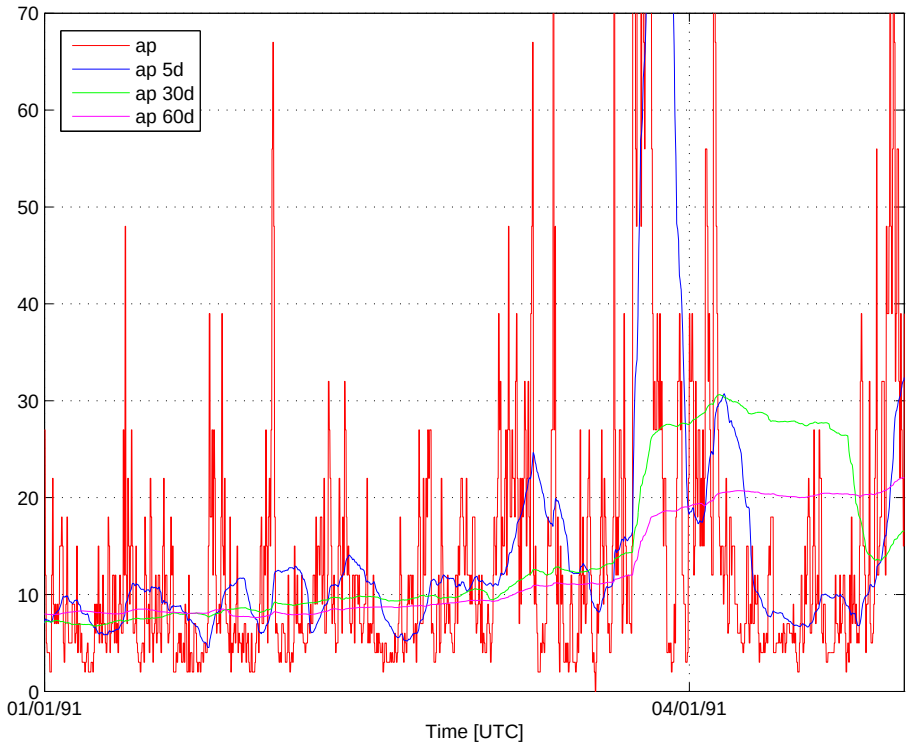


Figure 5.12 ap running mean [days], (year 1991 JAN-APR)

5.6 Normalization

To make values comparable and non-dominant it is important to normalize them. The values of all input and output parameters are normalized and stored in a dedicated file format suitable for the SNNS[5]. Values are normalized to a range of 0 to 1. More information about normalization is given in Ch. 11.1.2.

[5]Stuttgart Neural Network simulator

5.7 The day-night border

The day night border has to be calculated for every datapoint. The solar UV radiation which mainly causes ionization and photo detachment is absorbed in the ozone layer. That is the reason why we have to take a larger radius than the geometric Earth radius. This calculation is illustrated in Fig. 5.13.

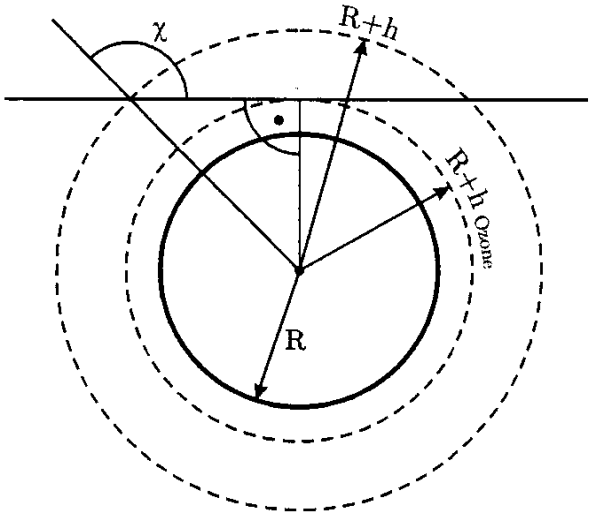


Figure 5.13 Geometry of day night border. The sun rises at the right hand side.

The ozone layers altitude h_{Ozone} is here assumed to be at 25 km in accordance with atmospheric models [Isaksen, (1973)]. For an altitude h the day-night border is calculated as follows:

$$\chi_O = 180 - \arcsin(\frac{R + h_{Ozone}}{R + h})$$

For example $h = 100$ km yields a day night border of 98° [Egger, (2004)].

With this calculation we can separate between day and night data because we have a defined sunset for every altitude.

5.7.1 Hour Number starting at Sunset

With the exact sunset time calculated by a regula falsi algorithm the time after sunset can be determined for every datapoint.

In Figure 5.14 and 5.15 all datpoints are plotted to show the relation between hour number after sunset and zenith angle.

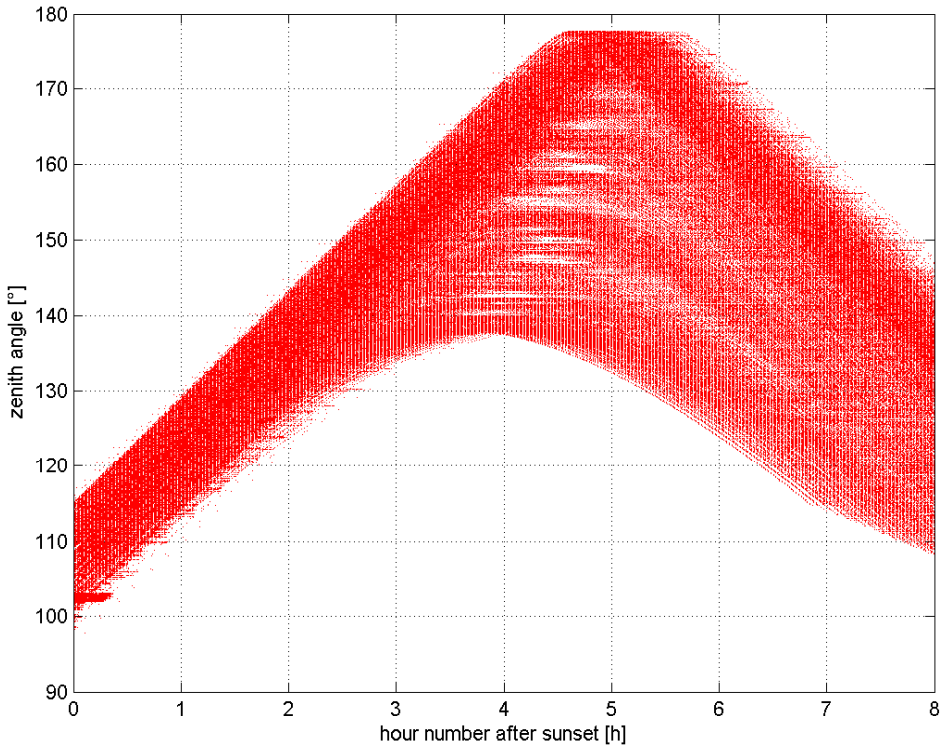


Figure 5.14 Arecibo, zenith angle vs. hours after sunset. The length of the night can be seen as the time after sundown to midnight. It varies from 4 hours in summer with a maximum zenith angle of 140°, to 5 hours in winter when the zenith angle can reach 180°.

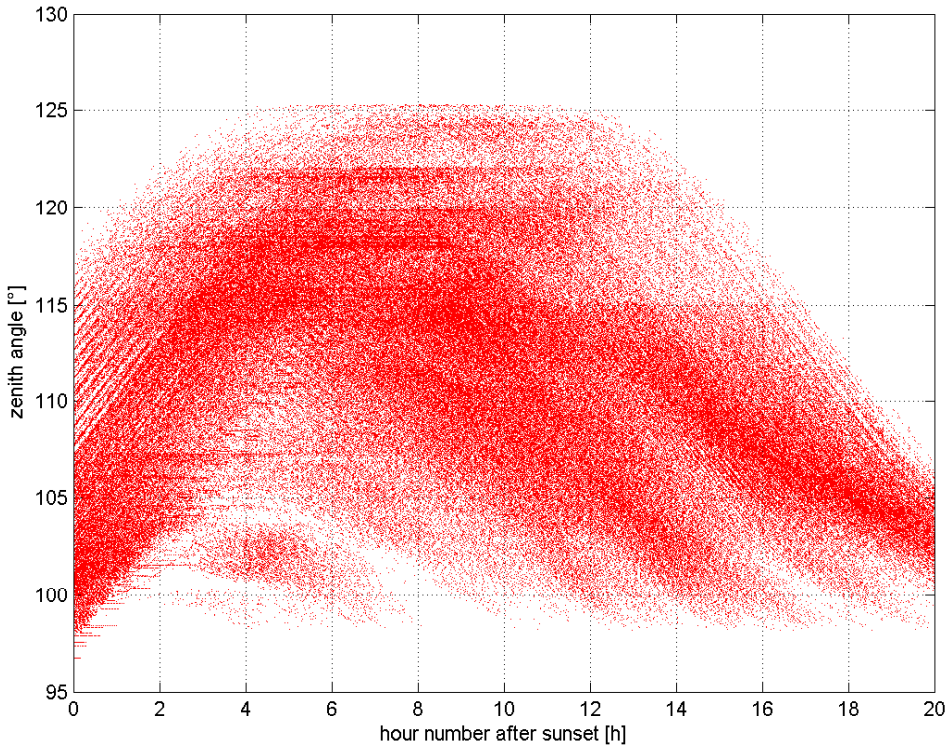


Figure 5.15 Svalbard, Zenith angle vs. hours after sunset. Note the different scales compared to the previous figure of both the zenith angle range (maximum 125°) and the length of the night.

5.8 Conclusion

In this chapter all calculated input parameters of the artificial neural network are described. Zenith angle, Chapman production, daily integrated insolation, solar activity and magnetic index are used to present the neural network data which determine the physical situation of the datapoint. For example, is for the seasonal variation the daily integrated insulation and a 30 day average of daily integrated insulation used, to ensure asymmetric behavior and a closed "loop" around the year as well. Closed "loop" means that at the end of the year is the same situation as at the beginning of a year. With day of year alone as input parameter the neural network is not able to learn these conditions.

6 Development of an empirical model using neural networks

The main problem of modelling physical processes using measured data is to balance two reqirements. On the one hand we want to fit the model as good as possible to the measured data and on the other hand we want to get a model which is able to generalize and make predictions.

On the following pages it is described how to account for these needs and find an optimum network architecture, input parameters and learning parameters. An introduction to artificial neural networks was given earlier in Chapter 2.2 Neural Networks.

6.1 Neural Network Architecture and Learning Parameters

In this section different neural network sizes, activation functions and training parameters are compared. At first it is important to mention that it is not easy to find a proper architecture and parameters for a specific application. A lot of empirical work has to be done because it is hard to find thumb rules in literature. In Table 6.1 all used neural networks are listed. Many different neural networks were trained with different input parameters and training parameters.

6.1.1 Importance of Architecture

In Figure 6.1 it is shown, which network yields which RMS at different altitudes. It is difficult to make a decision, because networks with a lower RMS at low altitudes have a higher RMS at higher altitudes and vice versa. So the overall RMS values are compared, which can be seen in Figure 6.2 and Table 6.2.

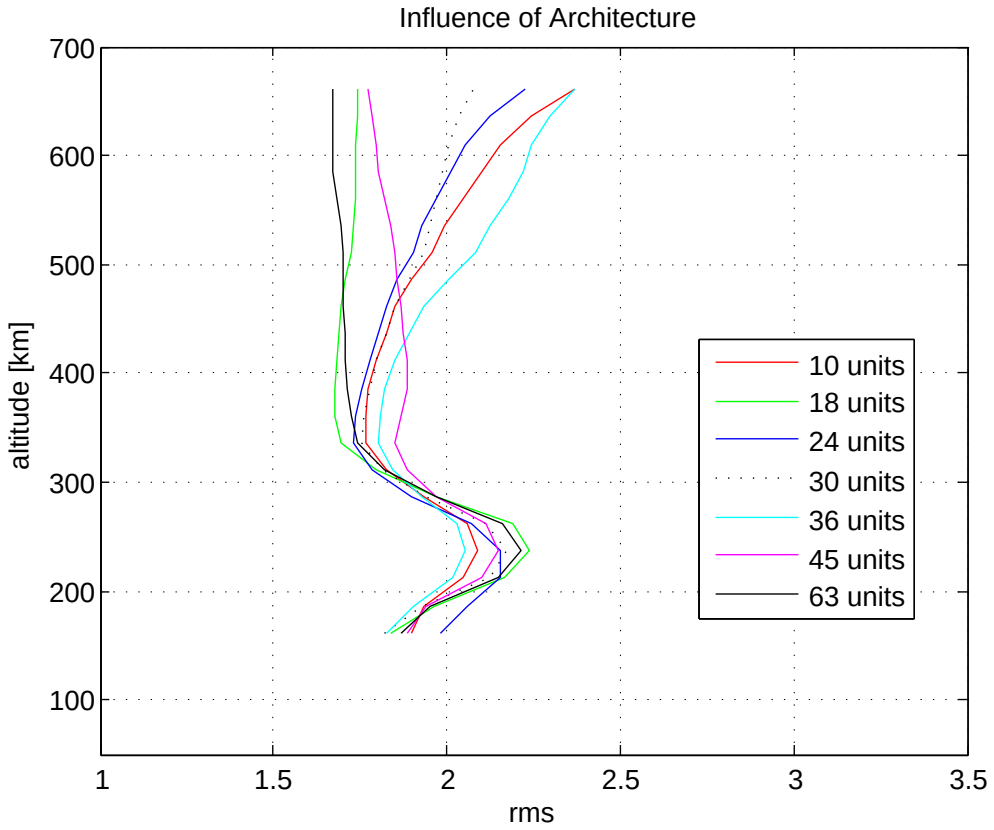


Figure 6.1 RMS error factor as a function of altitude for different network architecture (number of hidden layers/units). At higher altitudes more complex networks yield a better result.

hidden units	weights	hidden layers	units per layer
10	35	2	5
18	54	3	6
24	70	4	6
30	86	5	6
36	85	4	9
45	104	5	9
63	142	7	9

Table 6.1 Collection of neural networks which were used for model development, number of weights are calculated with 10 inputs

Description	$\mathbf{RMS}_{all}$	$\mathbf{RMS}_{train}$	$\mathbf{RMS}_{val}$	$\mathbf{RMS}_{test}$
10 units	2.31	2.51	1.88	1.97
18 units	2.13	2.29	1.84	1.80
24 units	2.19	2.32	1.88	1.94
30 units	2.17	2.32	1.84	1.93
36 units	2.26	2.42	1.85	2.01
45 units	2.17	2.32	1.85	1.90
63 units	2.12	2.28	1.85	1.79

Table 6.2 Influence of architecture

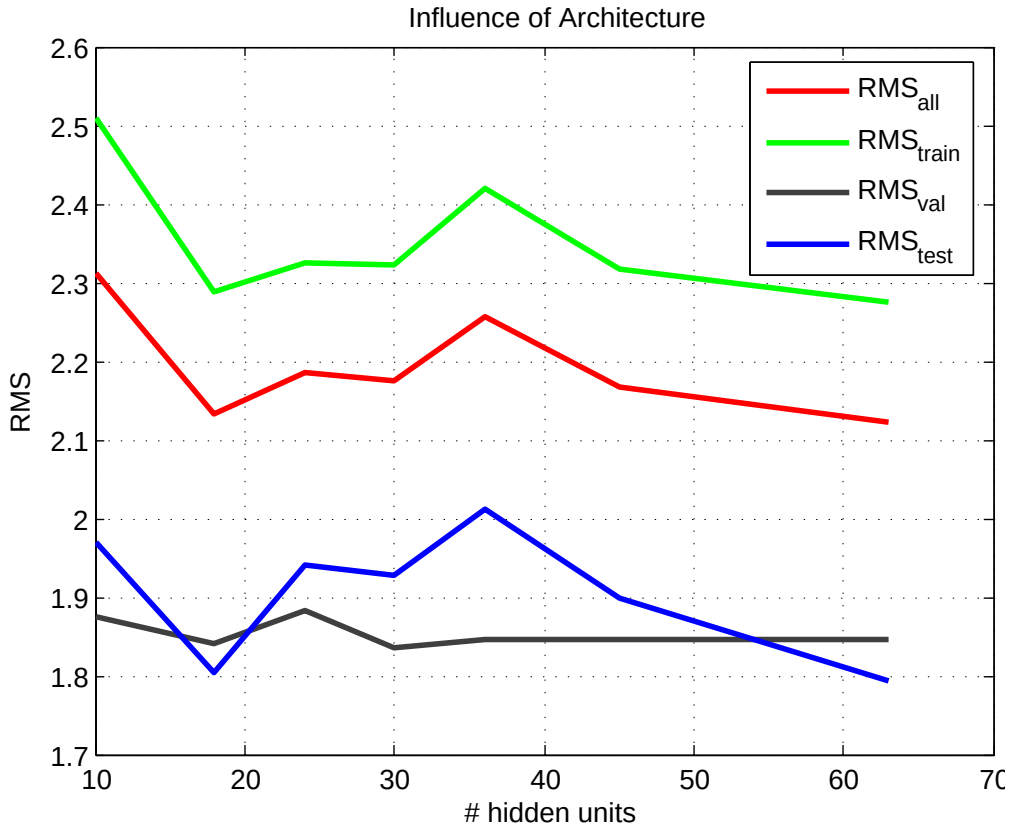


Figure 6.2 Arecibo, RMS error factors for different architectures and datasets (Arecibo).

As can be seen in Figure 6.2 there is an optimum for RMS at 18 hidden units. This neural network has 3 hidden layers with 6 neurons each and yields a good performance concerning RMS and as well training time. The training time for this neural network is about 4 hours with 400 training cycles and approximately 1.3 million training and validation datapoints (input patterns).

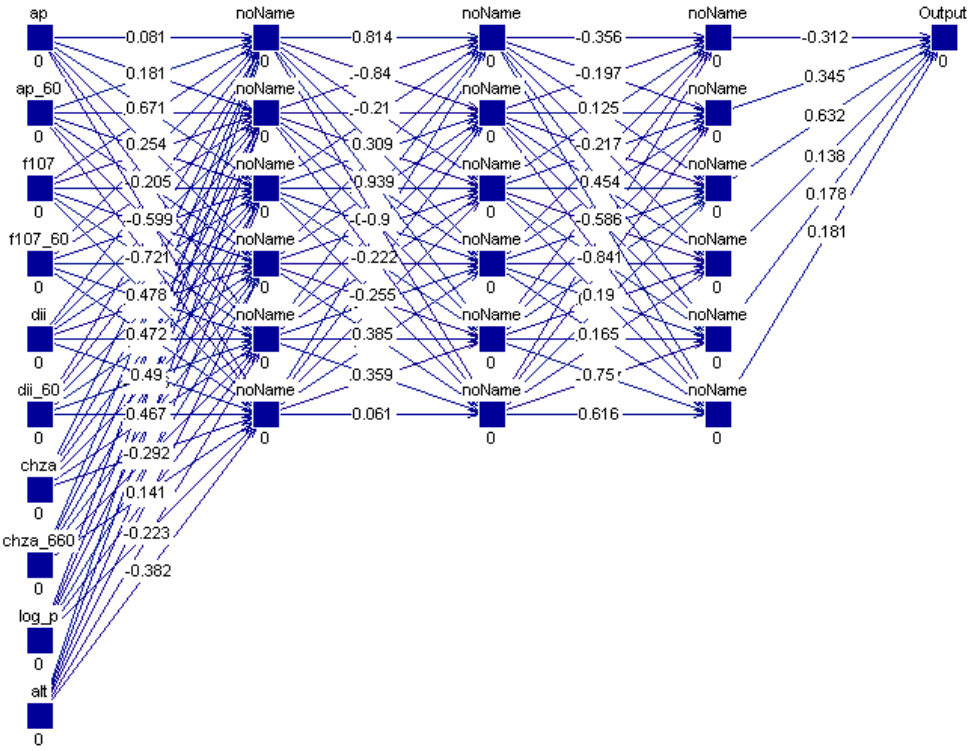


Figure 6.3 Architecture of a neural network with 10 input neurons and 3 hidden layers with 6 neurons each (Screenshot from SNNS Creator GUI).

A visualisation of a neural network is shown in Figure 6.3. On the left side there are the input neurons, then the signals propagate via weighted connections to the first hidden layer. The output of each neuron is connected to every neuron of the next layer. In a neuron all input values (weighted via connections) are summed and the output value is calculated via the activation function.

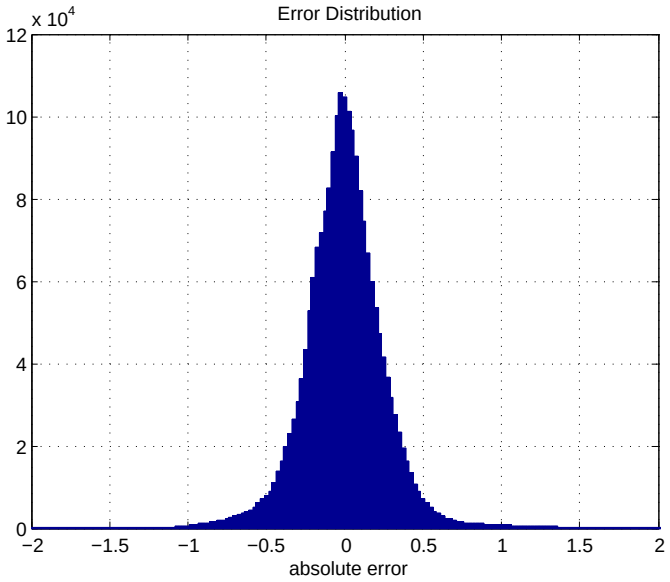


Figure 6.4 Arecibo, distribution of absolute error, 1 is an error of a factor of ten in electron densities.

6.1.2 Influence of Training Parameters

The question of choosing suitable training parameters for the neural network is very dificult. The only way to find parameters for a new modelling requirement is by trial and error.

The learning rate η is a factor which controlls how much the weights are adjusted dependent on the error (difference between the real output and the teaching input).

The learning rate η has to be low enough to find a local minimum and on the other hand it has to be high enough to ensure an acceptable learning time.

The maximum difference d_{max} between a teaching value and an output unit which is tolerated.

In all neural networks which are used here a η (learning rate) of 0.04 a d_{max} of 0.03 and number of training cycles of 400 is used. These values need long training times but they lead to reliable and compareable results. As an activation function there is always

tanh used. In Figure 6.4 the distribution of the absolute error is shown. All training data patterns are presented to the neural network, and prediction values are the result. These predictions are compared with the original measured data.

6.1.3 Influence of Pattern Selection

If as test data the most recent 20 % of data are taken(Fig. 6.6), the RMS for test data is lower than the RMS for training data. This may appear strange but is probably caused by different quality of the Arecibo data. I assume that there is less scatter in more recent Arecibo data.

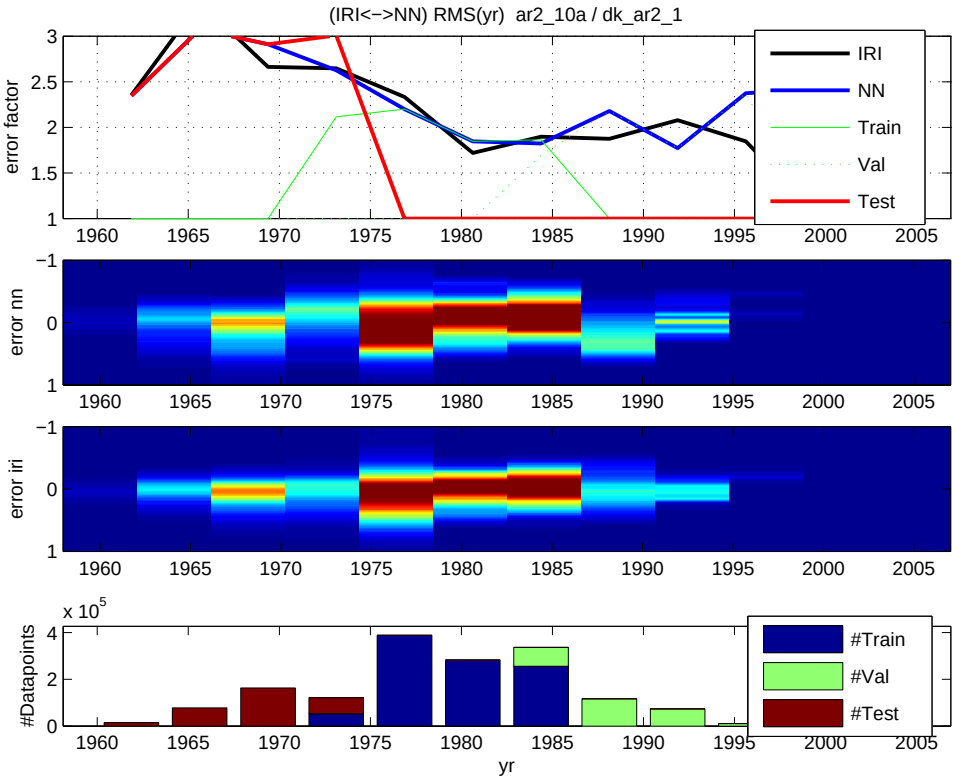


Figure 6.5 RMS (yr) 18 hidden units. Version A, oldest data is used for testing.

If as test data the oldest 20 % of data are taken, (Fig. 6.5), the RMS for test data is higher than the RMS for training data. To avoid this effect, training data and validation data are split profile-wise and randomly over the the entire time range (Fig. 6.7). To be sure, that there are no dependencies between training and test data the most recent 20 % of data are used for testing only in development process (Fig. 6.6).

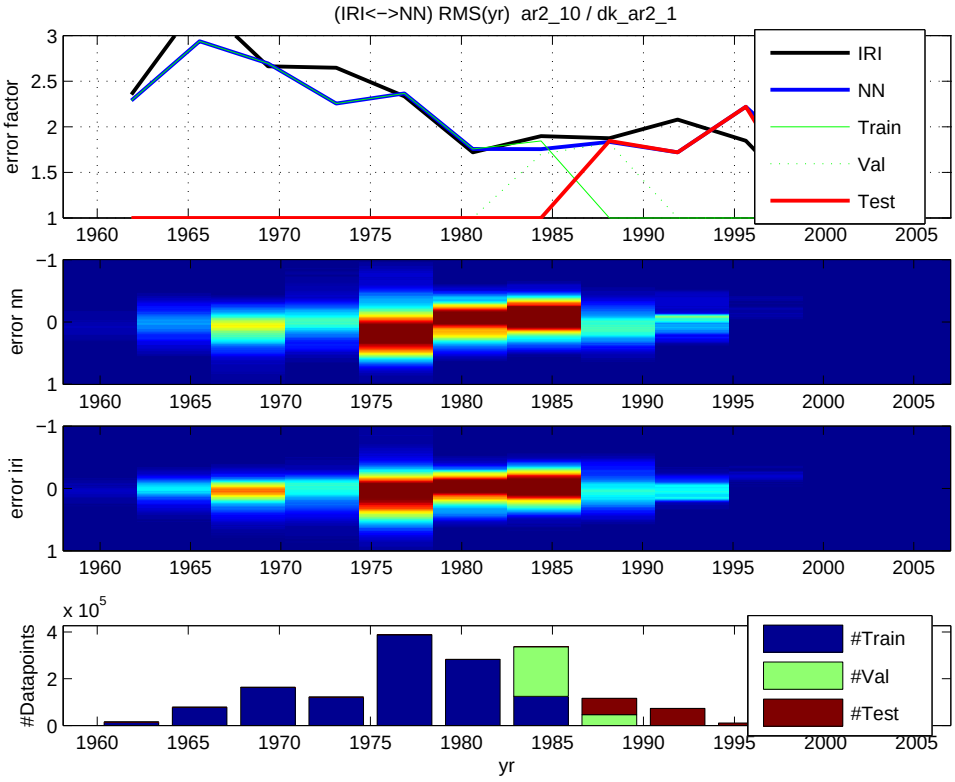


Figure 6.6 RMS (yr) 18 hidden units. Version B, youngest data for testing.

For the model version B is used as shown in Figure 6.6

The most recent 20 % of data is used for testing. the RMS_{test} obtains lower results than the training dataset what seems to be unlogical.

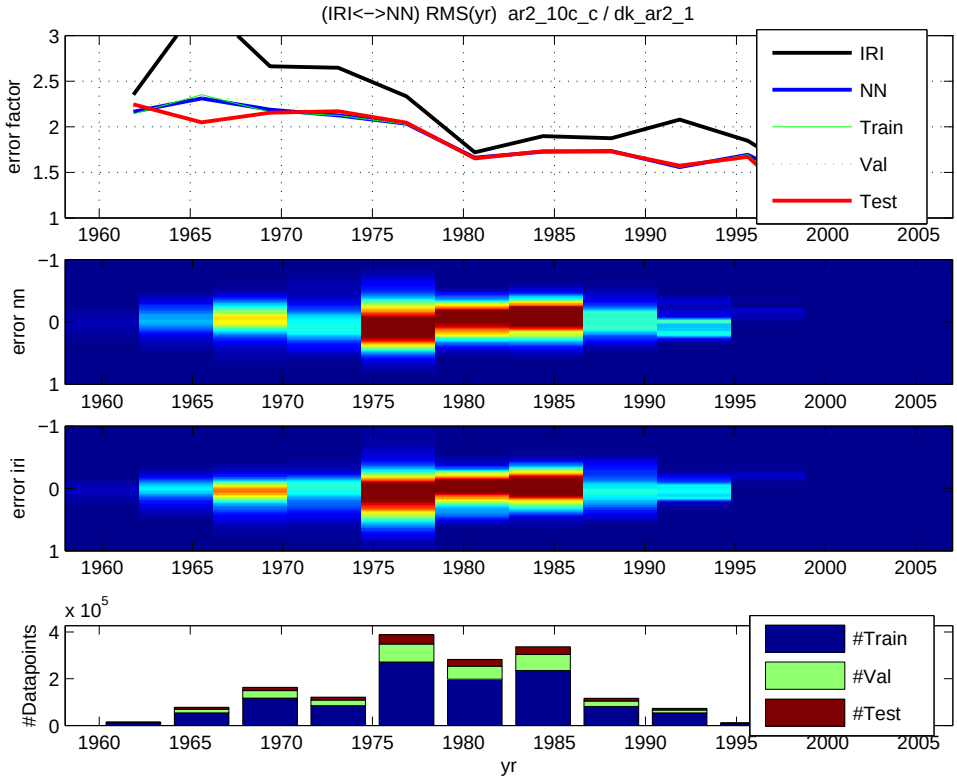


Figure 6.7 RMS (yr) 18 hidden units. Version C, all data (training, validation and test) are profilewise randomly distributed.

6.2 Combinations of Input Parameters

The main factor of developing a proper model is to choose the right input parameters. In Ch. 5 (Preprocessing) all input parameters are described. Here is described how average times and optimal combinations can be found.

6.2.1 Altitude and Pressure

The information of altitude for every datapoint comes with the dataset. To proof whether it is better to use altitude or air pressure, the air pressure is calculated for every datapoint with the MSIS model, described in Ch. 3.2. In Figure 6.8 three models are compared. One with altitude, one with logarithmic air pressure and one with both quantities as input parameters. Between 160 and 310 km altitude gives better results and otherwise logarithmic air pressure gives best result. Thus for further models *both* quantities are used as input parameters.

In theory pressure should provide better results because energetic radiation at a point in the atmosphere depends on the air column above this point (i.e. the local pressure).

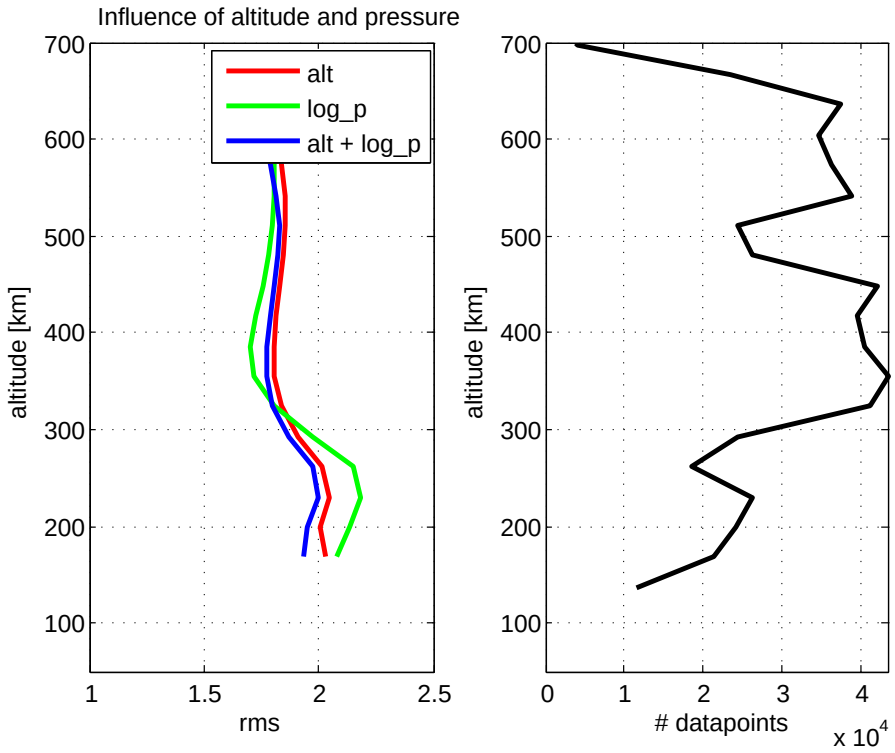


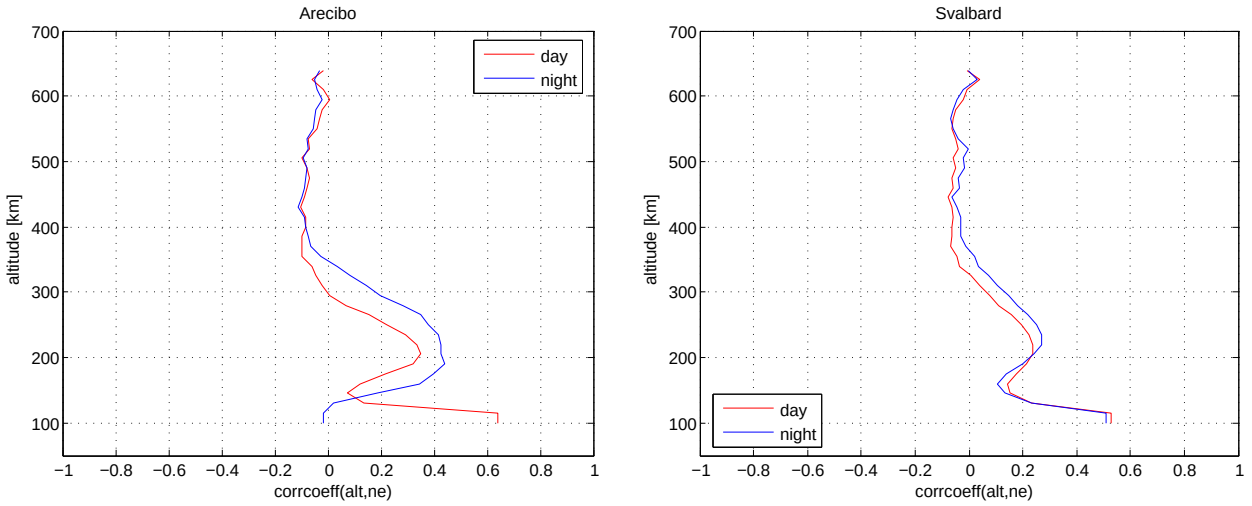
Figure 6.8 Influence of altitude and pressure. At low altitude the parameter altitude yields better results, whereas at high altitudes pressure yields better results.

The model with both quantities (pressure and altitude) as input gives as well the best overall RMS for test patternset of 1.82 (shown in Table 6.3).

Description	**RMS**$_{all}$	**RMS**$_{train}$	**RMS**$_{val}$	**RMS**$_{test}$
alt	2.16	2.31	1.84	1.87
log_p	2.20	2.37	1.87	1.85
alt + log_p	2.17	2.34	1.85	1.82

Table 6.3 Influence of altitude and pressure

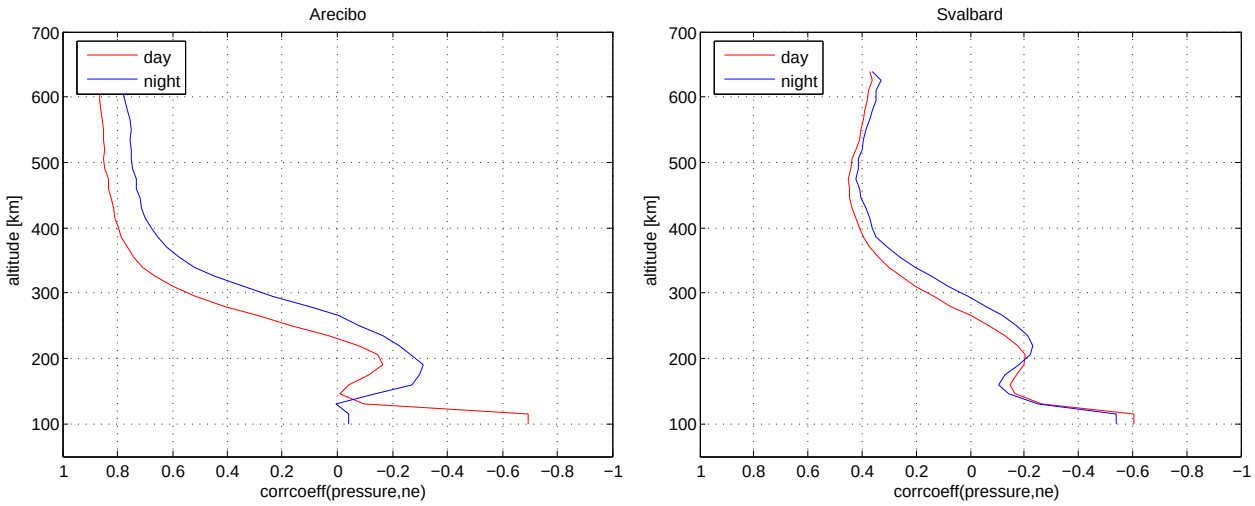
Altitude dependence on N_e for Arecibo and Svalbard.



(a) Correlation coefficient N_e(alt) day/night

(b) Correlation coefficient N_e(alt) day/night

Figure 6.9 Altitude dependence on N_e for Arecibo and Svalbard.



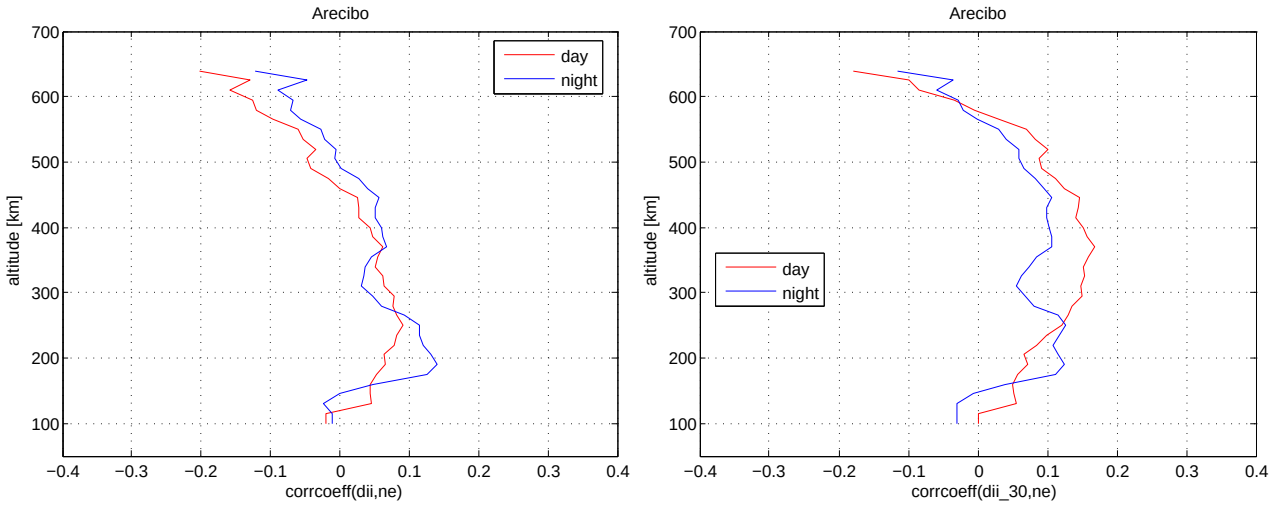
(a) Correlation coefficient N_e(pressure) day/night

(b) Correlation coefficient N_e(pressure) day/night

Figure 6.10 Air pressure dependence on N_e for Arecibo and Svalbard.

6.2.2 Seasonal Changes

If we look at seasonal variation of electron density we have the problem that we need to present the neural network a "looped" quantity. This means, that we have to find a way to "show" the neural network that the situation at the end of the year is the same as at the beginning. This is achieved by using the daily integrated insolation as input. For further information see Ch. 5.3.
The second problem is to provide information of an asymetric component to the neural network, because in autumn we have a different situation from that in spring even though there is the same daily integrated insolation. For this reason a 30 day average value of daily integrated insolation prior to the data to be predicted is used as second input parameter for seasonal variation to achieve an optimum of seasonal modelling capability. In Figure 6.11 the correlation between daily integrated insolation and N_e is shown.



(a) Correlation coefficient $N_e(dii)$ day/night

(b) Correlation coefficient $N_e(dii_{30})$ day/night

Figure 6.11 Daily integrated insolation for Arecibo.

6.2.3 Diurnal Changes

A very important influence on electron density is the Earth's rotation about its own axis. Similar to the seasonal changes of electron densities the diurnal changes demand a "looped" quantity to avoid the disadvantage of, for example, hour number. If we would take hour number as input parameter to present diurnal changes for the network it would not be able to learn, that 23.99 is nearly the same physical situation as 0.00.

Therefore the Chapman production function is used to provide the neural network an optimum of physical relevant information. As second input describing diurnal changes an eleven hour average value of the Chapman production function is used to make possible an asymmetric behavior of the model. Because the electron density distribution before sunset is different to electron density distribution after sunrise, at otherwise similar zenith angle.

In Figure 6.12 the influence of Chapman production function on electron densities in Arcibo is shown (raw data). The plot shows an increasing mean of N_e with increasing Chapman production function. This relation is very similar in Svalbard shown in Figure 6.13.

A diurnal "loop" can also be reached by using sinus and cosine [McKinnell, (2002)].

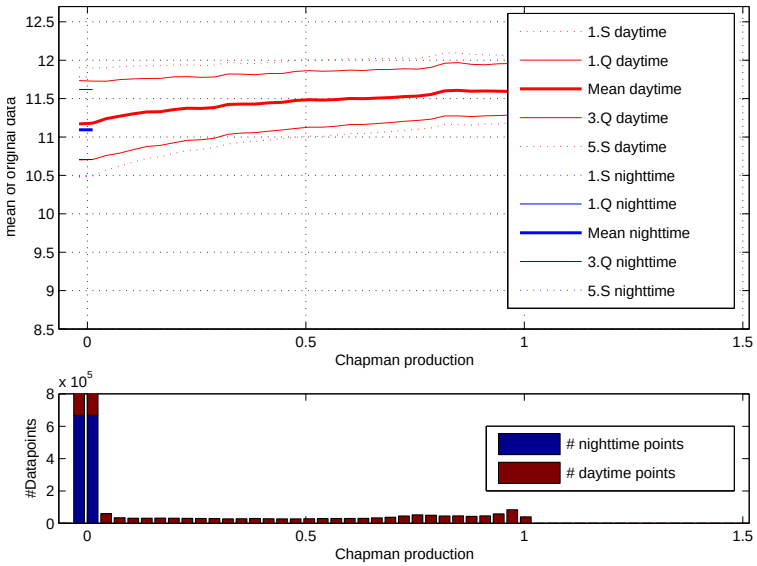


Figure 6.12 Original data from Arecibo $N_e(chza)$. A clear correlation between electron density and Chapman production is obvious. Note that there is a peak in the distribution at 0 because during nighttime Chapman production is 0.

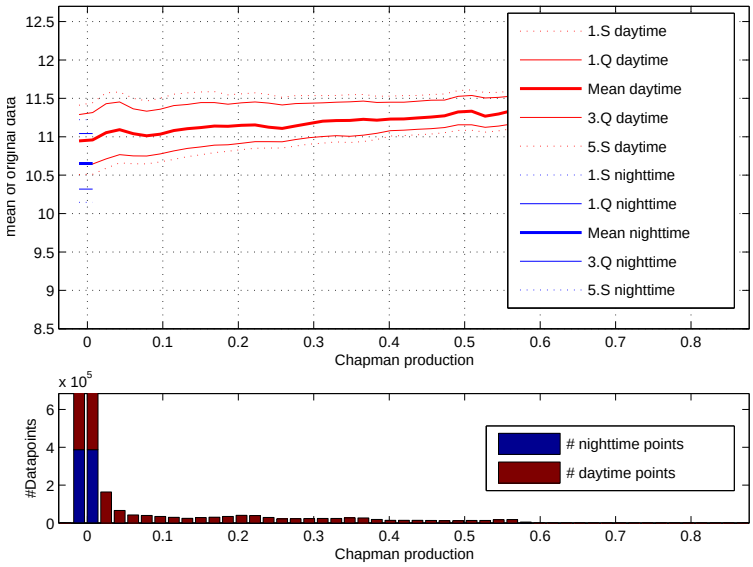
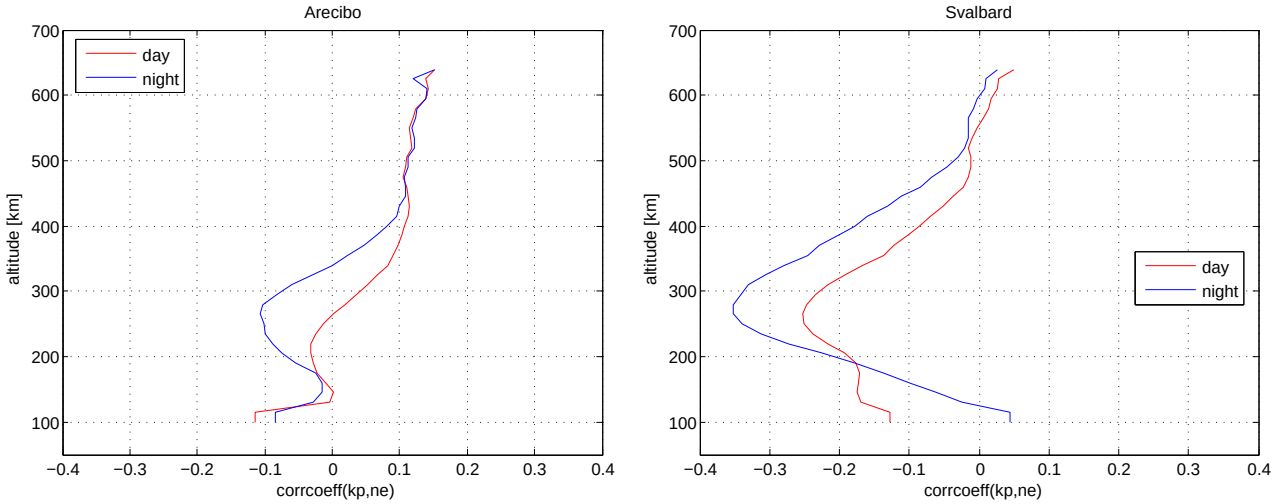


Figure 6.13 Original data from Svalbard $N_e(chza)$. As in Fig 6.12 a clear correlation between electron density and Chapman production is obvious. Note that the mean electron densities in Svalbard are lower than in Arecibo.

6.2.4 Correlation with kp, ap, Dst

In Figures 6.14 and 6.15 the correlation of N_e and kp respectively Dst is shown for Arecibo and Svalbard. As an input parameter for the neural networks ap is used.

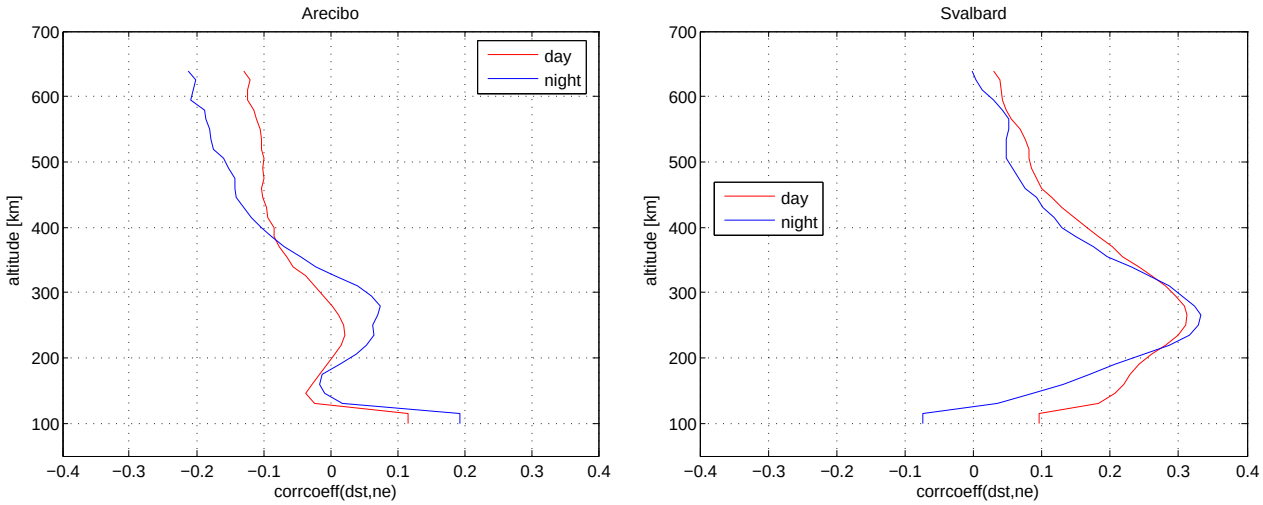


(a) Correlation coefficient $N_e(kp)$ day/night

(b) Correlation coefficient $N_e(kp)$ day/night

Figure 6.14 Geomagnetic index kp for Arecibo and Svalbard. Apparently at high latitudes kp is a more important parameter than near the equator.

Due to the high latitude of around 79° geomagnetic influence in Svalbard is dominating.



(a) Correlation coefficient $N_e(dst)$ day/night

(b) Correlation coefficient $N_e(dst)$ day/night

Figure 6.15 As Fig. 6.14, but for the index Dst.

6.2.5 Best Lead Time for the Solar Flux Index F107

Due to the longer lifetime of free electrons in higher regions of the ionosphere it is useful to add a timedependent input parameter. A first approach is to use a running mean of the solar flux preceding the time the data were taken.
To determine the best timespan of calculating the average different times were used. Several neural networks were trained with different average calculation times. An overview is given in Table 6.4.

As shown in column RMS_{test} a 60 days average value is taken in addition to the actual solar flux value as input. An optimum of 60 days average time was also found in [McKinnell, (2002)].

Description	$\mathbf{RMS}_{all}$	$\mathbf{RMS}_{train}$	$\mathbf{RMS}_{val}$	$\mathbf{RMS}_{test}$
0 days	1.99	1.99	1.99	1.99
5 days	1.99	1.99	2.00	2.00
15 days	1.99	1.98	2.01	1.98
30 days	1.97	1.97	1.99	1.95
60 days	1.97	1.97	1.98	1.94
120 days	1.98	1.97	2.00	1.99

Table 6.4 Influence of running mean of $F107$ (Arecibo).

In Figure 6.16 the dependence of the RMS on altitude is displayed. At higher altitudes longer lead times would be better.

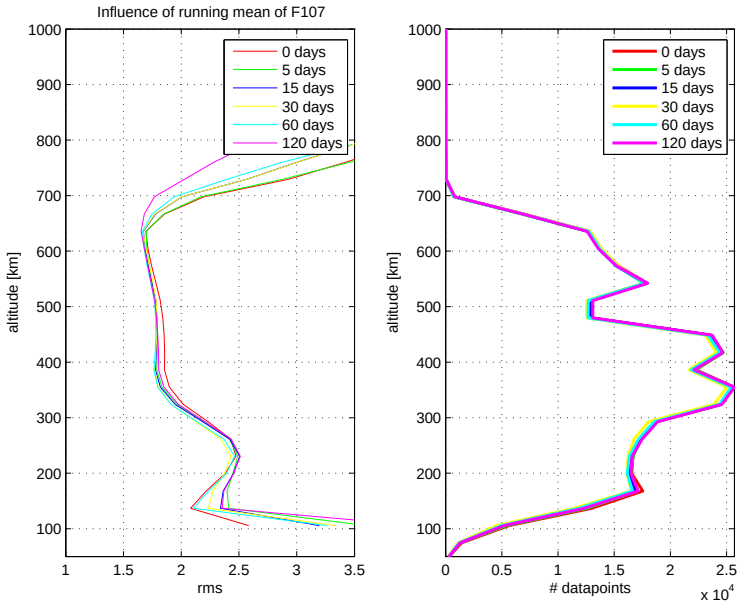


Figure 6.16 Influence of running mean of $F107$ (Arecibo).

The Figure 6.17 is a visualization of Table 6.4. An optimum RMS can be seen for 60 days.

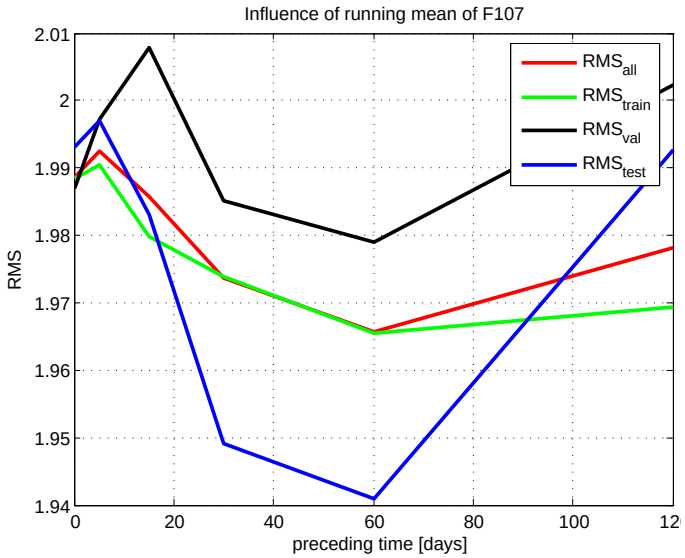
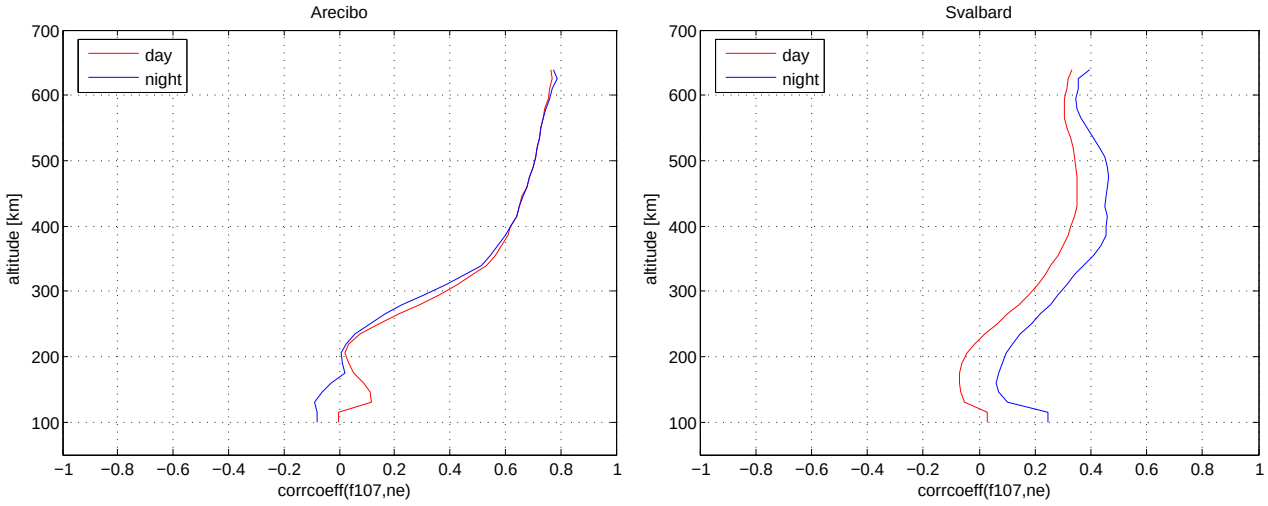


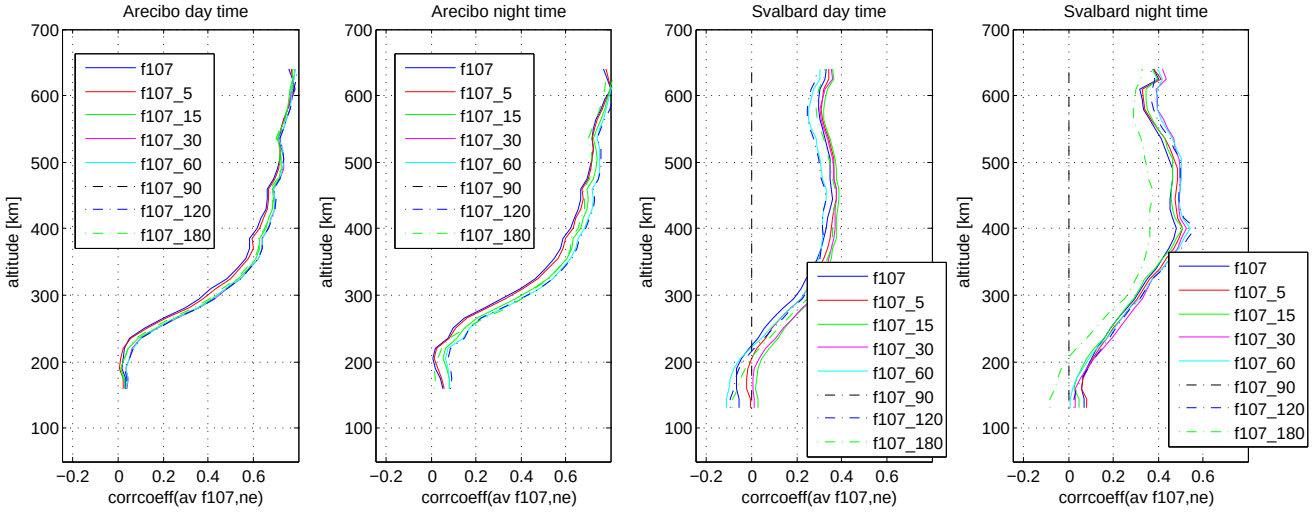
Figure 6.17 Influence of running mean lead time of $F107$ (Arecibo).



(a) Correlation coefficient $N_e(F107)$ day/night

(b) Correlation coefficient $N_e(F107)$ day/night

Figure 6.18 $F107$ Solar Flux for Arecibo and Svalbard



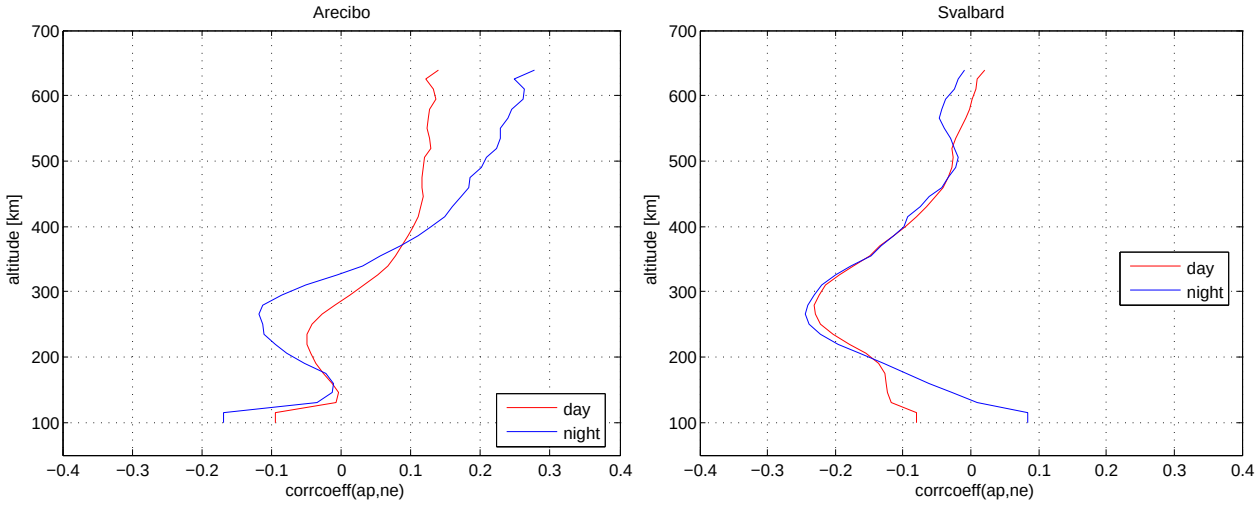
(a) Correlation coefficient $N_e(F107)$ x day average of $F107$

(b) Correlation coefficient $N_e(F107)$ x day average of $F107$

Figure 6.19 $F107$ Solar Flux for Arecibo and Svalbard with x day average.

6.2.6 Best Running Mean Lead Time for the Geomagnetic Index ap

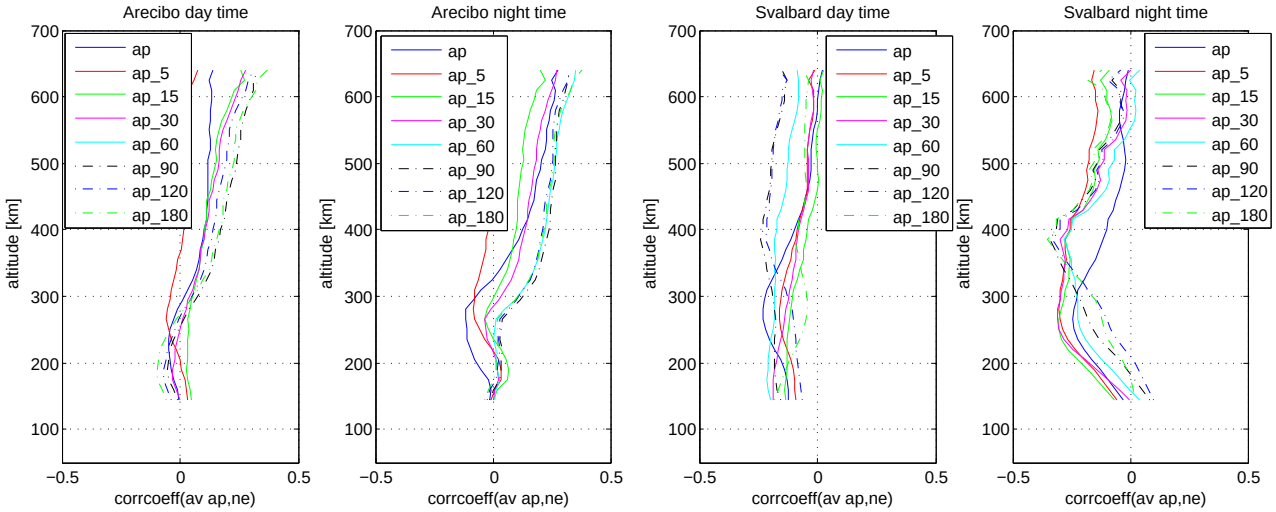
Due to lower and less distinct dependence on geomagnetic index ap it is difficult to decide which averaging time is best. Therefore 0 days and 60 days was taken which provide a good compromise between short time and long time dependence.



(a) Correlation coefficient $N_e(ap)$ day/night

(b) Correlation coefficient $N_e(ap)$ day/night

Figure 6.20 Geomagnetic index *ap* from Arecibo and Svalbard.



(a) Correlation coefficient $N_e(ap)$ x day average of ap (b) Correlation coefficient $N_e(ap)$ x day average of ap

Figure 6.21 Geomagnetic index ap from Arecibo and Svalbard with x day average.

6.3 Dummy Data

One problem of developing an empirical model is that measurement errors are not known. The empirical model should "learn" the physical relation, but due to measurement errors it is possible that contradictory data is presented to the neural network in the training phase.
Therefore a little trick is needed to test how the neural network would learn the relations between input parameters from clean data. In the present context clean means that there are no contradicting data and statistical spreading. IRI (2007) output is used to train the neural network. Instead of measured data "synthetic" data from a model are used.

As we see in Fig. 6.22, the RMS stays approximately at 1.7 over the years. It is also to mention that this is an indicator for the phenomenon of a quality improvement in Arecibo data from 1965 to 1995 as seen in Fig. 6.22.

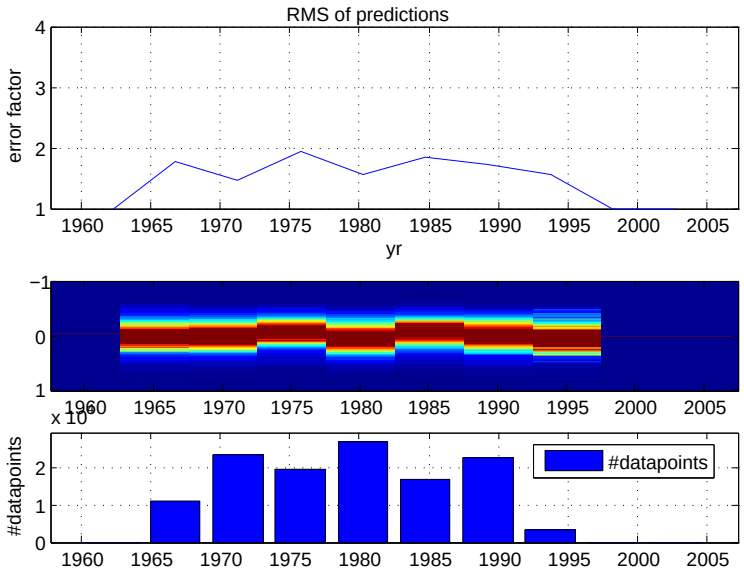


Figure 6.22 RMS diagram for neural network trained with dummy data (IRI model, Arecibo).

6.4 Conclusion

In this chapter the choice of the most suitable input parameters for the neural network is given. To obtain height data, altitude and air pressure is used. For diurnal changes the Chapman production and an 11 hour average value of Chapman production is used. To ensure information about seasonal changes the daily integrated insolation and a 30 day average was used. Solar activity is presented to the neural network by both the actual solar flux and a 60 day average preceding the time in question. And finally the geomagnetic index and a 60 day average is used to describe disturbances in Earth's magnetic field.
In sum these are 10 parameters used as inputs for the neural network which cover all important factors for electron densities in the ionosphere. An visualisation of the architecture is shown in Figure 6.3.

1. Chapman production
2. Chapman production 11 hour average

3. $F107$ solar flux
4. $F107$ solar flux 60 day average
5. Magnetic index ap
6. Magnetic index ap 60 day average
7. Altitude
8. Pressure
9. Daily integrated insolation
10. Daily integrated insolation 30 day average

7 Results of the Arecibo Model

In the first part of the results the developed model for Arecibo will be discussed. Figure 7.1 shows the mean characteristics of electron density as a function of altitude for day- and nighttime.

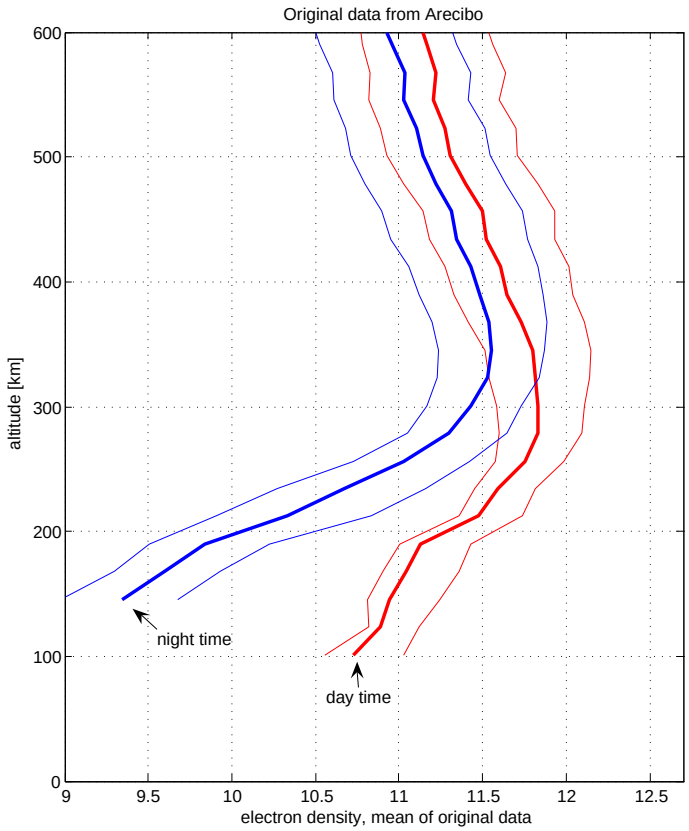


Figure 7.1 Logarithmic mean of the raw data electron densities in 20 km bins (Arecibo).

To illustrate the "learning" capability of the neural network, the learning- and output electron density values are compared in Figure 7.2. Similar plots can be seen in [Minow, (2004)].

If we compare all datapoints used for testing (youngest 10 %) with the model's prediction, we get an RMS_{test} of 1.74 ($\mathrm{RMS}_{test,IRI} = 1.95$). If we compare all datapoints with the model's prediction, we get an RMS of 2.09 ($\mathrm{RMS}_{IRI} = 2.20$).

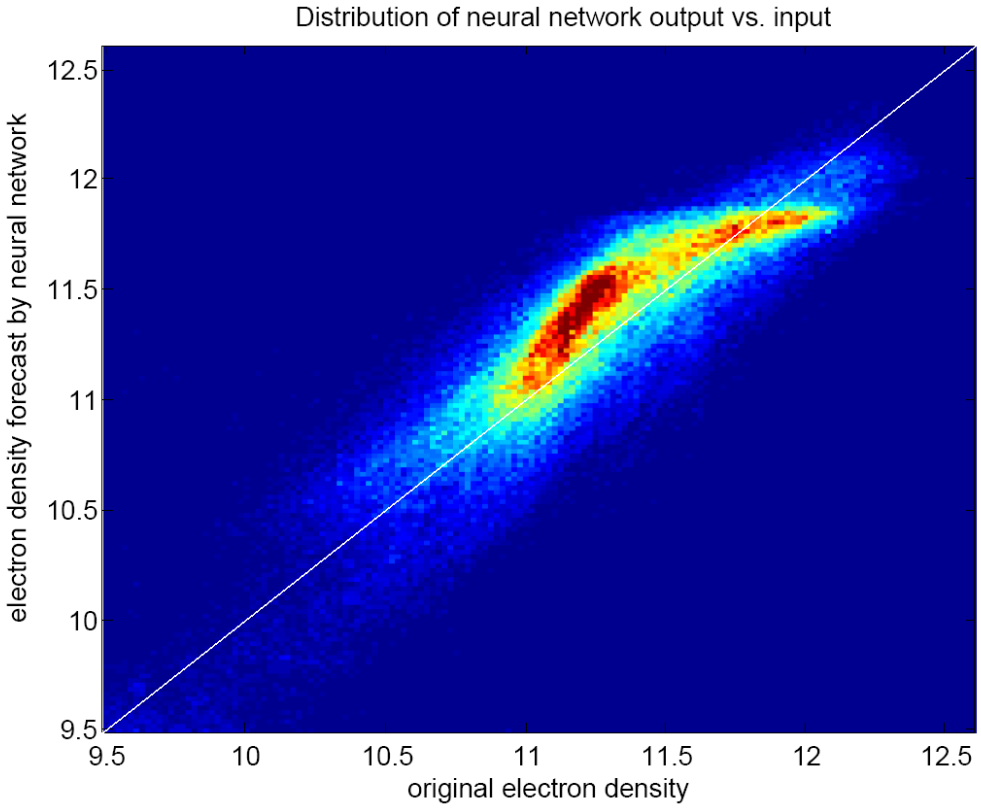


Figure 7.2 Distribution of neural network output vs. input (electron densities, Arecibo).

At higher electron densities it can be seen that the neural network provides too small values for the majority of the data. But the neural network is capable to "learn" such high electron densities which can be seen at the light blue areas.

7.1 Charts from model output

At first all parameters are fixed instead of the hour of the day [UTC]. In Figures 7.3, 7.8 and 7.9 the plots for several altitudes at spring equinox, midwinter and midsummer are shown.
The electron density is increasing steeply when sun rises and decreases again starting at sunset. Beyond the zenith angle and the daily integrated insolation is plotted. The parameters *ap* and solar flux are fixed to 17.2 and 130 Jy unless indicated otherwise.

In the model for Arecibo electron density profiles are extended by 5 datapoints from the empirical model FIRI when the lowest datapoint is higher than 150 km. With the extended FIRI data the model has a slightly better RMS error due to more stability at lower altitudes. Further details about the FIRI model can be found in [Friedrich and Torkar, (2001)].

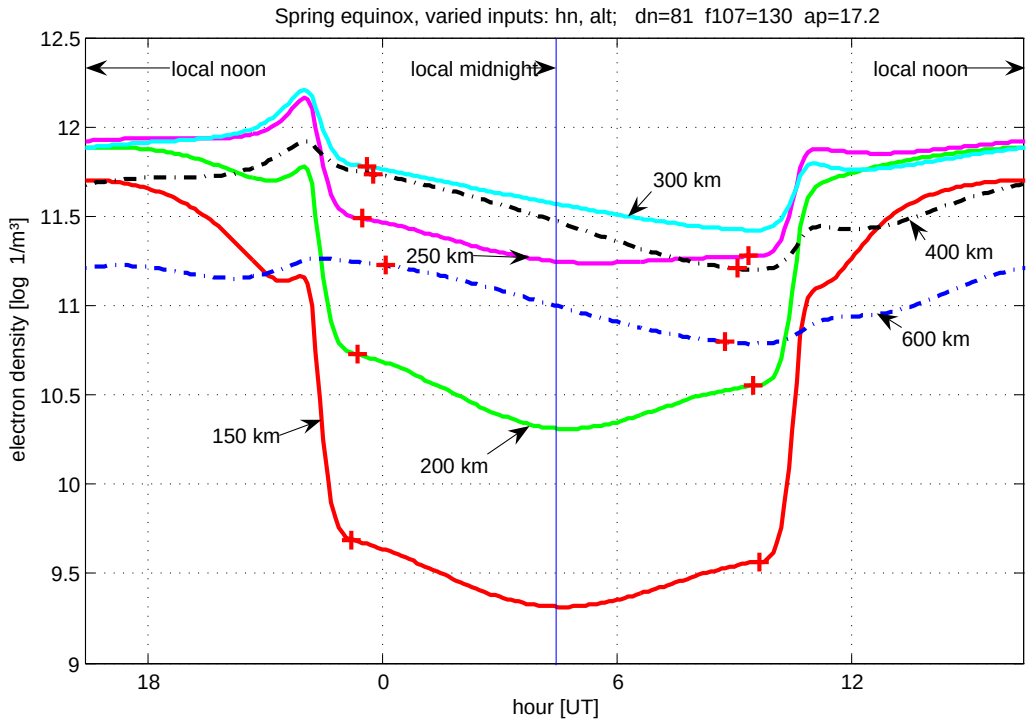


Figure 7.3 Arecibo, Electron densities at various altitudes during a day, centered at local midnight. With FIRI data, Spring equinox, dummy data (model output). Red crosses indicate the day-night border according to Ch. 5.7.

In Fig. 7.4 a dayplot of an model is shown were no FIRI data were added to the training data, Arecibo. Note that all other plots from Arecibo model here are with FIRI in training data.

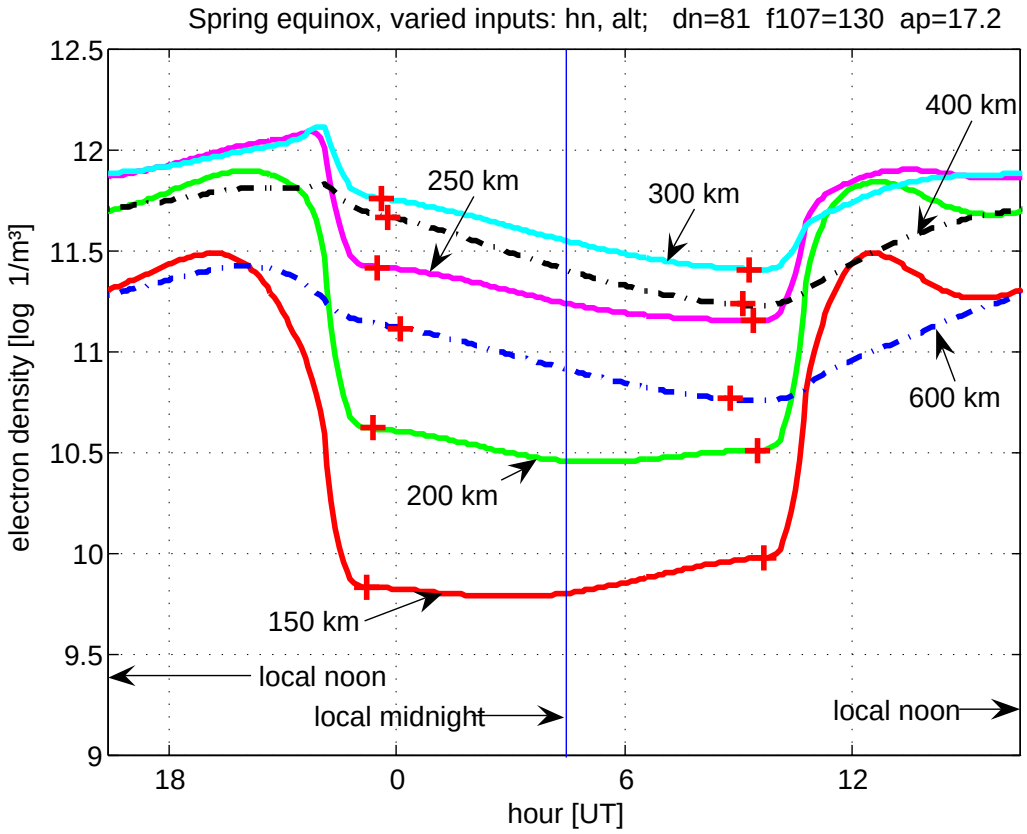


Figure 7.4 Arecibo, Electron densities at various altitudes during a day, centered at local midnight. Without FIRI data, Spring equinox, dummy data (model output). Note the unrealistic behaviour at the lowest altitudes which are only insufficiently covered by the IS radar. Red crosses indicate the day-night border according to Ch. 5.7.

In Fig. 7.5 a dayplot of an IRI 2007 is shown ($F107 = 168$ Jy). At 100 km electron density is limited to $9.510^6\ m^{-3}$. Remarkable is a increase of electron density at night in 200 and 250 km. This is apparently a bug in IRI 2007 software.

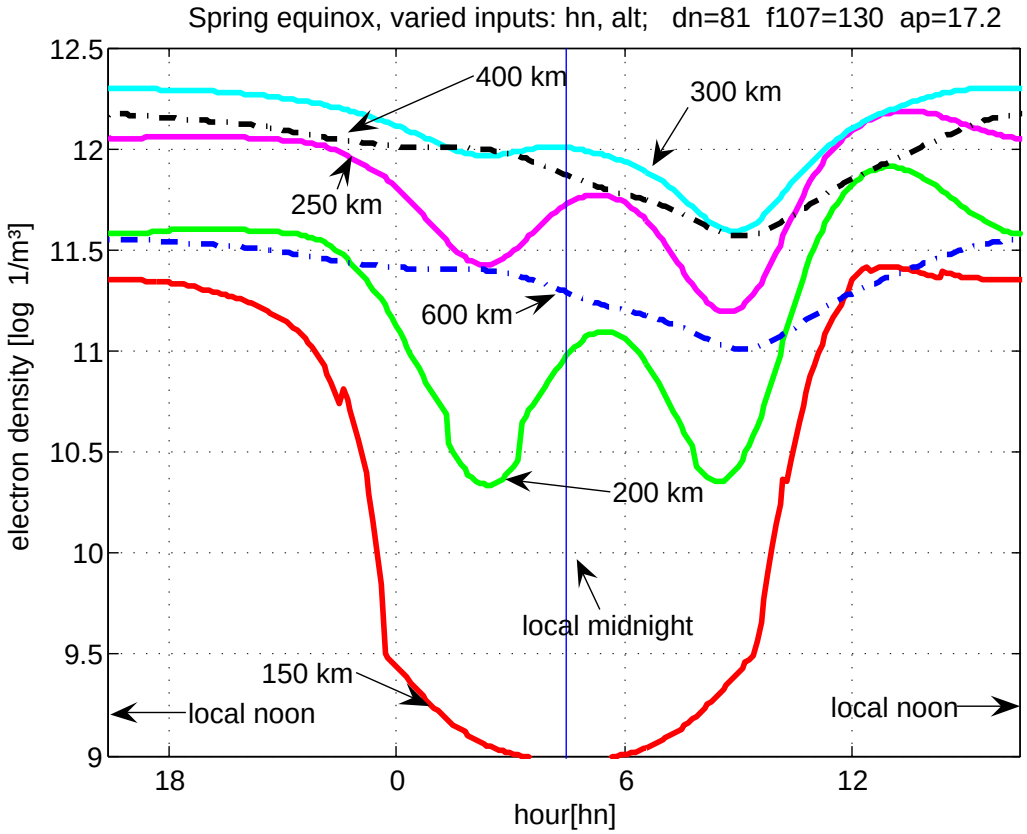


Figure 7.5 IRI 2007 output for spring day ($F107 = 168$ Jy).

These density peaks increase with lower latitude. At higher altitudes IRI shows a similar traces as the Arecibo model but with slightly higher electron densities.

In Figures 7.6 (including FIRI) and 7.7 (without FIRI) electron density profiles for different zenith angles are shown. With FIRI data the model produces lower electron densities at around 150 km where FIRI data was added. Note that all other plots shown from Arecibo model are extended by FIRI data due to more stability at lower altitudes.

Without FIRI data the neural network "learns" mainly electron densities where data coverage is good, and delivers poor results at the edge of the data range.

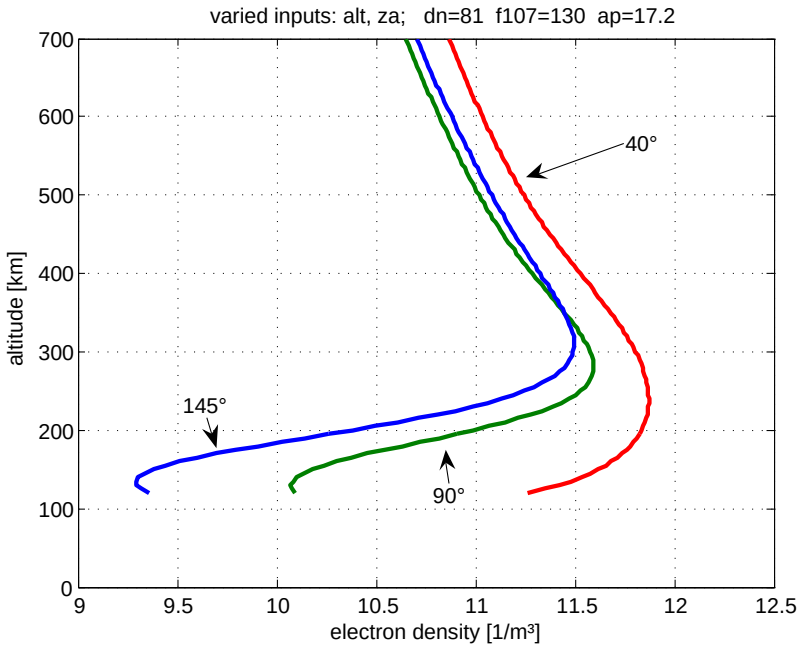


Figure 7.6 Arecibo, different zenith angles, including FIRI data (lower electron densities at around 150 km), AM (model output).

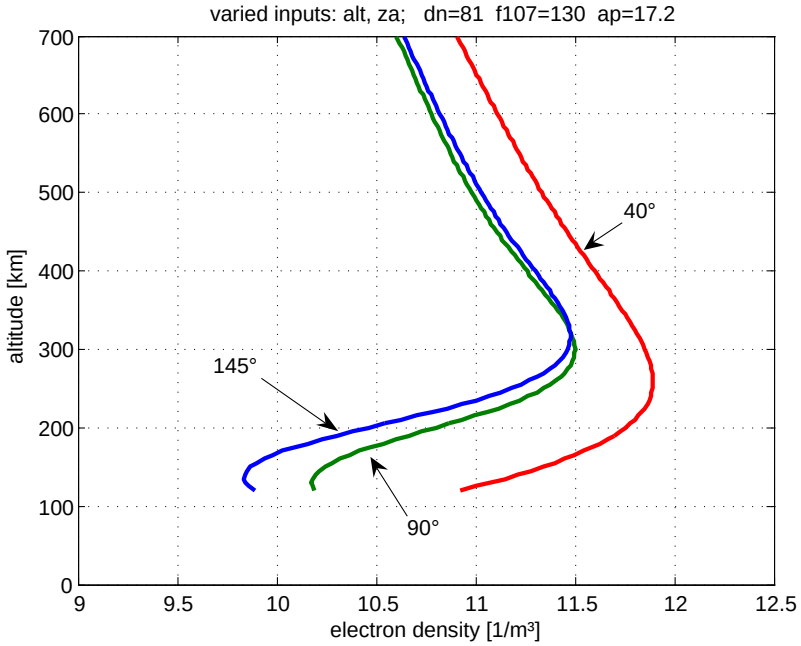


Figure 7.7 Arecibo, different zenith angles, without FIRI data, AM (model output). At the lowest altitudes results are biased to the instruments tresholds.

In Figures 7.8 and 7.8 the diurnal variation at Arecibo is shown for midwinter and midsummer. The electron desities corresponds to Chapman production. The difference between midsummer is not very strong because Arecibo's low latitude. The daily integrated insolation only varies between 0.22 (midwinter) and 0.33 (midsummer).

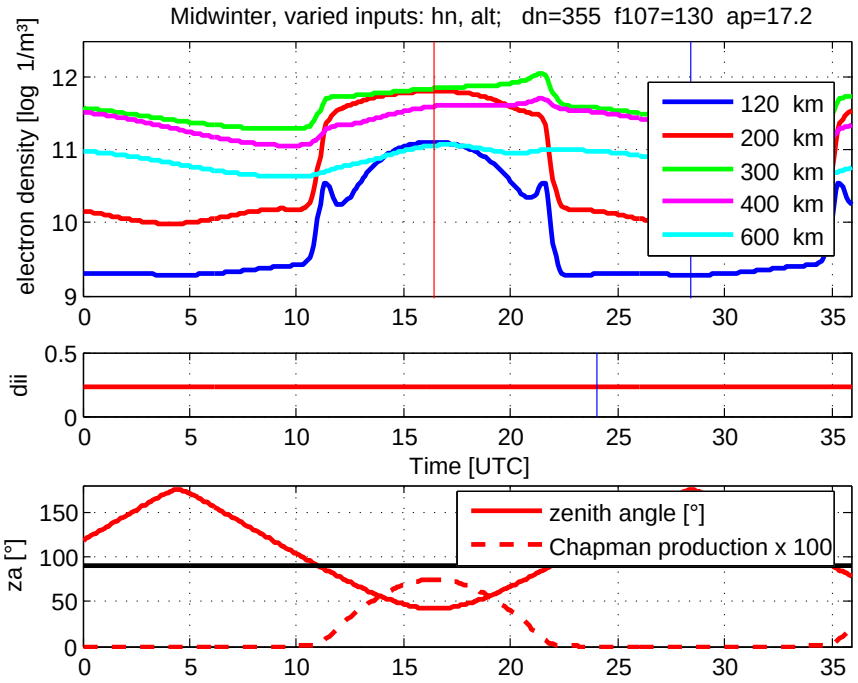


Figure 7.8 Arecibo, diurnal variation at midwinter (model output). Upper panel: electron densities, central panel: daily integrated insolation (DII), bottom panel: Solar zenith angle and Chapman production.

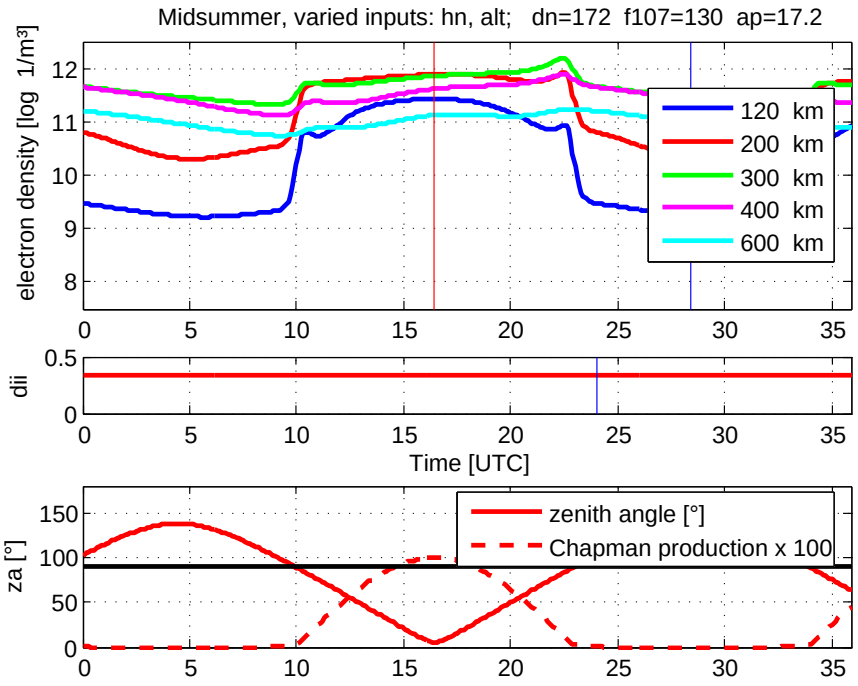


Figure 7.9 Arecibo, as Fig. 7.8, but for a midsummer day.

The next variant is to vary the day of the year to get an idea of seasonal variations in Arecibo. In Figure 7.10 and 7.11 electron densities for several altitudes at noon and midnight are plotted. Seasonal variations are not very pronounced due to the low latitude.

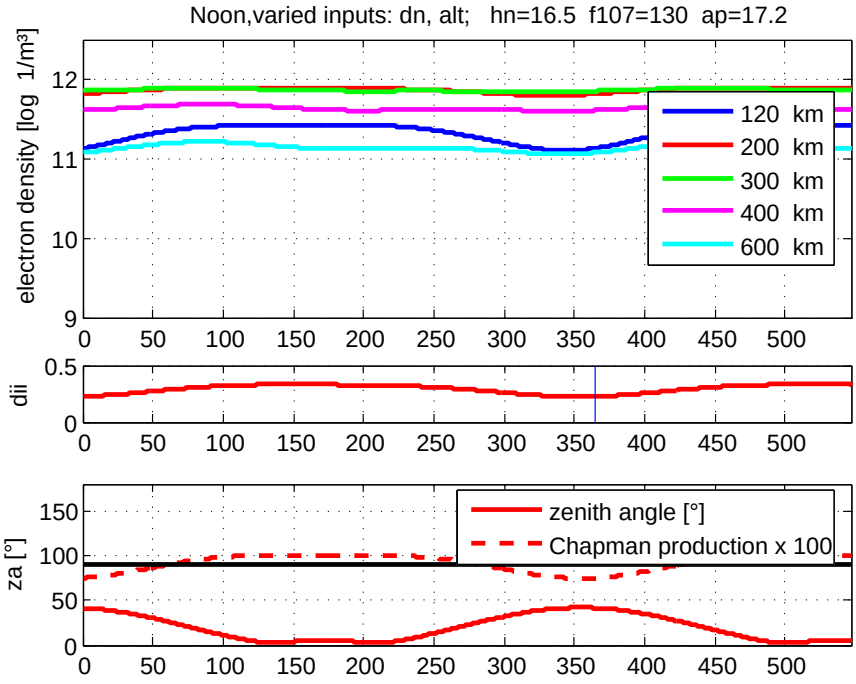


Figure 7.10 Arecibo, seasonal variation of noon electron densities for median geophysical conditions. Upper panel: electron densities, central panel: Daily integrated insolation (DII), bottom panel: Solar zenith angle and Chapman production.

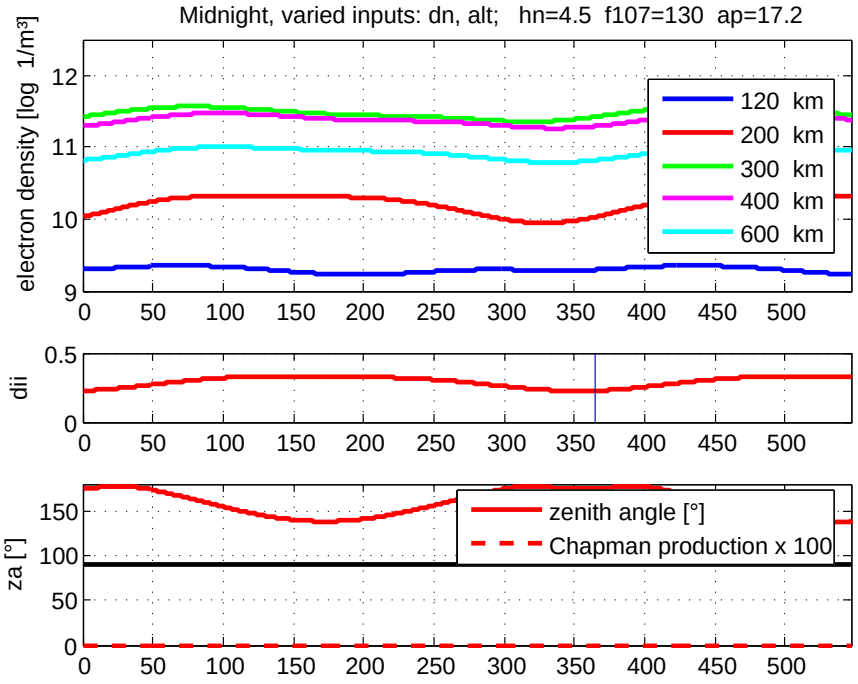


Figure 7.11 Arecibo, as Fig. 7.10, but for midnight.

At 200 km the solar zenith angle is always greater than 130°. Daily integrated insolation has more impact at night than the zenith angle.

The solar activiy expressed by solar flux is very high. In Figures 7.12 and 7.12 electron density profiles are plotted for several solar fluxes (80, 130 and 180 Jy) at spring equinox at noon and midnight.

The influence is as yet unresolved , but confirmed by an other model based on global foF2 measurements ([Friedrich et al. (2007)]). In Figure 7.14 a plot is shown where $F107$ influence is compared with a model from E. Oyeyemi for midnight and noon ([Oyeyemi, (2005)]).

But also at daytime there is an influence at altitudes higher than about 300 km. Plots for midsummer, autumnal equinox and midwinter can be found in appendix starting from Figure 11.1.

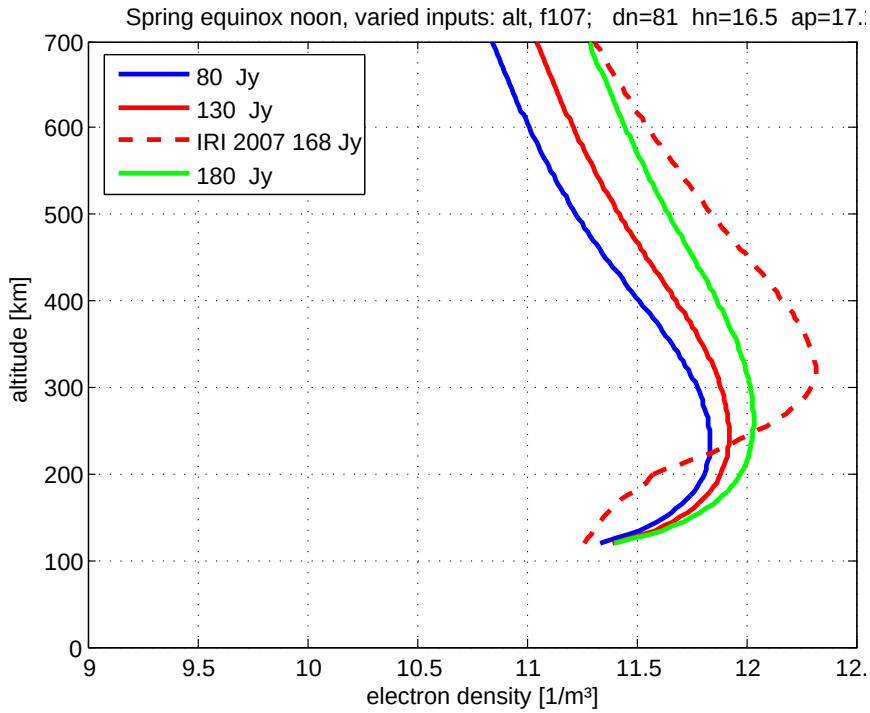


Figure 7.12 Arecibo, solar activity dependence of noon electron densities at spring equinox. IRI which is shown for comparison but for one solar activity only (due to a software problem intentionally entering other solar activities is not possible).

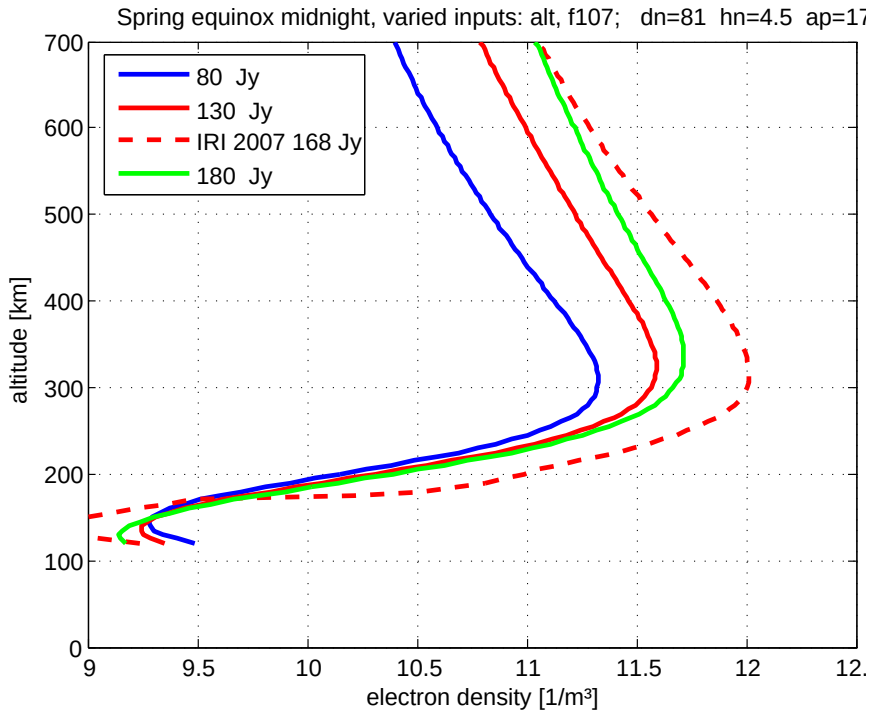


Figure 7.13 Arecibo, as Fig. 7.12, but for midnight.

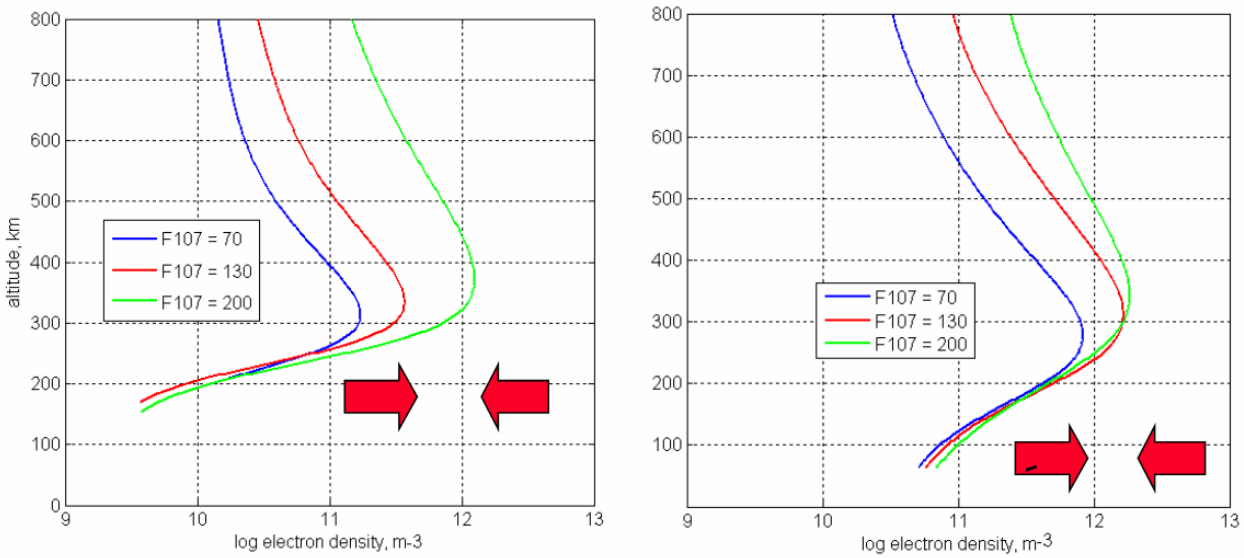


Figure 7.14 At night there is more electron density variation at different solar activity. Red arrows indicate the corresponding range of electron density in model by E. Oyeyemi ([Oyeyemi, (2005)] and [Friedrich et al. (2007)]).

The reversal of the solar flux dependence near 150 km is *not* physical, but with reflects the insufficient data coverage. However a physical reversal is expected below about 70 km.

The final parameter, the geomagnetical index ap is varied. In Fig. 7.15 and 7.16 several values of ap are plotted at noon and midnight for spring equinox. The influence of ap is very small only at midnight there is a marginal dependence at high electron densities. With more ap a slightly higher electron density occurs.

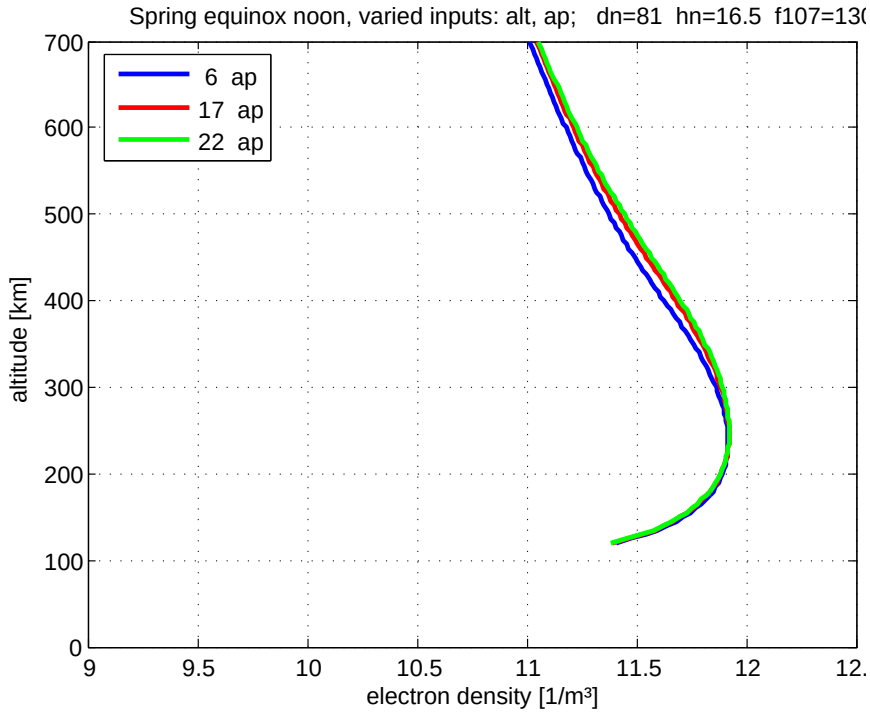


Figure 7.15 Arecibo, ap dependence of noon electron densities at spring equinox.

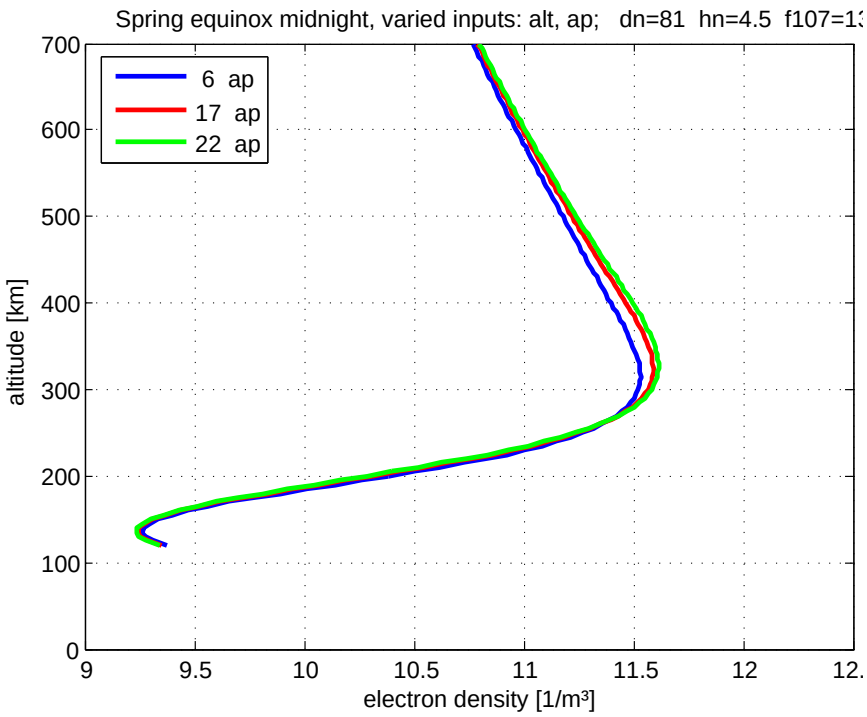


Figure 7.16 Arecibo, as Fig. 7.15, but for midnight.

7.2 Sunset Night Decay

In Figure 7.17 for every hour after sunset an electron density profile is plotted. The electron density decreases due to recombination.

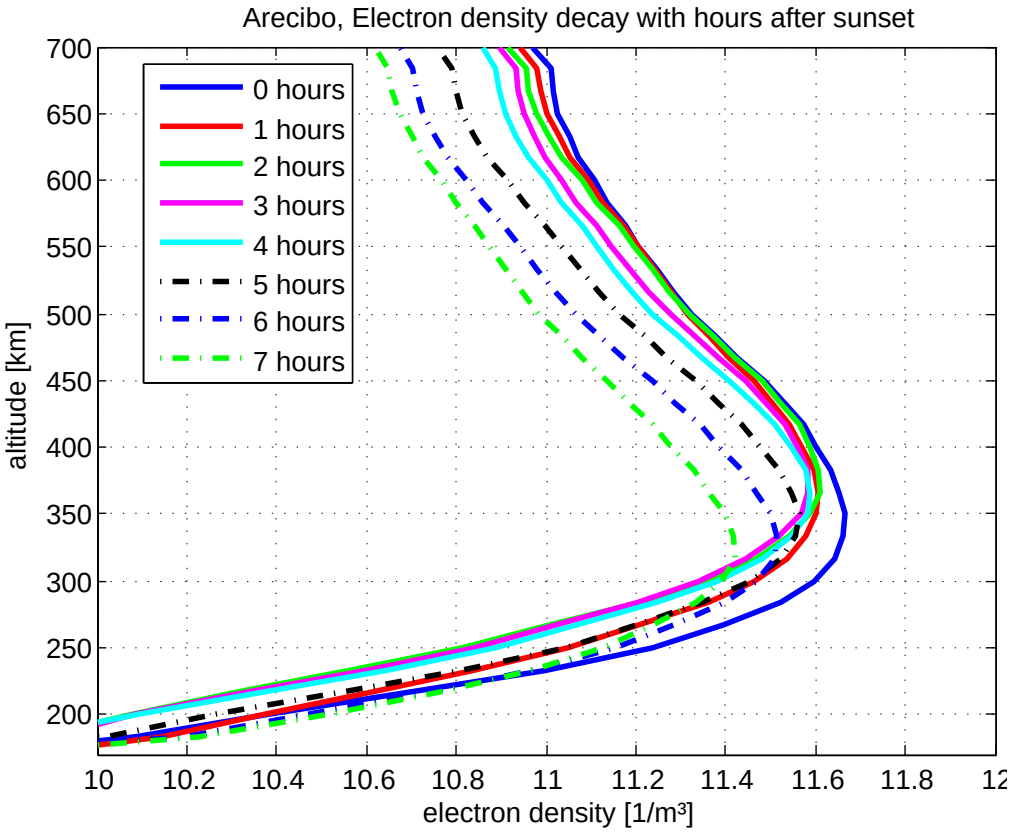


Figure 7.17 Arecibo, electron density decay with hours after sunset.

In Figures 7.18 and 7.19 it is shown how electron density decays with zenith angle and hours after sunset (exponential decay). Electron density N_e is displayed in different colors (9 to 12.5 m^{-3}). Note that the two Figures are based on measured date including all season, all ap and all $F107$.

In Fig. 7.19 electron density increases after midnight due to scattered ultraviolet radiation.

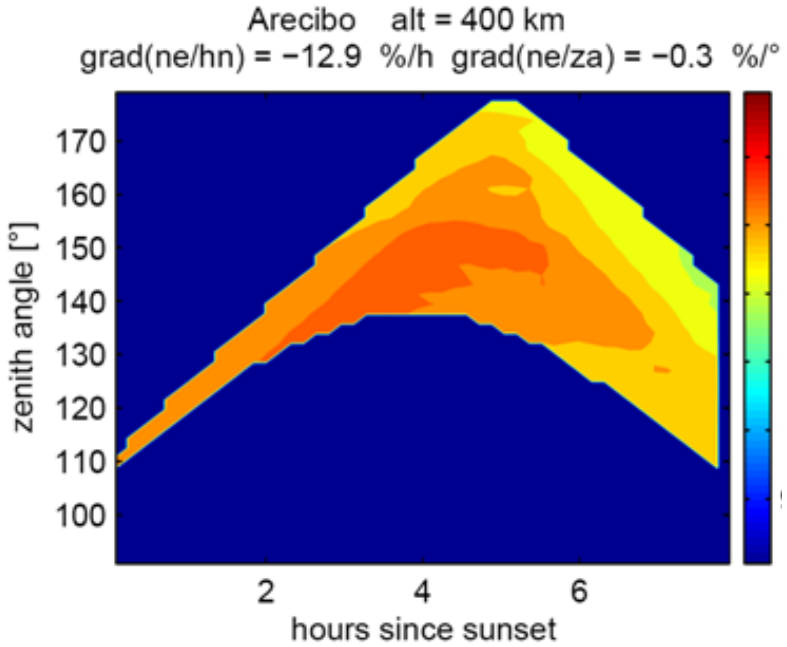


Figure 7.18 Arecibo, nocturnal electron density (400 km, measured data).

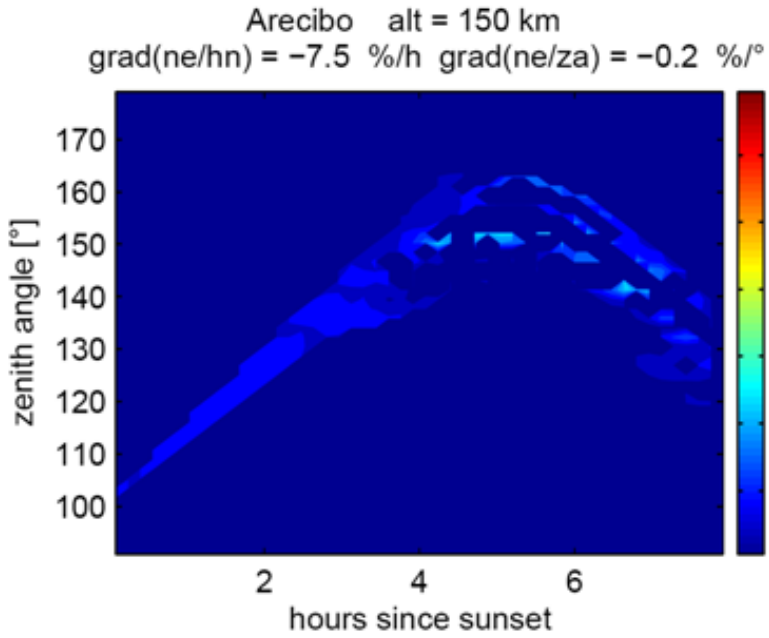


Figure 7.19 Arecibo, nocturnal electron density (150 km, measured data).

In Fig. 7.20 the model output for several times around sunset is compared with a plot from the literature [Knight, (1972)]. The Arecibo model delivers very similar results.

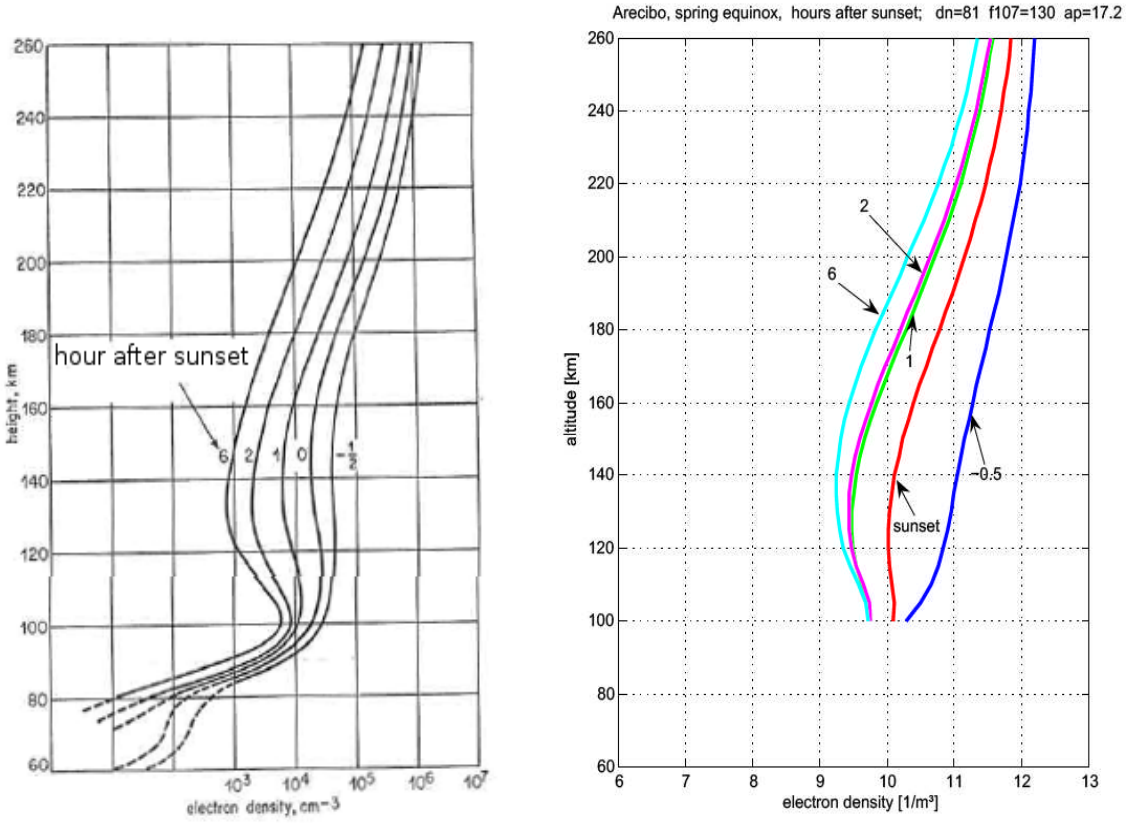


Figure 7.20 Comparison with plot from the literature, [Knight, (1972)] (left), and the Arecibo model for median conditions (right).

In Figures 7.21 and 7.22 the gradient of N_e is plotted as a function of altitude for hours after sunset and zenith angle (measured data).

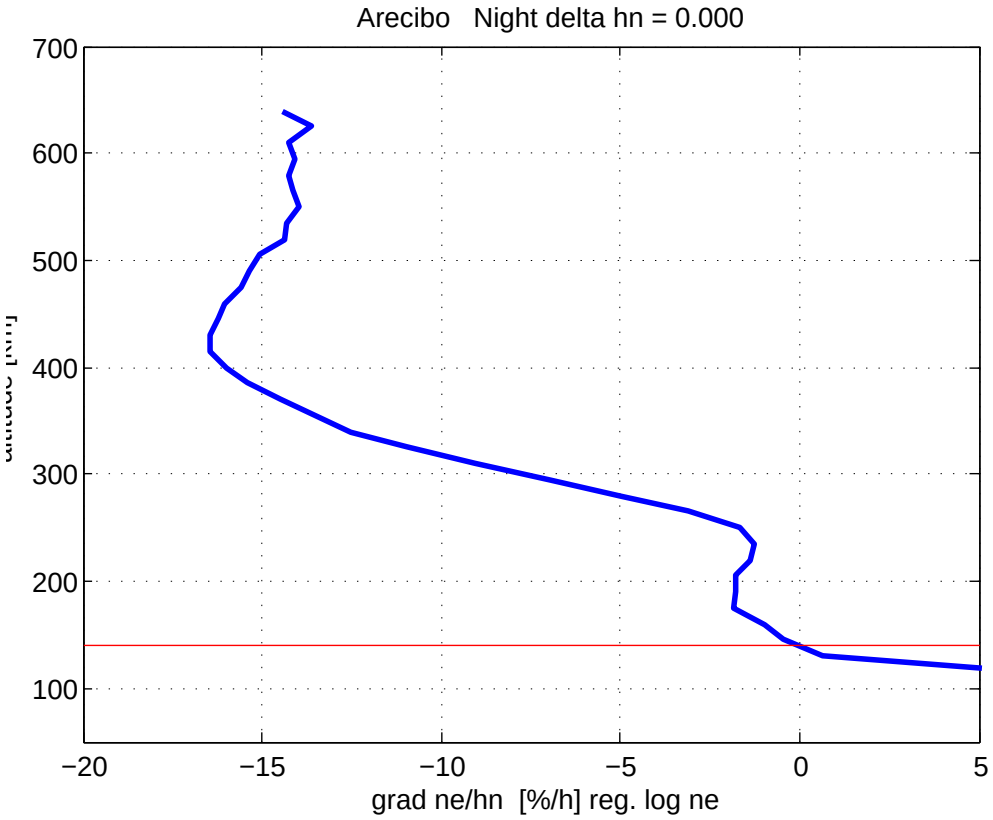


Figure 7.21 Gradient (N_e/hn), night(measured data). The red line indicates the lower quartile of the data distribution. Below data are not considered relevant.

At higher altitudes the electron density decays exponentially with about 15 % per hour. At lower altitudes electron density is predominantly controlled by scattered ultraviolet radiation at about 1 % per degree.

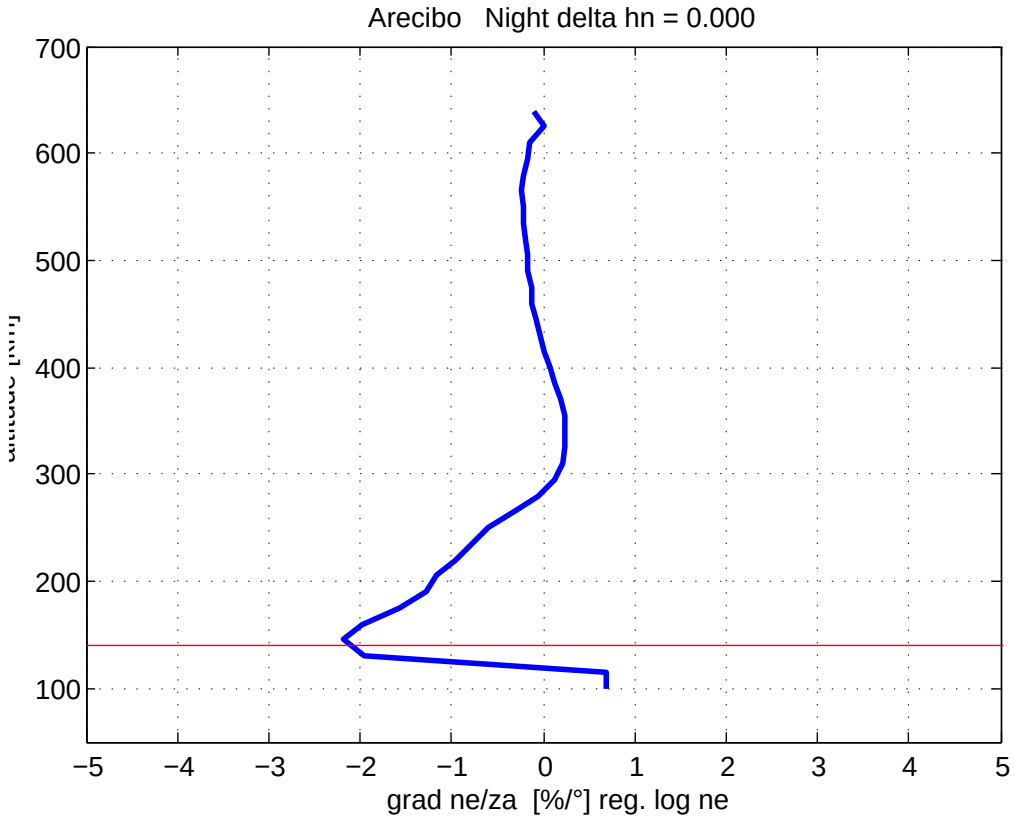


Figure 7.22 Gradient (N_e/za), night(measured data). The red line indicates the lower quartile of the data distribution. Below data are not considered relevant.

This exercise is not very successful with Svalbard data. Due to the small range of solar zenith angle and the not very pronounced transition from day to night (Fig. 8.14).

7.3 Seasonal and diurnal variations

In Figures 7.23 to 7.28 contour plots for 200, 400 and 600 km are shown. Each plot shows the electron densities for all seasonal and diurnal conditions. The plot from original data includes all ap and $F107$ for the regarding altitude layer. The model output is at mean condition. On the right side an intersection is shown for midsummer (red line) and mid-winter (blue).

The seemingly diferent shape of the contours is due to the uneven distribution of data and physical conditions.

In Figures 7.23 and 7.24 contourplots for 200 km are shown.

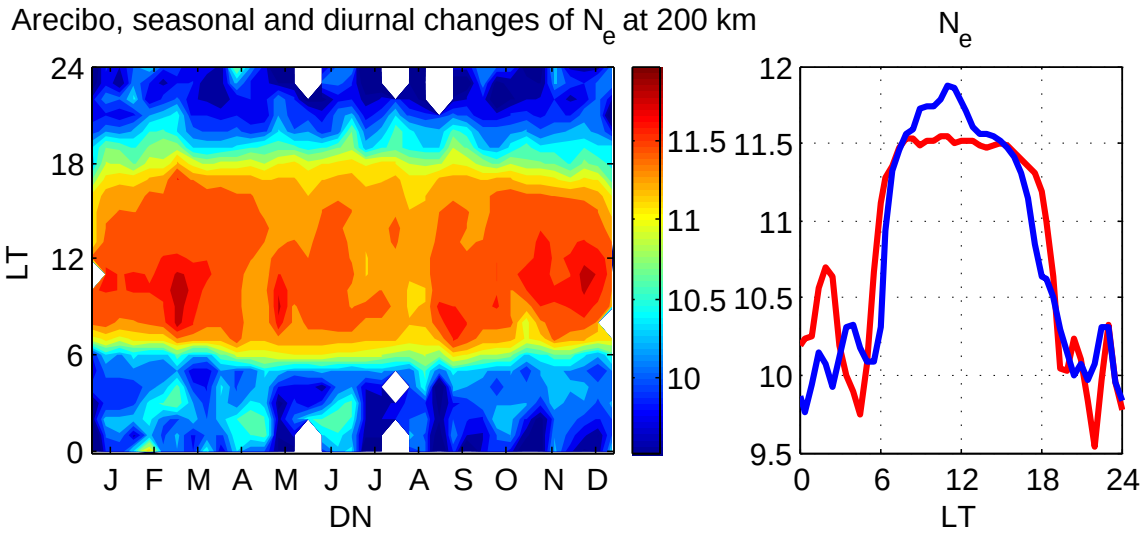


Figure 7.23 Arecibo, original data, 200 km. Lines of constant electron densities (left) and a crossection for midsummer and midwinter (right).

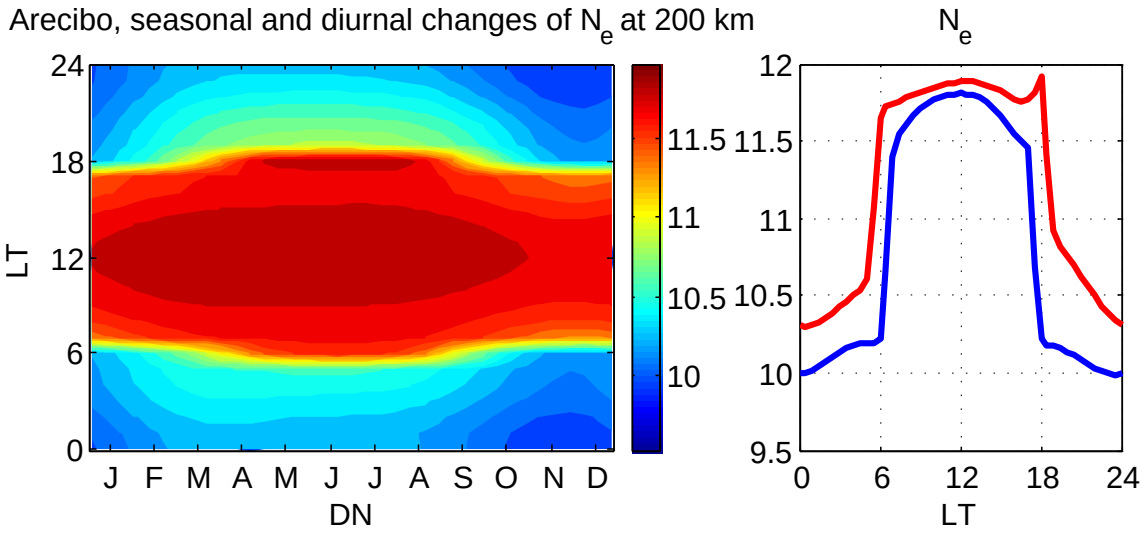


Figure 7.24 Arecibo, model data, 200 km.

In Figures 7.25 and 7.26 contourplots for 400 km are shown.

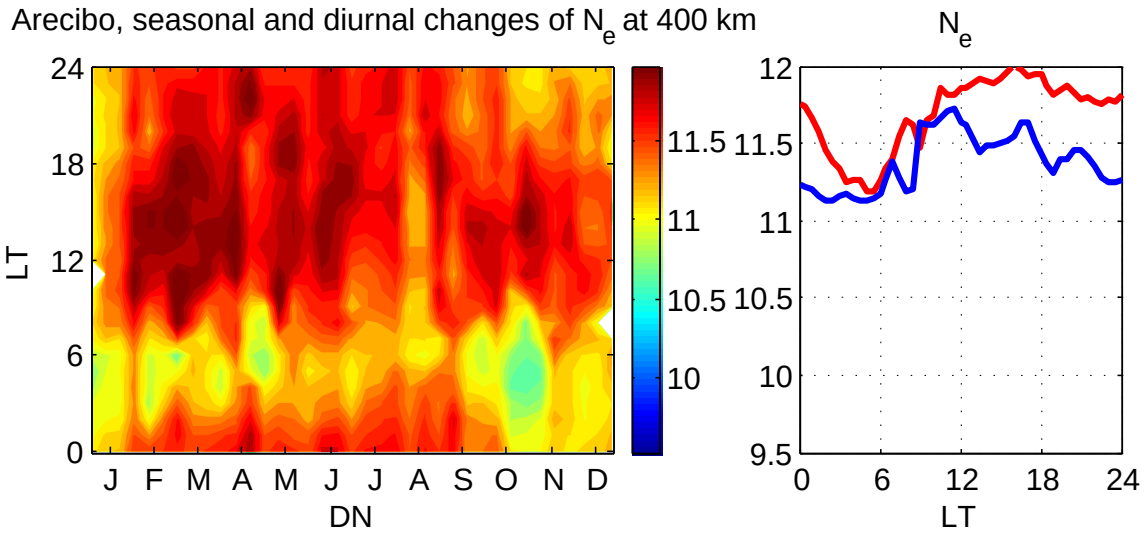


Figure 7.25 Arecibo, original data, 400 km. Lines of constant electron densities (left) and a cross section for midsummer and midwinter (right).

The seemingly diferent shape of the contours is due to the uneven distribution of data and physical conditions.

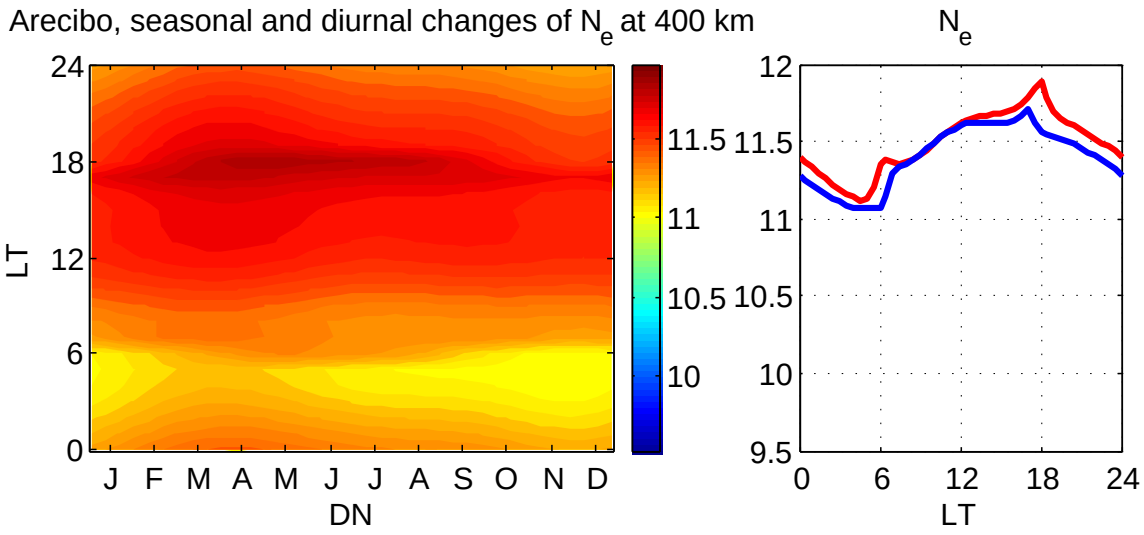


Figure 7.26 Arecibo, model data, 400 km.

In Figures 7.27 and 7.28 contourplots for 600 km are shown.

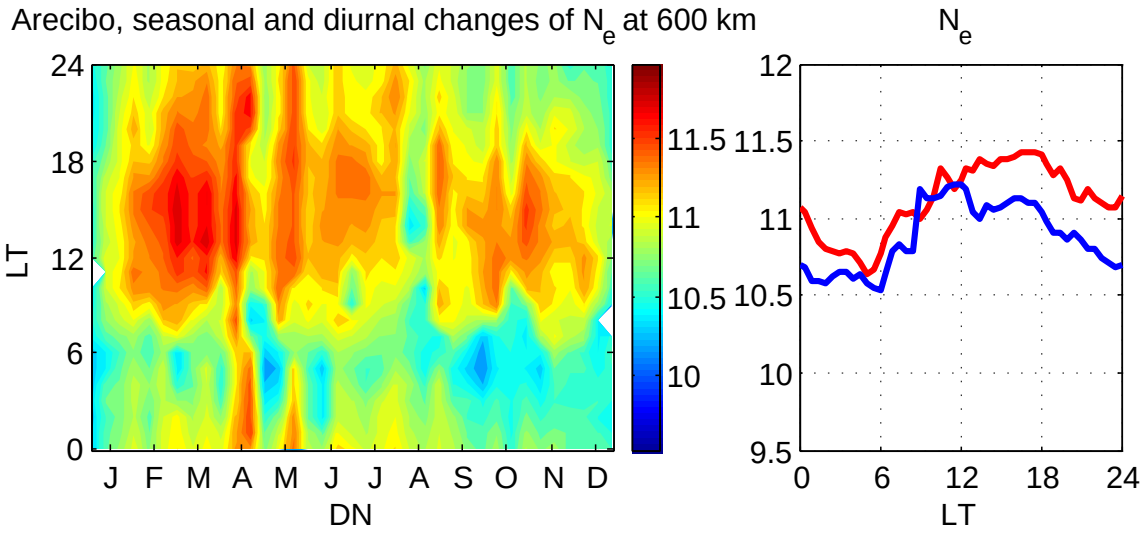


Figure 7.27 Arecibo, original data, 600 km. Lines of constant electron densities (left) and a cross section for midsummer and midwinter (right).

The seemingly diferent shape of the contours is due to the uneven distribution of data and physical conditions.

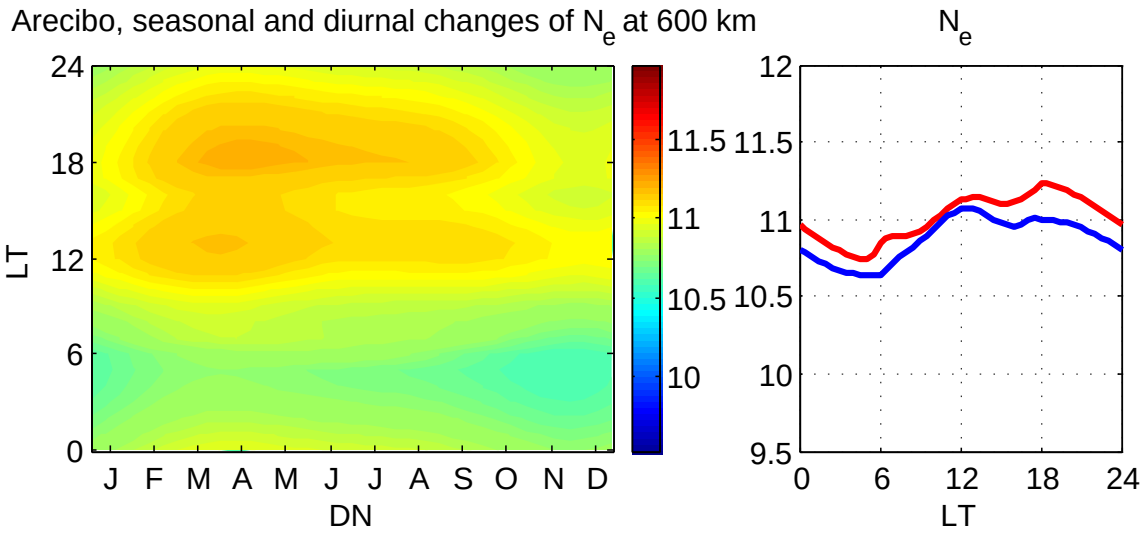


Figure 7.28 Arecibo, model data, 600 km. Note the double peak at 12 and 18 h LT most pronounced in April.

In general the model produces good outputs, sometimes higher electron densities than measured sometimes lower due to uneven data coverage. For example it can occur that in spring time there is more data at high solar activity because the incoherent scatter

radar was used for ionospheric measurements sometimes very sporadically and sometimes very often.

8 Results of the Svalbard Model

In the second part of the results we will discuss the model for Svalbard at 78° latitude. Model prediction plots will be discussed and explained.

8.1 The Ionosphere at high latitude

At high latitudes the geomagnetic field runs nearly vertical which causes a more complex ionosphere than at low- or middle latitudes. The magnetic field lines connect the ionosphere to the outer magnetosphere. The electron density at high latitude is very dynamic and controlled by the solar wind. Due to vertical magnetic field lines this region is more controlled by energetic particles emitted by the Sun.

In Figure 8.1 the dependence of electron density on altitude is plotted. All datapoints are taken and a logarithmic mean is calculated.

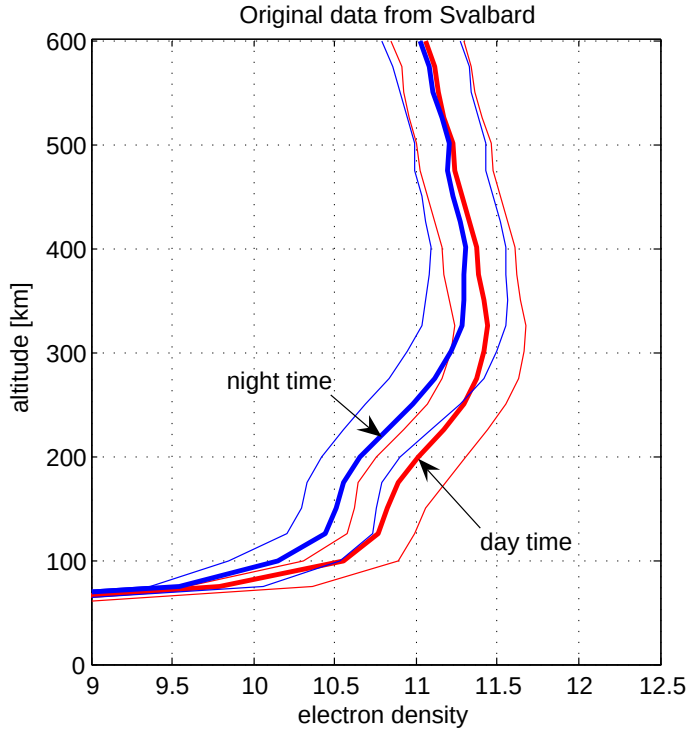


Figure 8.1 Logarithmic mean of the raw data electron densities, 20 km bins (Svalbard).

To illustrate the "learning" capability of the neural network, the learning- and output electron density values are compared in Figure 8.2.

If we compare all datapoints used for testing (youngest 10 %) with the model's prediction, we get an RMS_{test} error factor of 2.38 and for all datapoints we get an RMS_{all} error factor of 2.21.

As an example, the model for the lower ionosphere in the auroral zone (IMAZ), delivers an RMS error factor of 2.26 (IMAZ v2, [McKinnell and Friedrich, (2007)]). IMAZ is also an empirical model based on neural networks.

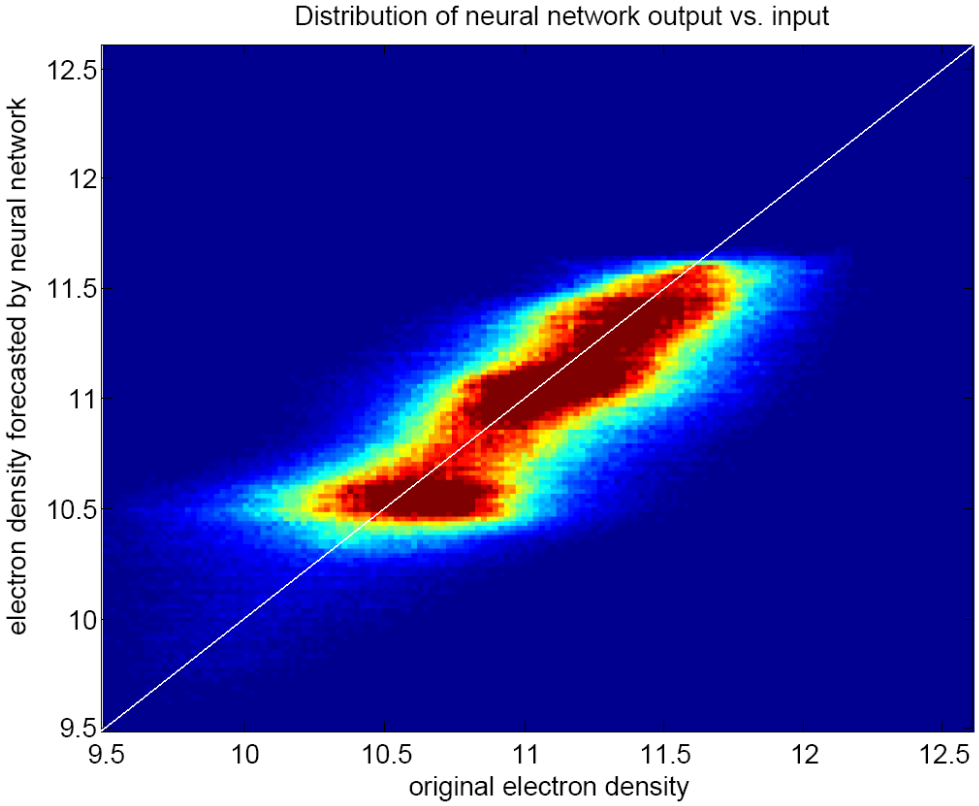


Figure 8.2 Distribution of neural network output vs. input (electron densities, Svalbard).

The predicted values appear to scatter equally below and above the input data, where there is a bias to predict too large values in the Arecibo model (Fig. 7.2). The neural network is unable to reproduce data near to the edge of the input space (e.g. data with 10^{12} are capped to $10^{11.7}$).

8.2 Charts from model output

At first all parameters are fixed except for the hour of the day [UTC]. In Figures 8.3, 8.4 and 8.5 the plots for several altitudes at spring equinox, midwinter and midsummer are shown.

Because of the slow sunrise/sunset at high latitudes the electron density increases gradually when the sun rises and decreases again starting at sunset. Beyond the zenith angle and the daily integrated insolation is plotted. The parameters *ap* and solar flux are fixed to 13 and 175 Jy unless indicated otherwise.

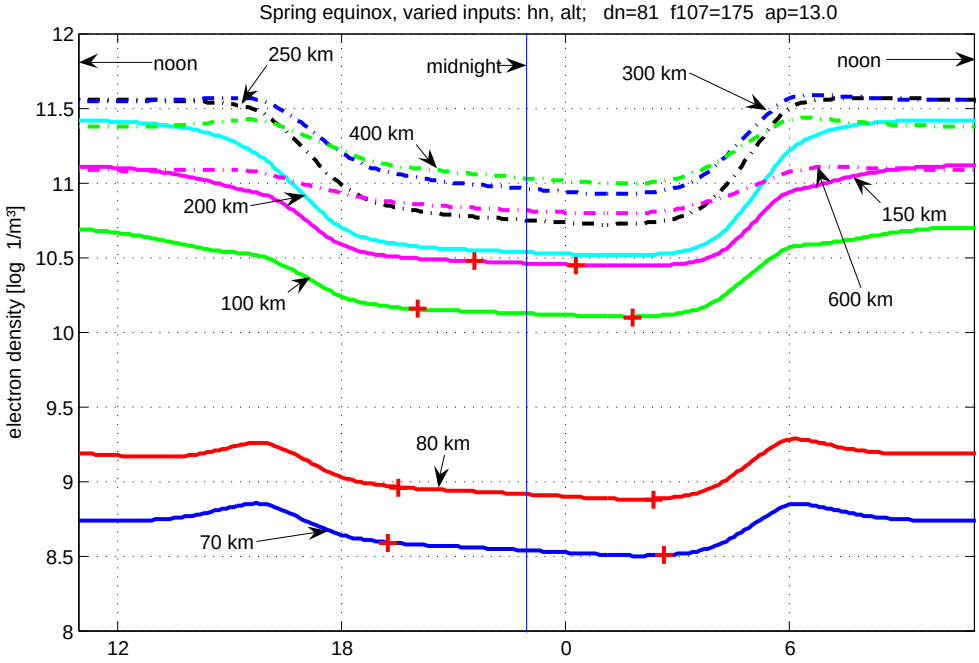


Figure 8.3 Svalbard, spring equinox, diurnal variations of N_e, hours [UTC], dummy data (model output). Red crosses indicate the day-night border according to Ch. 5.7.

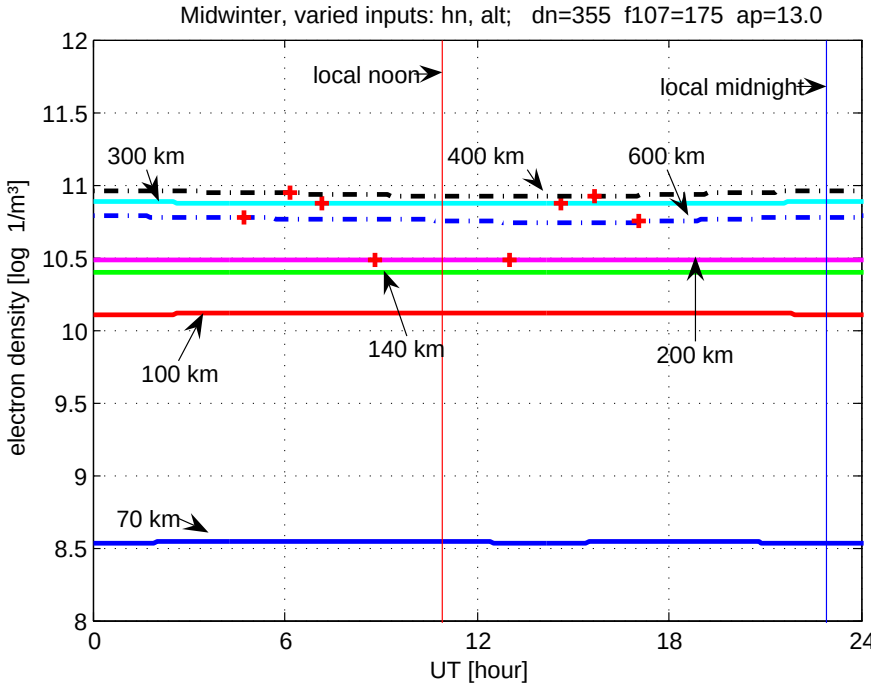


Figure 8.4 Svalbard, Midwinter, dummy data (model output). Obviously there is no diurnal variation due to full darkness for the whole day. Red crosses indicate the day-night border according to Ch. 5.7.

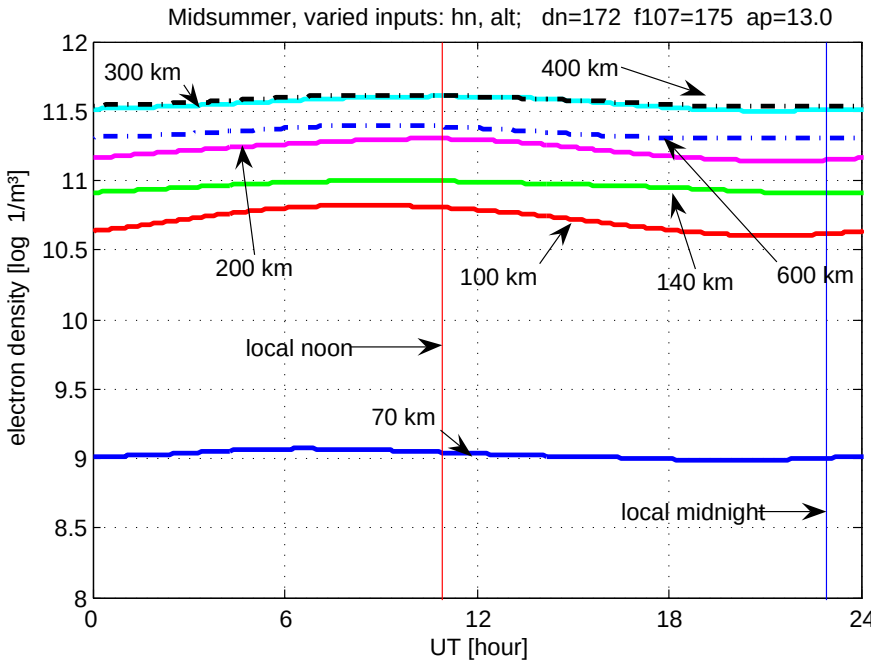


Figure 8.5 Svalbard, Midsummer, dummy data (model output).

The diurnal variation at midsummer in Fig. 8.5 is more pronounced than at mid-winter because it is always daylight and zenith angle changes between 50 and 80 degrees what can be seen in Figure 8.6 and 8.7.

The next variant is to vary the day of year to get an idea of seasonal variations in Svalbard. In Figure 8.6 and 8.7 electron densities for several altitudes at noon and midnight are plotted.

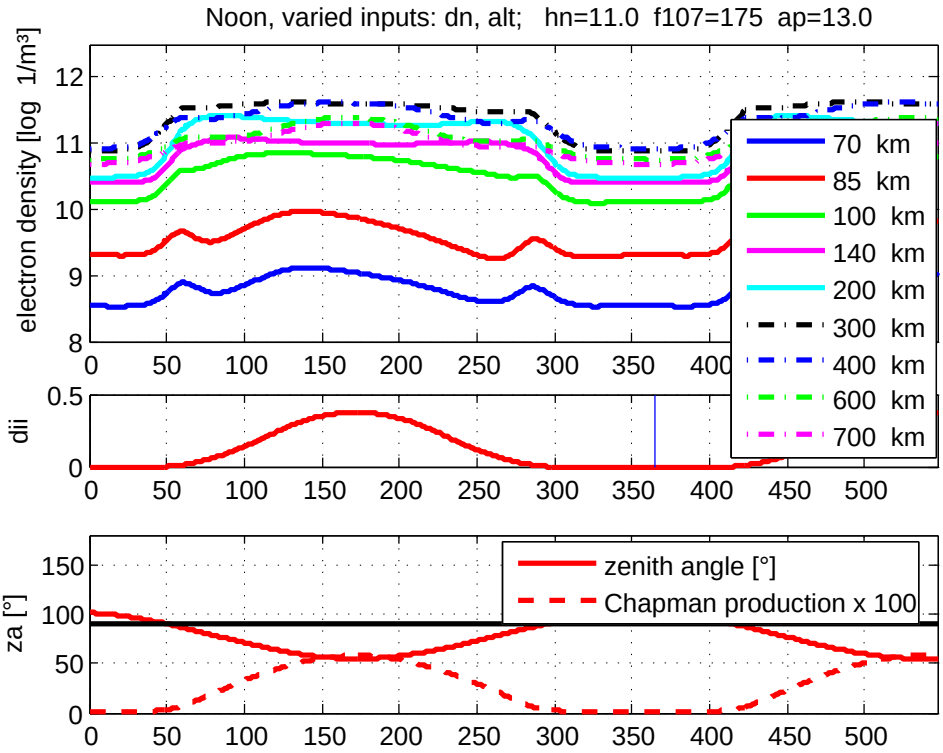


Figure 8.6 Svalbard, seasonal variation, Noon (model output).

Due to the latitude of 78° Svalbard has polar night and day which can clearly be seen in zenith angle characteristics and well as in the electron densities.

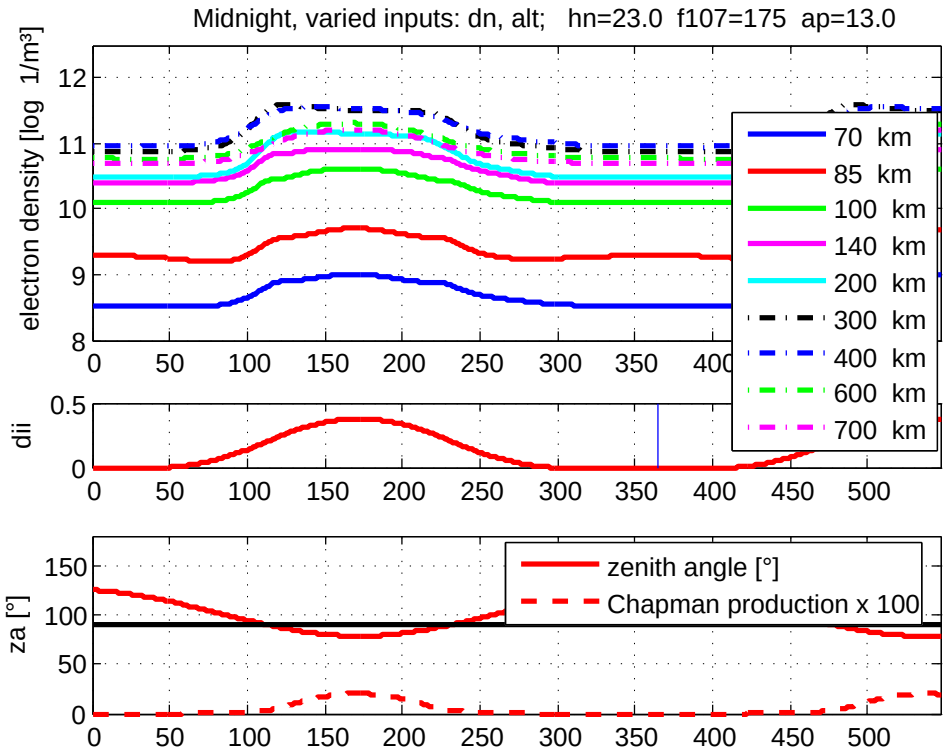


Figure 8.7 Svalbard, seasonal variation, Midnight (model output).

In Figures 8.8 and 8.9 electron density profiles are plotted for several solar fluxes (140, 175 and 200 Jy) at spring equinox, noon and midnight. Due to vertical magnetic field lines it is much easier for energetic particles to get into ionosphere.

Plots for midsummer, autumnal equinox and midwinter can be found in appendix starting at Figure 11.1.

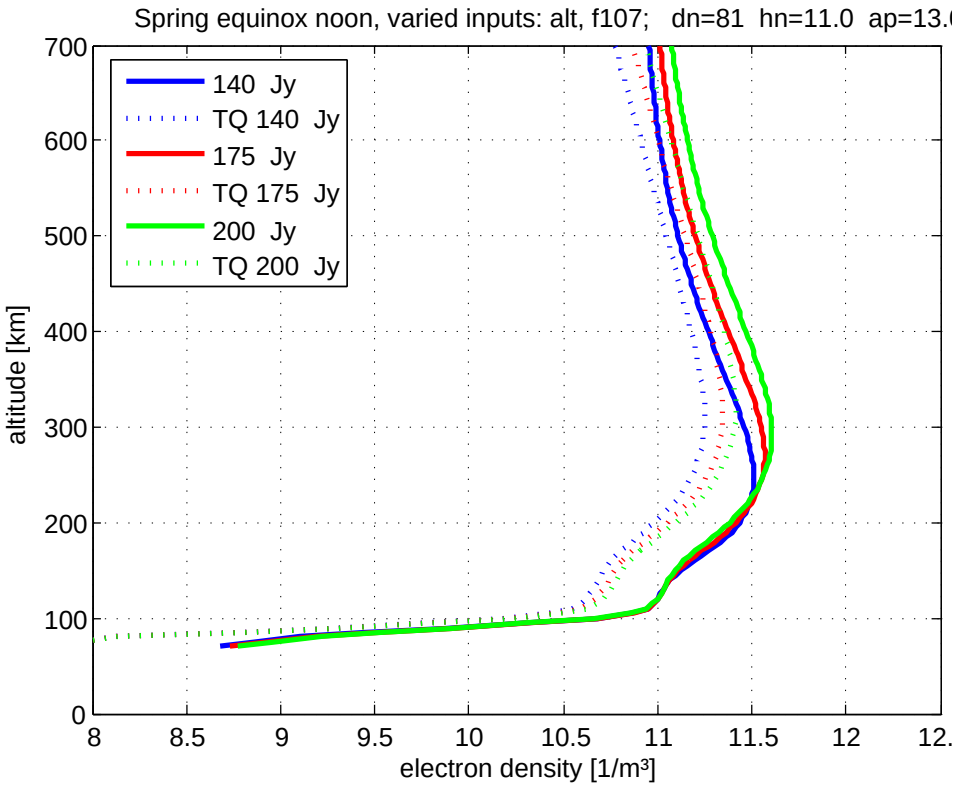


Figure 8.8 Svalbard, spring equinox noon (model output).

In Figures 8.8 and 8.9 the true quiet electron density is plotted (dotted lines). Further information about the true quiet electron density can be found in [Egger, (2004)].

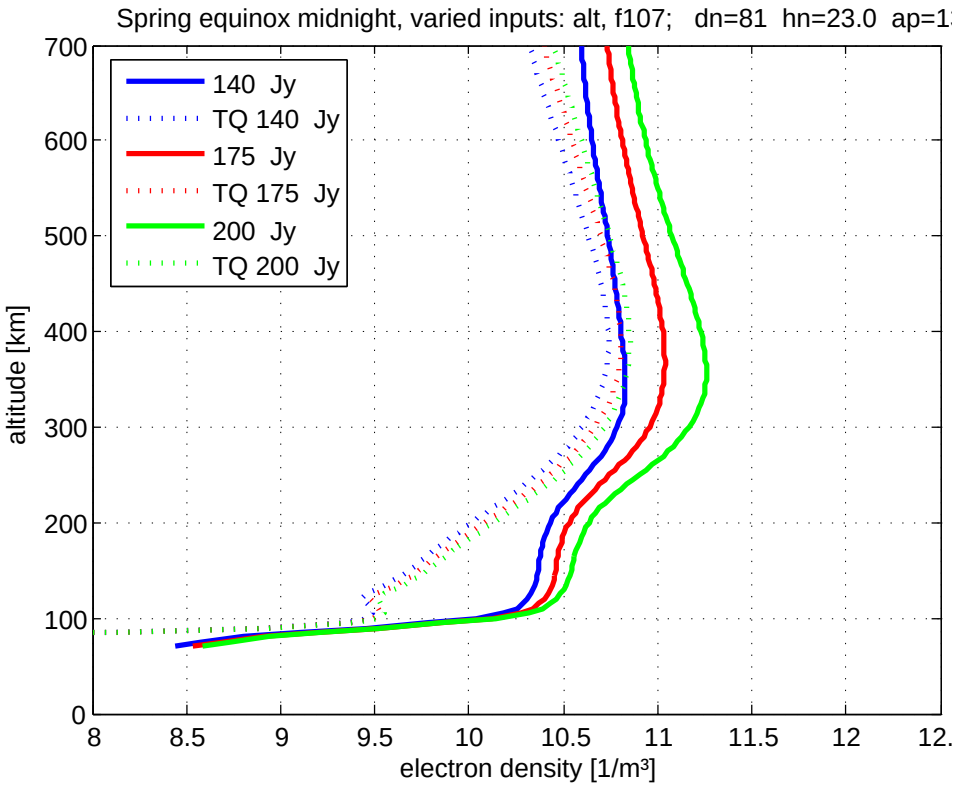


Figure 8.9 As Fig. 8.8, but for midnight.

The final parameter the geomagnetical index *ap* is varied. In Figures 8.10 and 8.11 several values of ap are plotted at noon and midnight for spring equinox. The influence of *ap* is very small only at midnight there is a little dependence near the F-region peak.

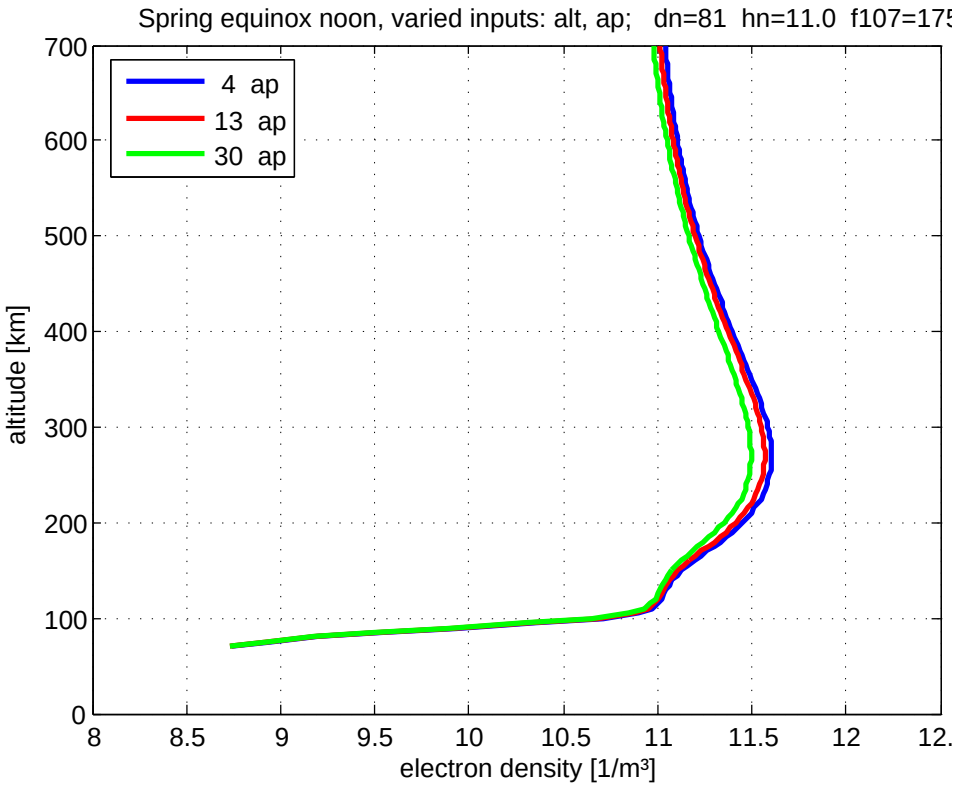


Figure 8.10 Svalbard, Spring equinox noon, dummy data (model output).

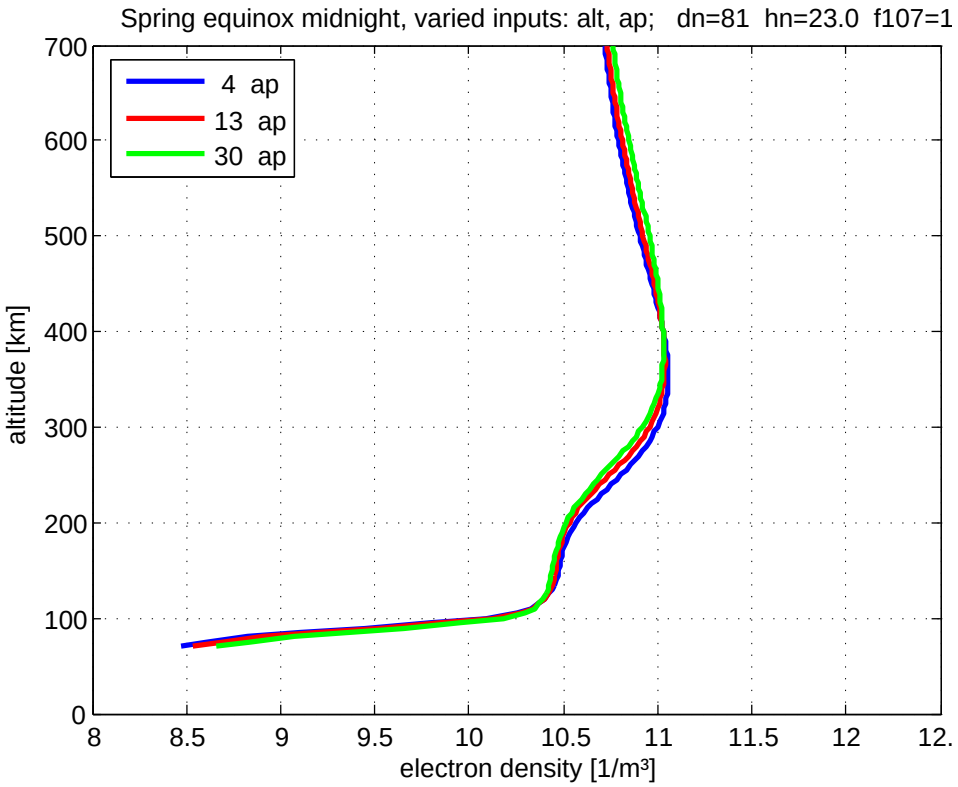


Figure 8.11 Svalbard, Spring equinox midnight, dummy data (model output).

8.3 Sunset Night Decay

In Figure 7.17 for every hour after sunset an electron density profile is plotted. The electron density decreases due to recombination. Note that the plot is produced from raw data.

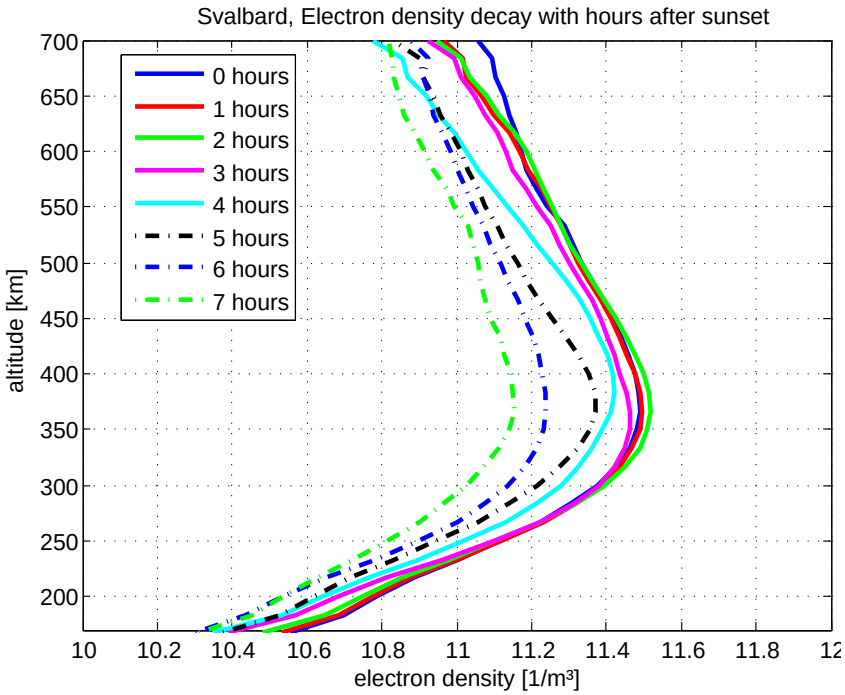


Figure 8.12 Svalbard, electron density decay with hour after sunset (raw data).

In Figure 8.14 and 8.13 the gradient of N_e is plotted as a function of altitude for hour after sunset and zenith angle. Note that the plots are produced from raw data.

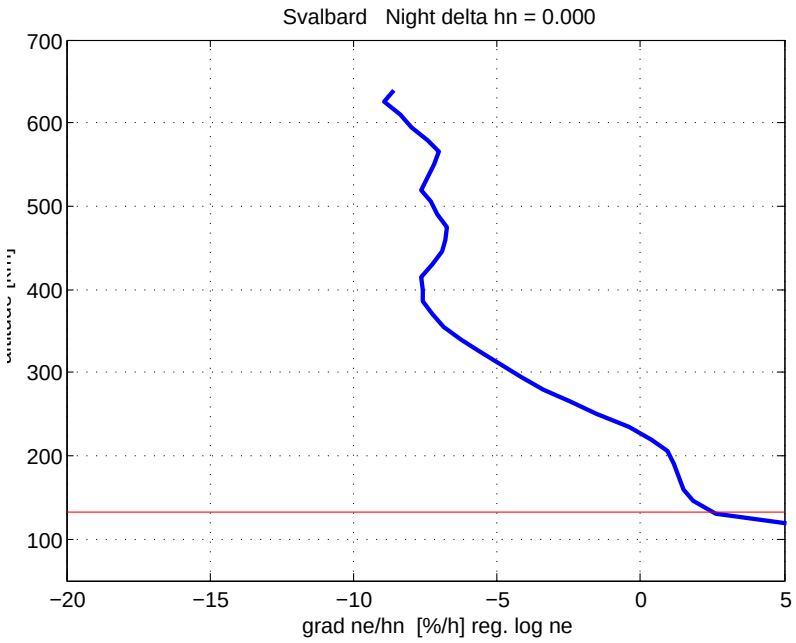


Figure 8.13 Gradient (N_e/hn) night(raw data). The red line indicates the lower quartile of the data distribution. Below data are not considered relevant.

The electron density decay is approximately only about half the value obtained for Arecibo (15 % per hour at higher altitudes).

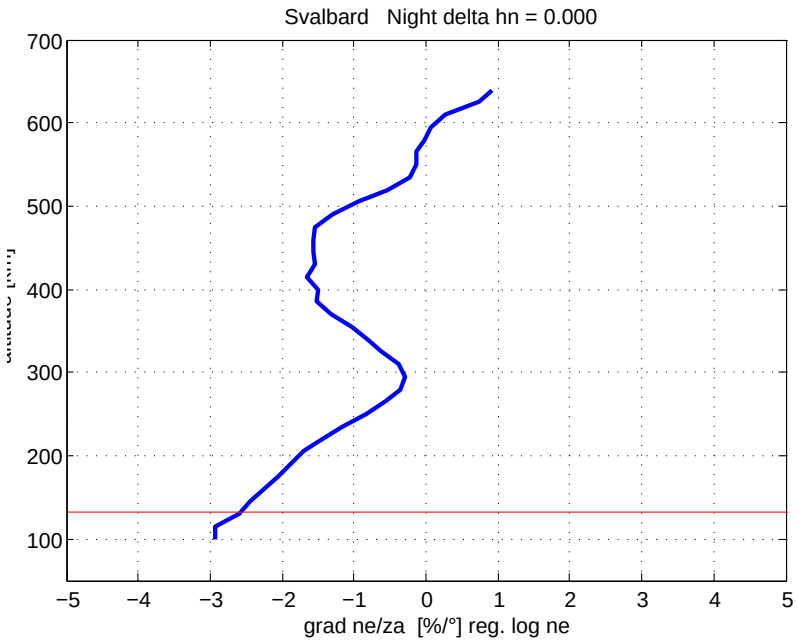
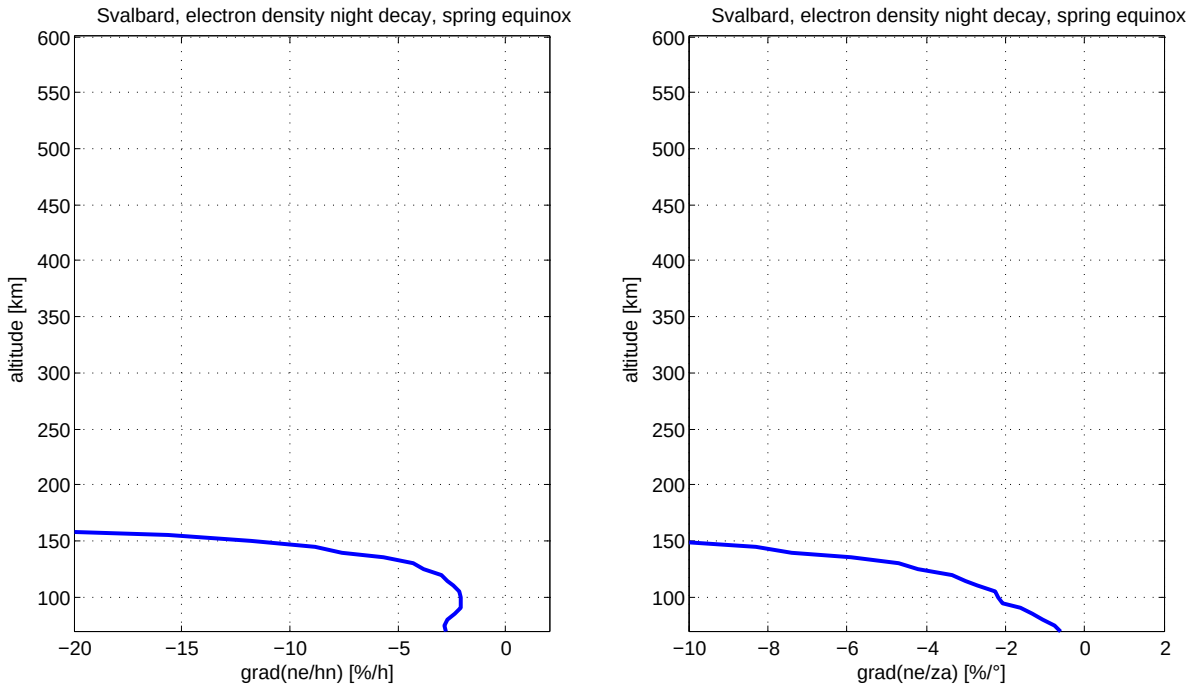


Figure 8.14 Gradient (N_e/za) night (raw data). The red line indicates the lower quartile of the data distribution. Below data are not considered relevant.



(a) Svalbard, spring equinox, night decay dependence on hour after sunset (model output).

(b) Svalbard, spring equinox, night decay dependence on zenith angle (model output).

Figure 8.15 Due to darkness for the whole day above 160 km no useful results can be produced.

8.4 Seasonal and diurnal variations

In Figures 8.16 to 8.21 contour plots for 200, 400 and 600 km are shown. Each plot shows the electron densities for all seasonal and diurnal conditions. The plot from original data includes all ap and $F107$ for the regarding altitude layer. The model output is at mean condition. On the right side an intersection is shown for midsummer (red line) and midwinter (blue line).In Figures 8.16 and 8.17 contourplots for 200 km are shown.

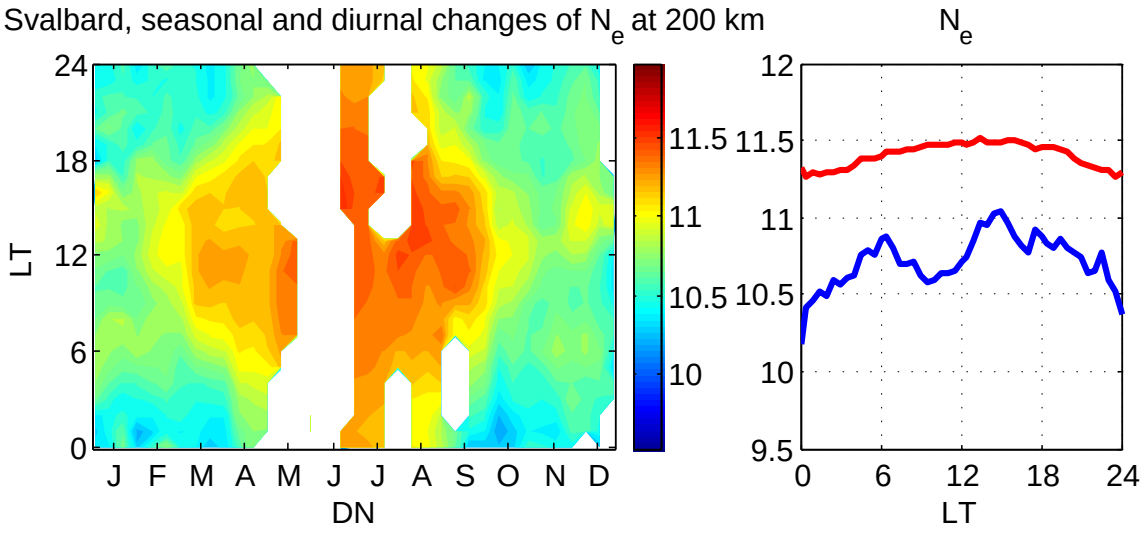


Figure 8.16 Svalbard, original data, 200 km. It seems that the data coverage in summer is not very good, which leads to an infirm model result for this season and height. Lines of constant electron densities (left) and a crossection for midsummer and midwinter (right).

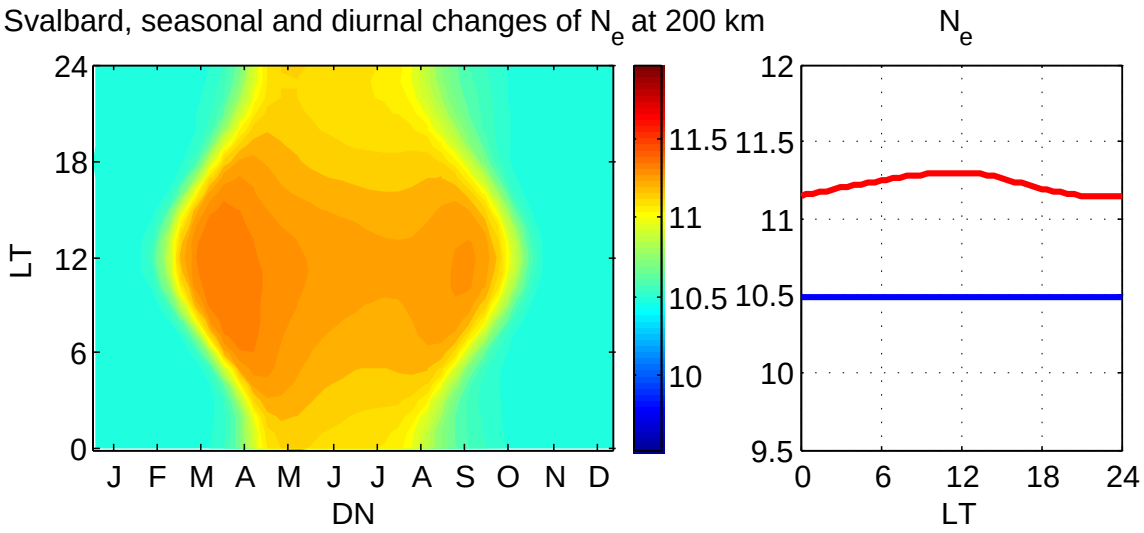


Figure 8.17 Svalbard, model data, 200 km. The shift of the peak can be explained with uneven data coverage. In the raw data plots all solar and geomagnetic conditions are included. The model output plot is produced at mean conditions.

In Figures 8.18 and 8.19 contourplots for 400 km are shown.

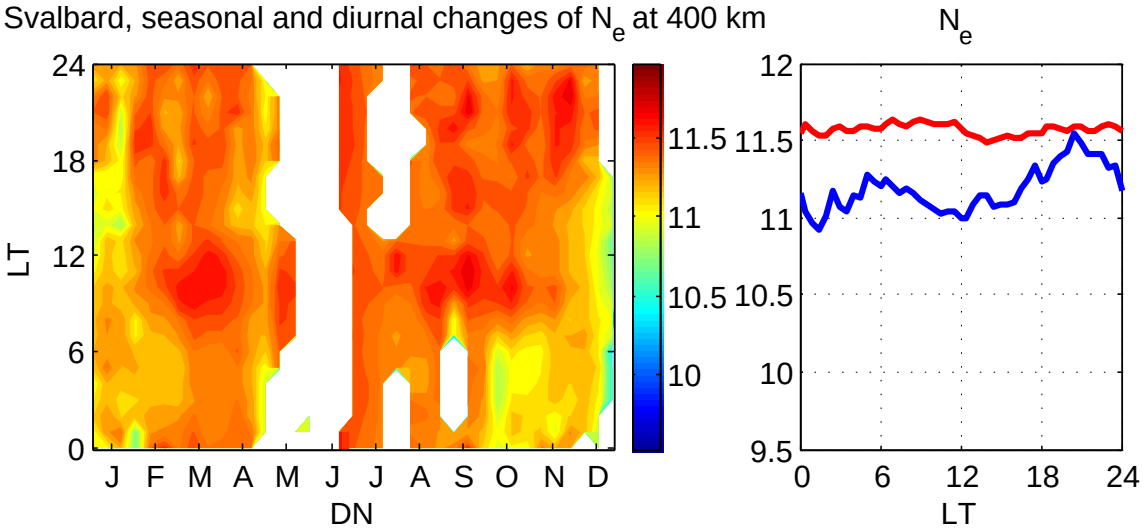


Figure 8.18 Svalbard, original data, 400 km. Lines of constant electron densities (left) and a crossection for midsummer and midwinter (right).

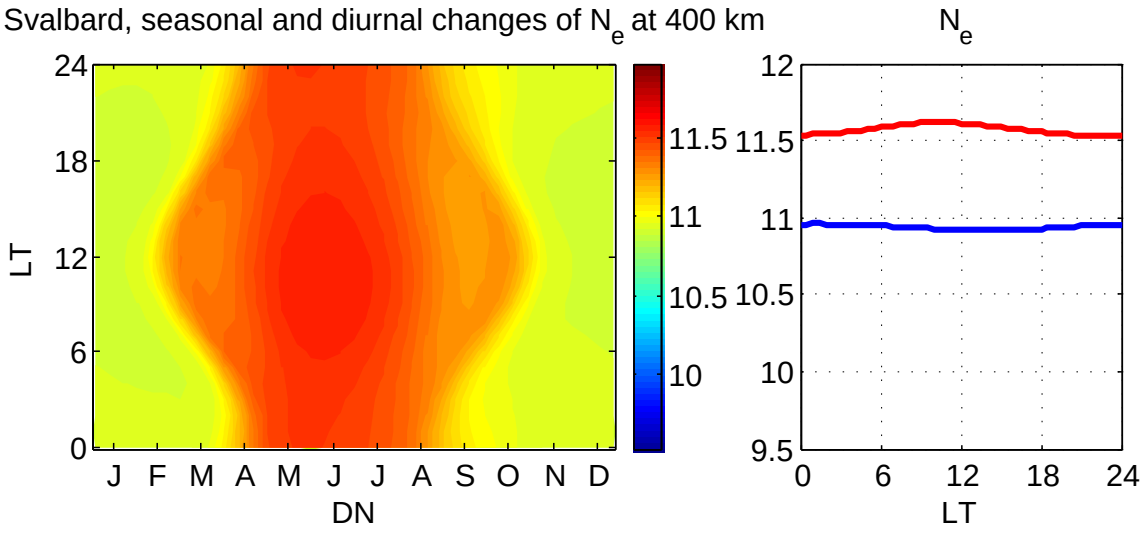


Figure 8.19 Svalbard, model data, 400 km.

In Figures 8.20 and 8.21 contourplots for 600 km are shown.

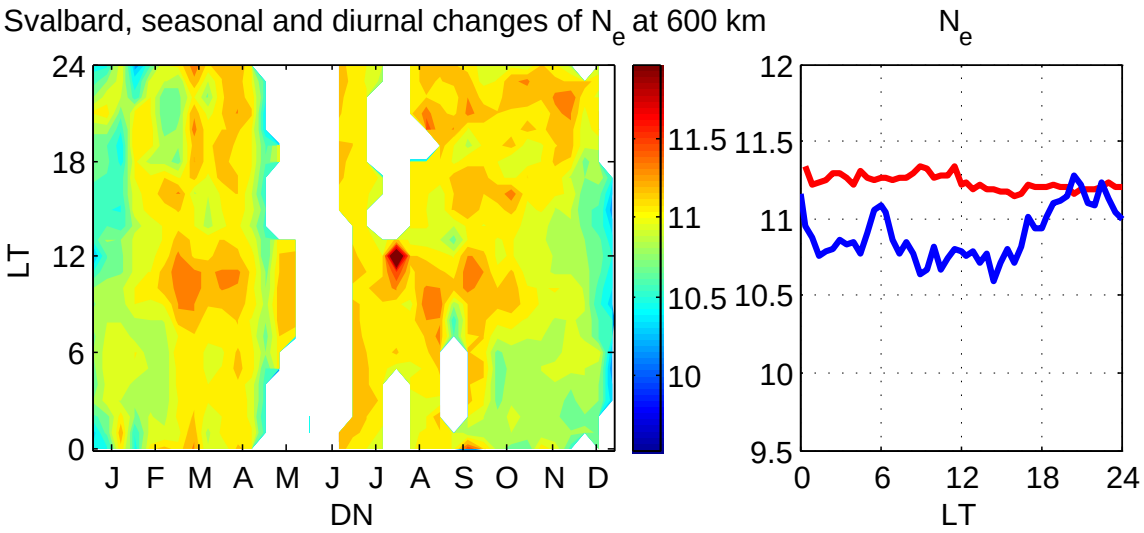


Figure 8.20 Svalbard, original data, 600 km. Lines of constant electron densities (left) and a crossection for midsummer and midwinter (right).

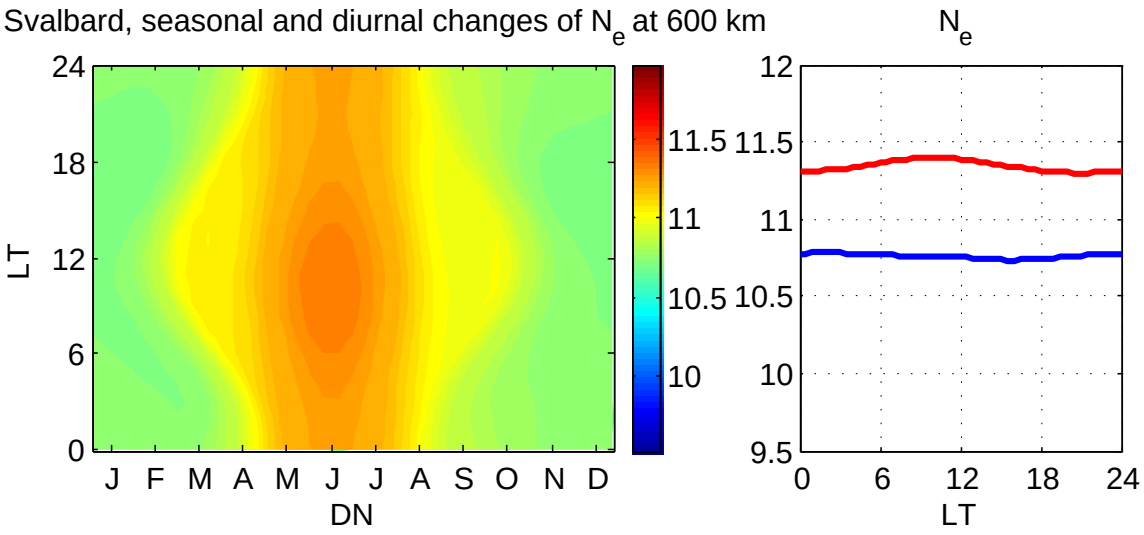


Figure 8.21 Svalbard, model data, 600 km.

In general at Svalbard the data coverage is not very good because most of the data is from few years with very uneven data distribution. Nevertheless the model produces useable predictions which corresponds well with seasonal and diurnal conditions like zenith angle. And at high latitude it is due to a disturbed ionosphere obvious that many data was measured under special conditions what makes it here difficult to compare model output and original data.

9 Conclusion

For both models (Arecibo and Svalbard) the same input parameters are used, which are listed in Table 9.1).

Altitude and pressure are used as input parameters because the combination of both produces the best RMS. Pressure is calculated using the MSIS model.

To get the geomagnetic influences into the model the geomagnetic index ap and an average value with 60 days preceding time (ap_{60}) are used as input parameter. 60 days does represent a pronounced optimum but was found to be a good compromise for all altitudes. In general both at Arecibo and Svalbard the geomagnetic dependence is much less than the dependence on solar flux. In general in Svalbard the geomagnetic (ap/kp, Dst) influence is larger than in Arecibo.

As input for the solar activity the 10.7 cm solar flux ($F107$) and an average value ($F107_{60}$) with 60 days preceding time are used. Again 60 days was found to be a good compromise.

For seasonal and diurnal variation of electron density it is important to find a way to "teach" the neural network a cyclic behaviour, for example the last day of year is 1 day before the first day of the year and the model should produce similar results.

For seasonal variation the Daily Integrated Insolation (dii) and a 30 day averaged (dii_{30}) value are used as input parameter.

To model diurnal variation the Chapman production (ch) and an 11 hour averaged value (ch_{11}) are used as parameter for the neural network. Two inputs are used to make it possible for the neural network to "learn" asymmetric characteristics.

To model the diurnal and seasonal variation alternatively sine and cosine functions may be used which produce good results as exercised by [McKinnell, (2002)]. Here the Chapman production and Daily Integrated Insolation is used due to a better correlation with electron density of these measures. Further details can be found in Ch. 6 Development.

Parameter	Description
alt	Altitude
$logp$	Logarithm. air pressure
ap	Geomagnetic index ap
ap_{60}	Geomagnetic index ap, averaged, 60 days preceding time
$F107$	10.7 cm solar flux
$F107_{60}$	10.7 cm solar flux, averaged, 60 days preceding time
dii	Daily integrated insolation (seasonal variations)
dii_{30}	Daily integrated insolation, averaged, 30 days preceding time
ch	Chapman production (diurnal variations)
ch_{11}	Chapman production, averaged, 11 hours preceding time

Table 9.1 Input parameters for the neural networks.

In Table 9.2 the models for Arecibo and Svalbard are compared. Further details can be found in Ch. 4 Datasources. The data for both models come from ISR's (Incoherent Scatter Radar). An other comparison of two datasets (EISCAT and Millstone Hill) was done bei [Lei et al.].

The data for the Svalbard model is augmented by data from 372 rocket flights from Heiss Island (local time is corrected for Svalbard) and one rocket flight from Svalbard.

The electron density profiles for the Arecibo model are extended by approximately 20 % FIRI data (if minimum altitude of a profile reaches below 160 km) to increase stability at low altitude.

For the Arecibo model a neural network with 18 hidden nodes (3 hidden layers, 6 nodes each) is used. This architecture was found as optimum. For the Svalbard model this type of neural network produces poor results due to a more disturbed ionosphere and greater range of altitude. Therefore a neural network with 30 hidden nodes (5 hidden layers, 6 nodes each) is used for the Svalbard model.

Criteria	Arecibo	Svalbard
latitude	18.2°	78.2°
data source	ISR	ISR + rocket data (372)
data points	1,585,824 +20 % FIRI	1,459,262
profiles	98,554	162,507
hidden nodes	18	30
hidden layers	3	5
data time range	1966 to 2002	1996 to 2004
		+ rocket flights since 1979 (below 100 km)
data distribution	sporadic, within 36 years	sporadic
		most data within 8 years
altitude range	130 km to 700 km	60 km to 700 km
RMS error factor	2.09	2.21

Table 9.2 Comparison of Arecibo and Svalbard model.

The RMS error factors (model output is compared with all measured data) are 2.09 (Arecibo) and 2.21 (Svalbard) which is usable for predictions and is compareable to other models as IMAZ (RMS = 2.26, [McKinnell and Friedrich, (2007)]) or IRI (RMS = 2.20, when compared with the Arecibo data).

Due to a more disturbed ionosphere at higher latitude and a greater altitude range the RMS error factor is slightly higher. Probably the RMS difference would be greater if the measured data from Arecibo would have a constant good quality like data from the 80's or younger (shown in Ch. 6.1.3).

The Arecibo model produces good results from 130 to about 700 km due to less coverage at low altitude. The Svalbard model produces good results from 80 to about 700 km due to rocket data down to 60 km. Due to the uneven data coverage of the first years of operation used here more data over a longer time range would be needed to improve this model. Notably low solar activity conditions currently are only poorly covered.

10 Table of variables and abbreviations

Name	Description
η	Learning rate of a neural network
χ	Zenith angle,
alt	Altitude
ap	Geomagnetic index ap
$ch, chza$	Chapman production
d_{max}	Maximum tolerated difference between input and output
Dst	Geomagnetic index dst
$F107$	10.7 cm solar flux
I	The angle of the wave relative to the horizon
kp	Geomagnetic index kp
N_e	Electron density $[m^{-3}]$
IMAZ	Ionosphere Model for Auroral Zone
IRI	International Reference Ionosphere
EISCAT	European Incoherent Scatter Radar
NAIC	National Astronomy and Ionosphere Center
NOAA	National Oceanic and Atmospheric Administration
PCA	Polar Cap Absorption
RMS	Root Mean Square
SNNS	Stuttgart Neural Network Simulator
SSE	Summed Squared Error

Table 10.1 Used variables and abbreviations.

11 Appendix

In appendix the most important Matlab functions are explained and additional plots are shown.

11.1 Matlab Software Architecture

In this section it is briefly desrcibed how the Matlab (R14) software works and what its main functions do.

The training of the neural network itself is done with SNNS due to its much better performance concerning speed and ability to process large amounts of training data. But all process steps around are done with MATLAB. At first all model parameters are adjusted in d_indunet (11.1.1). If we run this script all process steps are done auomatically, the filter file is written(11.1.2), the pattern files are written and the SNNS batch file is generated and started.

The script d_write_pattern saves a MATLAB matrix in a SNNS pattern file format. These files are read by SNNS batch processor.

Every neural network is stored in an own directory tree, 5 different networks are trained and after that compiled together with a MEX File to get an MATLAB function. With this MATLAB function it is possible to directly produce predictions by passing normalized input parameters to the function. At the end the d_indunet script evaluates the neural network by comparing all datapoints with the predictions.

In another script called d_rms_batch all dummy plots are produced by fixing most input parameters and vary just the parameters which should be observed.

11.1.1 d_indunet

In the script d_indunet_ar respectively d_indunet_sh all parameters for a new model(network) are to be edited.

11.1.2 d_writefilter

The function d_writefilter contains all values which are important for filtering and norming input and output data for the neural network.

All values are stored in a matrix, in the function standard values are coded which can be edited by parameters of the function. So in d indunet filter values can be edited, for example data should be filtered for datapointswhere altitude is lower than 170 km.

11.2 Additional Plots

11.2.1 Arecibo

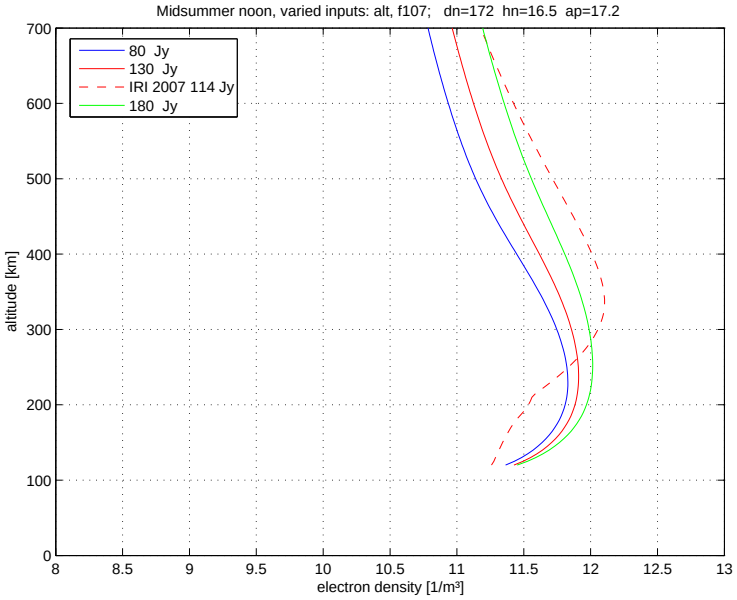


Figure 11.1 Arecibo, Midsummer noon, dummy data (model output).

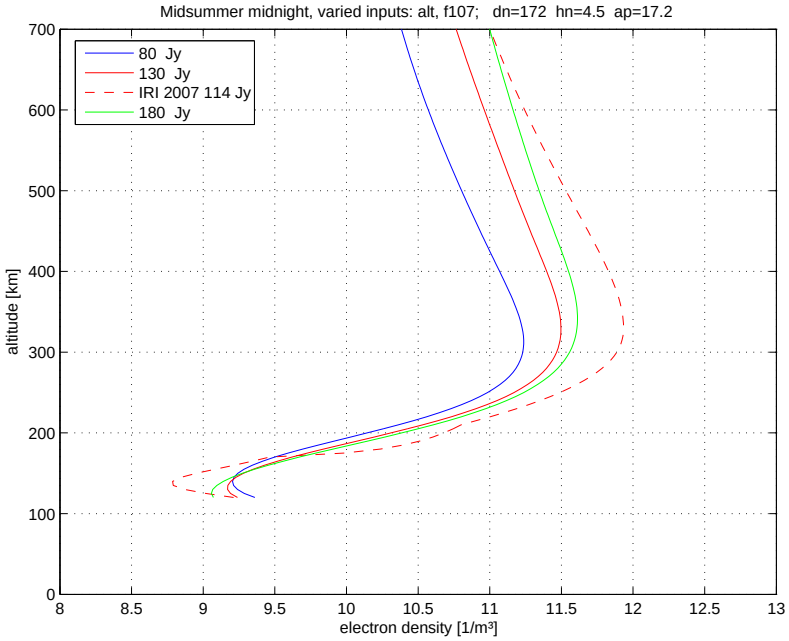


Figure 11.2 Arecibo, Midsummer midnight, dummy data (model output).

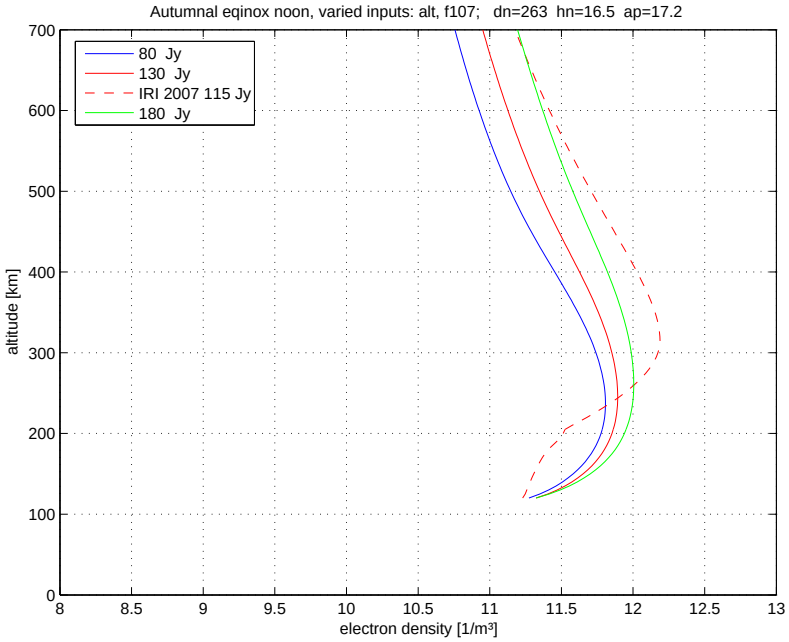


Figure 11.3 Arecibo, Autumnal eqinox noon, dummy data (model output).

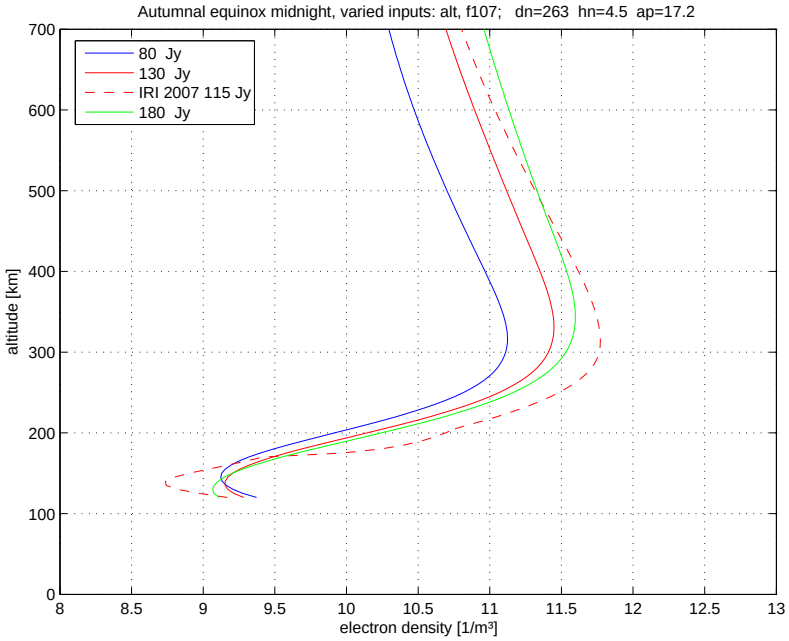


Figure 11.4 Arecibo, Autumnal equinox midnight, dummy data (model output).

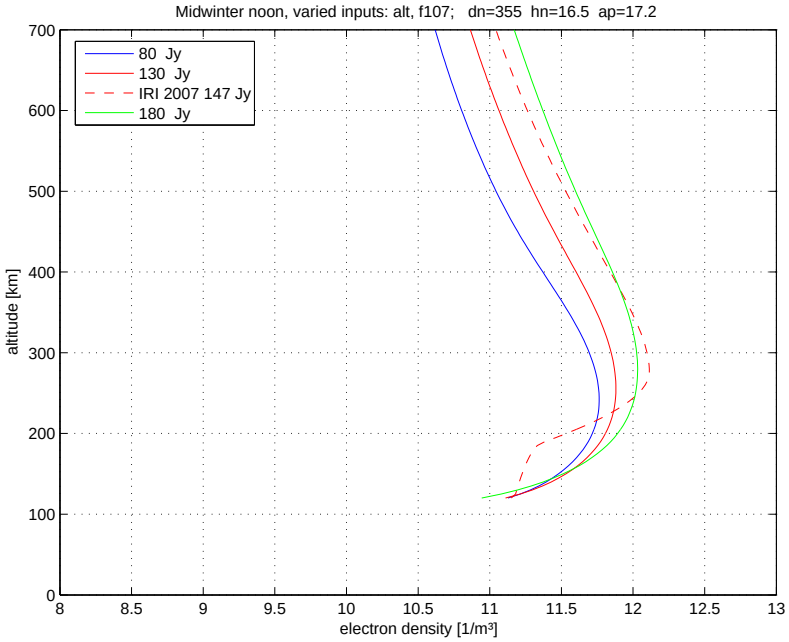


Figure 11.5 Arecibo, Midwinter noon, dummy data (model output).

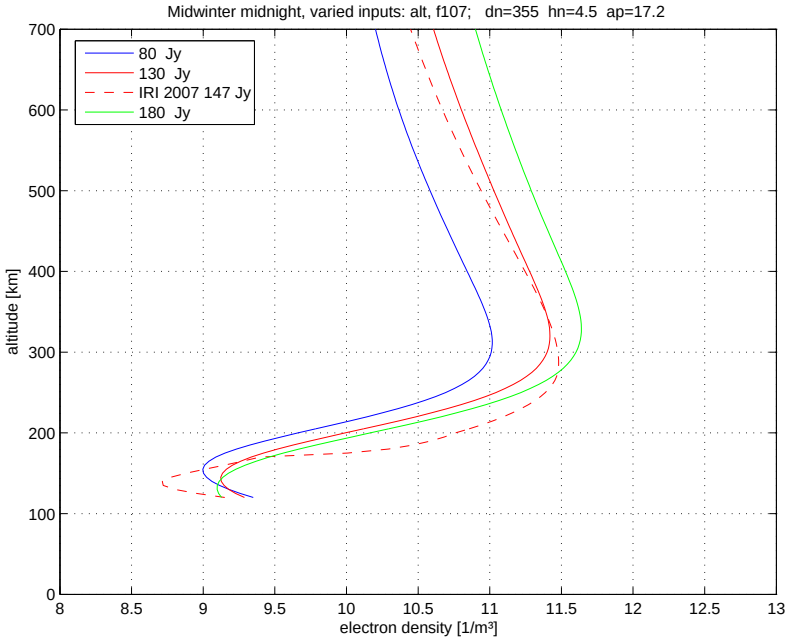


Figure 11.6 Arecibo, Midwinter midnight, dummy data (model output).

11.2.2 Svalbard

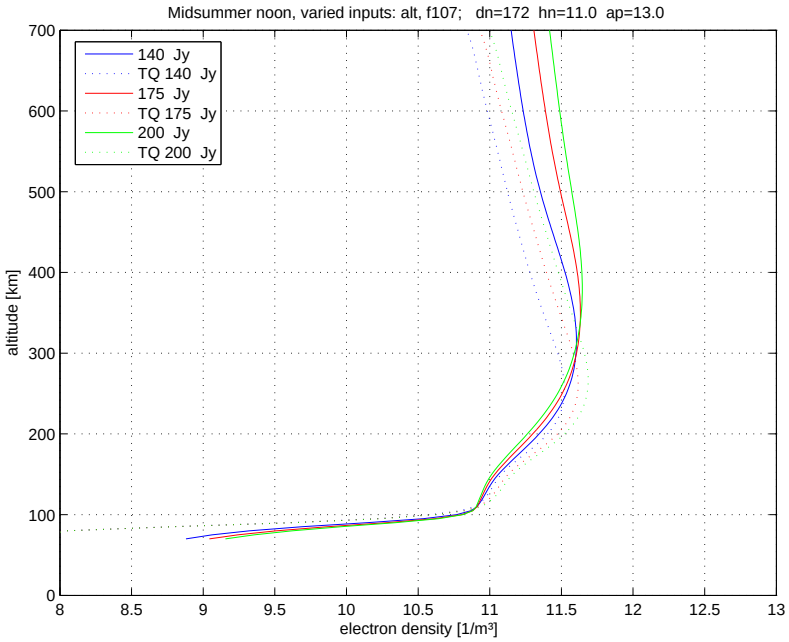


Figure 11.7 Svalbard, Midsummer noon, dummy data (model output).

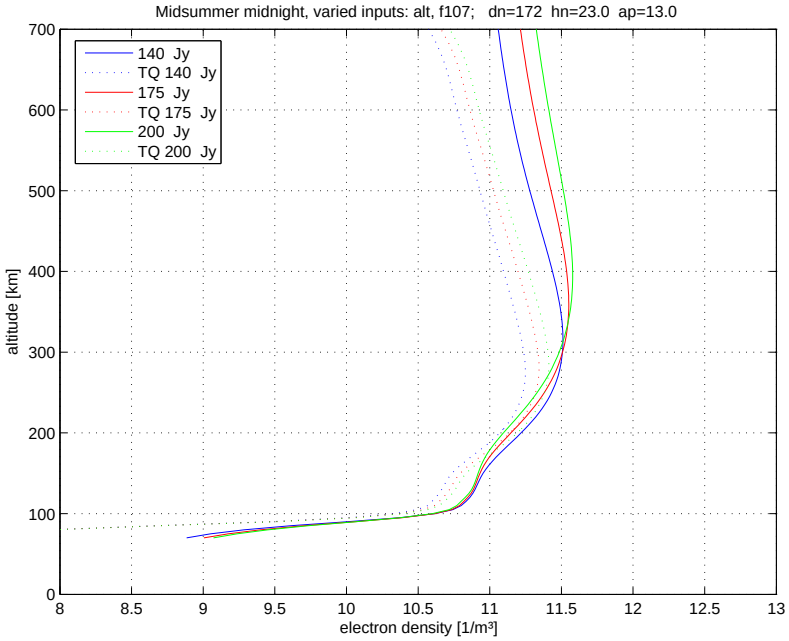


Figure 11.8 Svalbard, Midsummer midnight, dummy data (model output).

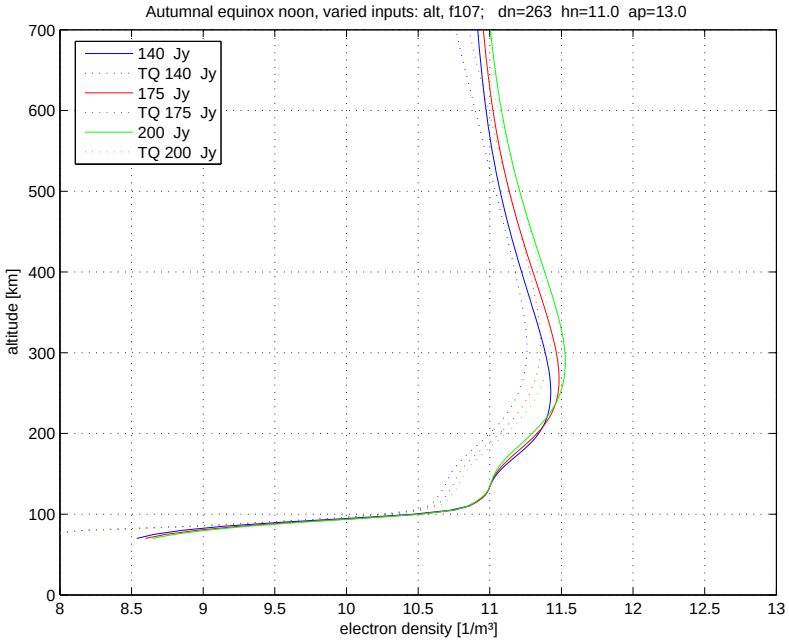


Figure 11.9 Svalbard, Autumnal equinox noon, dummy data (model output).

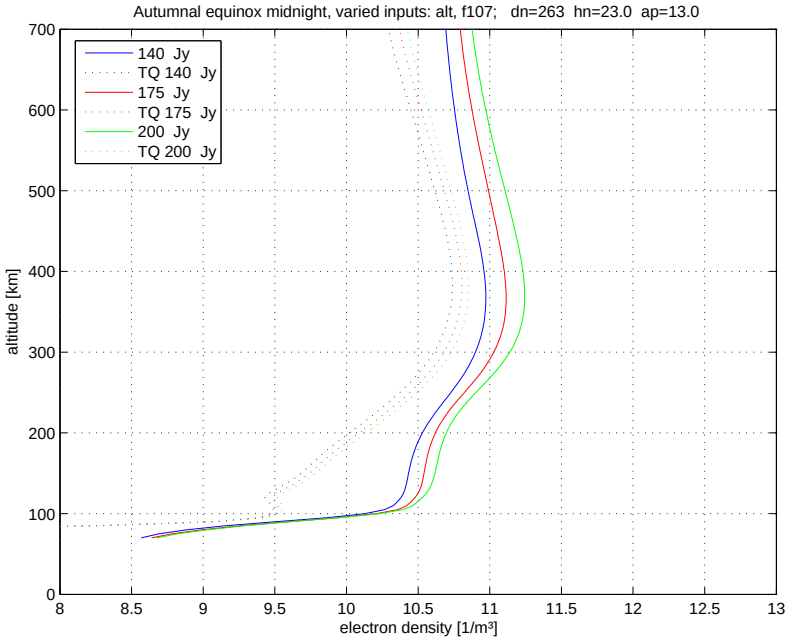


Figure 11.10 Svalbard, Autumnal equinox midnight, dummy data (model output).

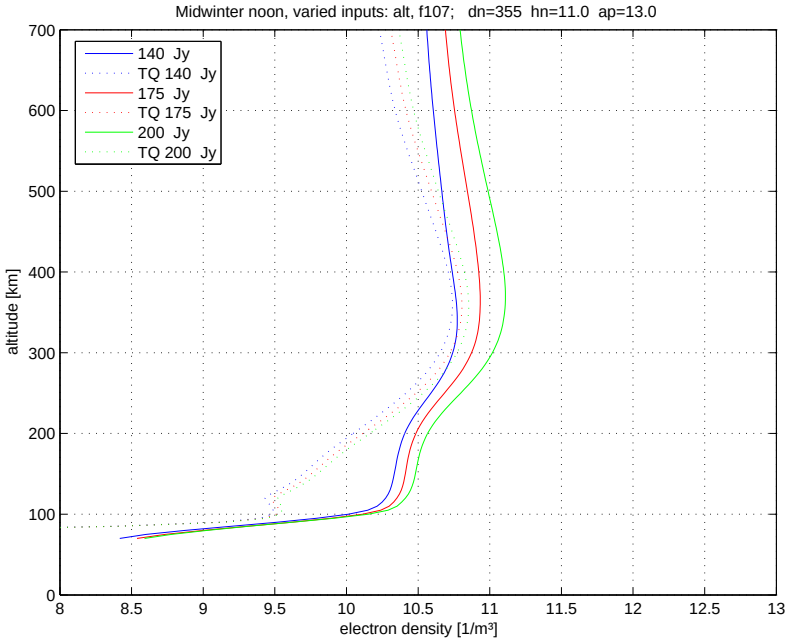


Figure 11.11 Svalbard, Midwinter noon, dummy data (model output).

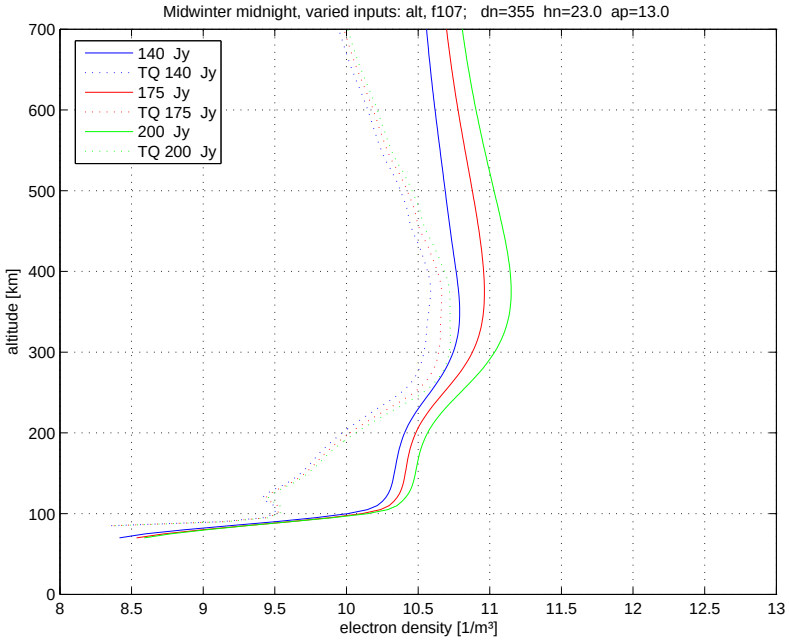


Figure 11.12 Svalbard, Midwinter midnight, dummy data (model output).

List of Figures

List of Tables

Bibliography

[Bilitza, (2009)] D. Bilitza: Evaluation of the IRI-2007 model options for the topside electron density, Advances in Space Research, Volume 44, Issue 6, Ionosphere - Modelling, Forecasting, and Telecommunications II, Pages 701-706, 2009.

[Boston University] Boston University, http://www.bu.edu/cism/CISM_Thrusts/ionosphere-atmosphere.html

[Danilov et al. (2003)] A. D. Danilov, N. V. Smirnova, T. A. Blix, E. V. Thrane, L. B. Vanina: Some features of electron density behaviour in the high latitude D-region derived from in situ measurements, Journal of Atmospheric and Solar-Terrestrial Physics, Volume 65, Issue 4, Pages 417-427, 2003.

[Callan, (2003)] R. Callan: Neuronale Netze im Klartext, Pearson Studium, Prentice Hall, 2003.

[Egger, (2004)] G. Egger: Empirical Model of the Polar Cap Ionosphere, pp. 38 + 66, TU Graz, Department of Communications and Wave Propagation, 2004.

[EISCAT] Geographical and geomagnetic positions of the EISCAT facilities, EISCAT, https://e7.eiscat.se/locations.html, 2010.

[Friedrich et al. (2007)] M. Friedrich, M. Fankhauser, E. Oyeyemi, L. A. McKinnell: A neural network-based ionospheric model for Arecibo, Advances in Space Research 42 776-781, 2007.

[Friedrich and Torkar, (2001)] M. Friedrich and K.M. Torkar: FIRI: A Semiempirical Model of the Lower Ionosphere, Journal of Geophysical Research, 106, No. A10, pp. 21,409-21,418, 2001.

[Hargreaves, (1992)] J. K. Hargreaves: The Solar-Terrestrial Environment. Cambridge University Press, 1992.

[Hargreaves and Hunsucker, (2003)] R. D. Hunsucker and J. K. Hargreaves: The High-Latitude Ionosphere and its Effects on Radio Propagation, Cambridge Atmospheric and Space Science. Series, 2003

[Harrich, (2001)] M. Harrich: Empirical Modelling of Electron Densities in the High Latitude Mesosphere, Phd Thesis, Department of Communications and Wave Propagation, TU Graz, 2001.

[Isaksen, (1973)] I.S.A. Isaksen: Diurnal Variation of Atmospheric Constituents in an Oxygen-Hydrogen-Nitrogen-Carbon Atmospheric Model, and the Role of Minor Neutral Constituents of the Lower Ionosphere, Geophys. Publ., 30,pp. 1-63, 1973.

[Jakowski et al. (2002)] N. Jakowski, S. Heise, A. Wehrenpfennig, S. Schlueter, R. Reimer: Journal of Atmospheric and Solar-Terrestrial Physics, Volume 64, Issues 5-6, Pages 729-735, 2002.

[Lei et al.] J. Lei, L. Liu, W. Wan, S-R. Zhang and A.P. Van Eyken: Comparison of the first long-duration IS experiment measurements over Millstone Hill and EISCAT Svalbard radar with IRI2001.

[Knight, (1972)] P. Knight: A classification of night-time electron-density profile, Journal of Atmospheric and Terrestrial Physics, Vol. 94, pp. 401-410, 2004.

[Lassen, (1926)] H. Lassen: Zeitschrift fuer Hochfrequenztechnik, Volume 28, pp. 109-113, 1926.

[Minow, (2004)] J.I. Minow: Development and implementation of an empirical ionosphere variability model, Advances in Space Research, Vol. 33, 2004.

[McKinnell, (2002)] L.A. McKinnell: A neural network based ionospheric model for the bottomside electron density profile over Grahamstown, South Africa, Phd Thesis, Rhodes University, 2002.

[McK04] L.A. McKinnell, M. Friedrich: Results from a new auroral lower ionosphere model, Advances in Space Research, 2004.

[McKinnell and Friedrich, (2007)] L.A. McKinnell, M. Friedrich: A neural network-based ionospheric model for the auroral zone, Journal of Atmospheric and Solar-Terrestrial Physics, Volume 69, Issue 12, Pages 1459-1470, 2007.

[McKinnell, (2008)] L.A. McKinnell: Development and implementation of an empirical ionosphere variability model. Advances in Space Research, 2008.

[NAIC 2010] National Aeronomy and Ionosphere Center, Arecibo Observatory, Puerto Rico, http://www.naic.edu, 2010.

[NASA Goddard] NASA Goddard Space Flight Center, MSIS http://omniweb.gsfc.nasa.gov/vitmo/msis_vitmo.html, 2010.

[Oyeyemi, (2005)] E.O. Oyeyemi, A.W.V. Poole, L.A. McKinnell: On the global model for foF2 using neural networks, Radio Science 40 doi:10.1029/2004RD003223, 2005.

[Schlegel, (1984)] K. Schlegel: HF and VHF Coherent Radars for Investigation of the High-latitude Ionosphere. May Planck Institut für Aeronomie, Katlenburg-Lindau, 1984.

[Schlegel, (1995)] K. Schlegel: EISCAT and the EISCAT data base – A tool for ionospheric modeling (E- and D-region), Advances in Space Research, Volume 16, Issue 1, The High Latitudes in the International Reference Ionosphere, Pages 147-154, 1995.

[Titheridge, (1988)] J. E. Titheridge: An Approximate Form for the Chapman Grazing Incidence Function, Journal of Atmospheric and Terrestrial Physics, 50, No. 8, pp. 699-701, 1988.

[USGS] USGS, http://geomag.usgs.gov/

[USGS, (2005)] USGS Geomagnetic indices. Technical report, US Geological Survey, http://geomag.usgs.gov, 2005.